Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Hannah K. Ratner Date

Functional Versatility of CRISPR-Cas Systems in a Bacterial Pathogen

By

Hannah Kwe Ratner
Doctor of Philosophy

Graduate Division of Biological and Biomedical Science
Microbiology and Molecular Genetics

David S. Weiss
Advisor

Graeme Conn
Committee Member

Charles Moran
Committee Member

Phil Rather
Committee Member

William Shafer
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Functional Versatility of CRISPR-Cas Systems in a Bacterial Pathogen

By

Hannah Kwe Ratner
B.S., University of Puget Sound, 2011

Advisor: David S. Weiss, PhD

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Microbiology and Molecular Genetics
2019

Abstract

Functional Versatility of CRISPR-Cas Systems in a Bacterial Pathogen
By Hannah Kwe Ratner

CRISPR-Cas systems are widespread in prokaryotes and function as adaptive immune systems that use small RNAs (CRISPR RNAs; crRNAs) to guide Cas proteins to recognize and cleave foreign nucleic acids. There is increasing evidence for broader roles of these systems. The bacterial pathogen *Francisella novicida* naturally encodes a CRISPR-Cas9 system that plays a critical role in bacterial virulence. We determined that Cas9 enables virulence by repressing the transcription of four endogenous genes and providing protection from phagosome produced antimicrobials. This regulation is mediated by a non-canonical small RNA (scaRNA), rather than a crRNA, that guides Cas9 to bind DNA targets and block transcription. scaRNA binds endogenous targets without lethally cleaving the bacterial chromosome due to reduced scaRNA:DNA complementarity. We harnessed this activity to reprogram scaRNA to repress other genes. Furthermore, with engineered, extended complementarity to an exogenous target, the repurposed scaRNA:tracrRNA-FnoCas9 machinery can also be licensed to direct cleavage of the invading DNA. These findings highlight that a cleavage-competent Cas9 complex can exist in two distinct functional states in bacteria: binding to endogenous DNA as a transcriptional repressor and cleaving foreign DNA to prevent infection. With this knowledge, we investigated Cas12a activity, a second CRISPR system in *F. novicida*. Despite the differences in timing of expression, Cas12a prevented infection with foreign DNA as efficiently as Cas9 when both systems were reprogrammed to prevent infection by the same DNA target. Synthesizing the requirements for DNA targeting and scaRNA-based repression, we demonstrated that the crRNAs of the *F. novicida* CRISPR-Cas9 and CRISPR-Cas12a systems can be commandeered to direct transcriptional repression in addition to DNA targeting, ultimately reprogramming Cas12a to repress endogenous targets. The functional versatility of *F. novicida* Cas9 and Cas12a indicates that subtle, single base changes in the crRNAs can direct the mechanics of CRISPR protein function. The shift between DNA targeting and transcriptional repression via DNA binding likely underpins a broad class of underappreciated CRISPR functions, one that is potentially critical to the physiology of the numerous Cas9- and Cas12a-encoding pathogenic and commensal organisms.

Functional Versatility of CRISPR-Cas Systems in a Bacterial Pathogen

By

Hannah Kwe Ratner
B.S., University of Puget Sound, 2011

Advisor: David S. Weiss, PhD

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Microbiology and Molecular Genetics
2019

Acknowledgements

I would like to dedicate this thesis to my mother, Judie Ratner, who sees the world with spectacular curiosity, excitement, and joy. Her fascination for the natural world has always inspired me. I am grateful for her constant love and support, the opportunities she has given me, and for trusting me to follow a path that may not have always made sense. More than anything, I am thankful for all the times she forced me out of the lab by devising non-negotiable adventures.

This thesis would not have been possible without the professional support of my mentors, colleagues, and labmates. My advisor, Dr. David Weiss, has shaped me into the scientist that I am today. It has been a privilege to work for someone who so thoughtfully and generously shares their knowledge and time, to help their students become better scientists, thinkers, writers and people. Thank you for giving me the space to be persistent, pushing me to go one step further, and going above and beyond to make sure that I could realize my goals. Thank you to the Weiss lab (past and present) for the exceptional scientific input, major support, hall beers, and for sharing their skills and knowledge so readily. Everyone in the lab has contributed to my PhD in critical and distinct ways, and made my time in graduate school a blast. In particular I would like to thank Drs. Chui Yoke Chin, Tim Sampson, and Emily Crispell, who each took time to train me and/or provided important scientific insight at different stages of my thesis projects. To our collaborators in the Charpentier lab, Drs. Emmanuelle Charpentier, Anaïs Le Rhun and Andrés Escalera-Maurer, thank you for your brilliant scientific contributions, feedback, and expertise, all of which I learned a lot from and were essential for the success of these projects. I am grateful to my committee for their feedback and support, especially Charlie Moran and Phil Rather, who frequently provided insight on my projects. Lastly, thank you to all of my former mentors and colleagues who continue to support my career, especially Bill Jacobs for convincing me to go to grad school by showing me that career in science can be awesome, and equally to Alex Stone, for telling me the truth about this path and advising me to turn back before I got hooked.

The love and support that my family and friends have cocooned me with, especially during the harder times, helped keep things in perspective and made this stage of my life a lot of fun. To my Atlanta fam, I love you to the moon and back and it has been a blast to share this time with you. To my family and friends in the North West, Europe, and all over the world, thank you for the visits to Atlanta, the trips, and for refusing to let me go MIA in the busier moments. I would like to specially thank Drs. Michele Daly, Alice Grossman and Lucy Brennan for going to such great lengths to keep things sunny, for listening, and for advising me personally and scientifically. To the Ratners, thank you for bringing the party (and also for the complementary legal advice), for always grounding me, and for putting up with me from the start. You are truly the most loving and spirited group of people.

Above all, my partner Simon Berrebi is the reason this document exists. Through the ups and downs of graduate school he constantly pushed me to trust and stand up for myself. You have always seen things so clearly, and believed in and supported me with overwhelming love, patience, and kindness. Thank you for staying in Atlanta for a little longer. You have filled this time with adventure, fun, and balance, and made this a wonderful process. I love you so much.

Table of Contents

Chapter 1. Introduction

CRISPR-Cas Biology

Adapted from: Ratner HK,* Sampson TR,* Weiss DS. Overview of CRISPR-Cas9 Biology. CRISPR-Cas: a Laboratory Manual. Chapter 1. Cold Spring Harbor Press. 2016. doi: 10.1101/pdb.top088849

History

CRISPRs (clustered, regularly interspaced, short, palindromic repeats) were detected in 1987 when an array of short, repetitive DNA sequences (~20-40 bp in length, termed "repeats") interspaced with non-repetitive sequences (termed "spacers"), was identified following the sequencing of the gene encoding alkaline phosphatase isozyme conversion enzyme (*iap*) (1). At the time, the function of these sequences was unknown, hindered largely by the lack of available DNA sequencing data (2). Soon after, interspaced repetitive palindromic regions of the genome were identified in more bacteria, including other species of *Enterobacteriaceae* species and *Mycobacterium tuberculosis* (3, 4). In *Mycobacteria*, a method of strain subtyping, spoligotyping, was developed using the variations in the spacer sequences in the CRISPR arrays (5-7). This tool has since been applied to other CRISPR-encoding species and newer CRISPR-based strain identification methods have been developed (7).

The role of these repetitive regions in prokaryotic biology remained an enigma for almost two decades, when computational analyses led to the discovery that they were present in numerous bacteria and archaea, and notably, that the spacers were identical to many sequences present in exogenous mobile genetic elements such as plasmids, transposons, and bacteriophages (8, 9). Further bioinformatic studies revealed that these arrays, termed CRISPR arrays, were often associated with a core set of Cas genes (10, 11). Many of the Cas genes had sequence

similarity to endonuclease and helicase families or genes encoding other nucleic acid binding proteins (10-12). In conjunction with the fact that many spacers were identical to mobile genetic elements, these findings gave rise to the postulation that CRISPR-Cas systems may act as a form of RNA-directed interference against foreign genetic elements (12). This hypothesis was solidified in 2007 by a set of foundational experiments that provided the first direct evidence that CRISPR sequences and the associated Cas proteins directed interference against bacteriophage infection (13). Perhaps even more interestingly, new spacer sequences were naturally acquired into the CRISPR array following bacteriophage infection, subsequently facilitating sequence-specific resistance to the offending phage, and revealing a mechanism of adaptive immunity in prokaryotes (13-17).

Overview

Since their discovery as prokaryotic adaptive immune systems, conserved features of RNA-directed interference by CRISPR-Cas systems have been uncovered (18-20). Briefly, CRISPR-mediated interference occurs in three primary stages: 1) spacer acquisition, 2) crRNA transcription and maturation, and 3) target identification and cleavage (Figure 1). During spacer acquisition, foreign nucleic acids are identified and processed into short, spacer-sized sequences that are inserted into the CRISPR array, to be flanked by a pair of repeat sequences (Figure 1A-D) (21). The CRISPR array is then transcribed and processed into mature small RNAs, called crRNAs, that each contain portions of the repeat sequences and a single spacer that facilitates identification of a target nucleic acid with significant sequence complementarity to the spacer sequence (Figure 1 E-F). The crRNAs complex with Cas protein(s) and, in some cases, additional RNAs to bind the target, resulting in target cleavage (Figure 1G-H) (18-20).

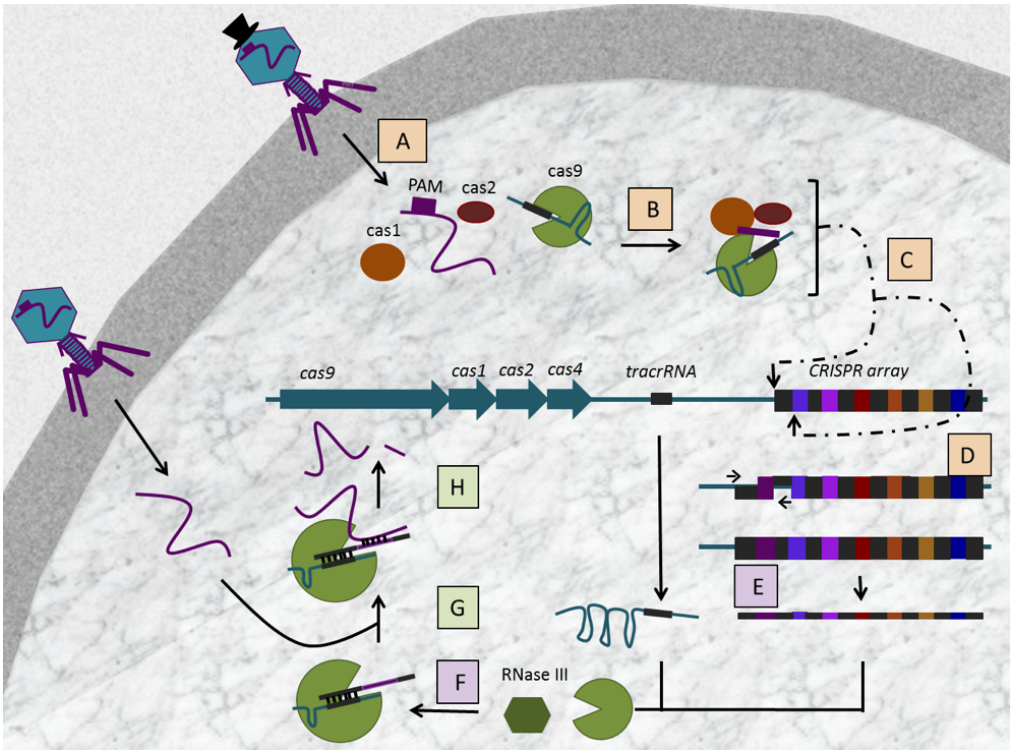


Figure 1. The three stages of adaptive immunity by CRISPR-Cas9 systems. (A-D) Spacer acquisition: (A) foreign DNA (dark purple) enters the cell, and (B) Cas1, Cas2, and Cas9 in complex with tracrRNA (blue) select a spacer sequence on the target through Cas9-mediated identification of a protospacer adjacent motif (PAM; dark purple rectangle on the foreign DNA). The PAM adjacent sequence is processed into a spacer-sized fragment. (C) The Cas protein complex attached to the spacer identify the CRISPR array and creates staggered single stranded breaks on each side on a repeat. (D) The new spacer sequence is inserted into the array and the single stranded repeats on either side of the new spacer are repaired by DNA polymerase I. (E-F) crRNA transcription and maturation: (E) the CRISPR array and tracrRNA are transcribed. (F) Cas9 binds tracrRNA and the CRISPR transcript, which is then cleaved into mature, spacer-specific crRNAs by RNase III. The mature dual crRNA:tracrRNA remains bound to Cas9 as a heteroduplex. (G-H) Target identification and cleavage: (G) Upon re-infection with foreign DNA, the spacer on the crRNA of the Cas9:RNA heteroduplex binds to its complementary sequence on the foreign nucleic acid. (H) Cas9 adopts a conformationally active state and cleaves both target DNA strands.

The field of CRISPR-Cas biology continues to rapidly expand. Numerous groups have elegantly revealed not only the molecular function of CRISPR-Cas systems in defense against foreign nucleic acids (13-15, 22-25) but also uncovered clues about the evolution of these systems (26-29), and their functions in other physiological processes (30-33).

In parallel to studies of the biological importance of CRISPR-Cas systems and molecular mechanisms that underlie their functions, these systems have been engineered for myriad biotechnological applications, revolutionizing molecular biology. In particular, technological development has focused on the CRISPR-associated endonucleases Cas9 and Cas12a. The roles of Cas9 and Cas12a in prokaryotic biology remain minimally explored. The objective of this thesis is to understand the functional flexibility of Cas9 and Cas12a systems in the context of one of their native bacterial hosts, the pathogen *Francisella novicida*. This work sheds light on the potential diversity in CRISPR-Cas system functions extending beyond DNA defense, and on how uncovering these new functions can be utilized to develop new CRISPR-Cas based tools.

Classification and Mechanistic diversity

CRISPR-Cas systems are widespread in prokaryotes, and based on whole genome datasets, can be found in 87% of sequenced archaea and 45% of bacteria (CRISPRdb 2017). CRISPR-Cas systems are structurally and mechanistically diverse (26, 34, 35). Each system has Cas genes from two functional categories, genes involved in spacer acquisition ‘adaptation,’ and genes involved in target recognition, cleavage and (to variable extent) crRNA processing that are called ‘effectors’ (36). As genome mining tools have improved and the number of sequenced prokaryotic genomes has increased, new CRISPR-Cas systems have been discovered and characterized, leading to the periodic re-classification of these systems. CRISPR-Cas systems are currently divided into two Classes, characterized by the nature of the effector protein(s) used for

nucleic acid targeting; Class 1 systems have a multi-protein effector complex, while Class 2 systems have a single protein effector (37). Each Class can be further subdivided into types and subtypes based on the unique Cas proteins, phylogeny of the most conserved Cas protein (Cas1), and the organization and structure of the genomic locus and associated RNAs (37). Class 1 is composed of Types I, III, and IV and Class 2 is composed of Types II, V, and VI (37).

The adaptation stage of immunity is the most conserved amongst CRISPR-Cas systems (38). The majority of CRISPR-Cas systems have a core adaptation module of composed of Cas1 and Cas2 (Figure 1A-C) (37) (21, 36, 38). These two metal-dependent nucleases are both necessary and sufficient for spacer acquisition, but dispensable for target interference (39-42). Cas1 and Cas2 form stable, heterodimeric complexes *in vitro*, and *in vivo*, the interaction between Cas1 and Cas2 is necessary for recognizing the DNA secondary structure of the CRISPR repeat sequence during integration of new spacers (41). Evidence from multiple types of CRISPR-Cas systems indicates that Cas1 and Cas2 may form complexes with diverse Cas proteins involved in target identification and cleavage (39, 43-46). Spacer acquisition may require these other Cas proteins to accurately select sequences in a way that prevents the CRISPR-Cas system from targeting its own chromosomal spacer sequences with the crRNAs transcribed from it; the details of this are described in the "in depth" section below for CRISPR-Cas9 systems (13, 39, 46).

The differences between the distinct types of CRISPR-Cas systems become increasingly clear at the crRNA maturation, target identification and interference stages of immunity. Notably, Class 1 systems (Type I, II, and IV), which comprise of ~90% of identified CRISPR systems, utilize large, multimeric protein complexes for these activities (45, 47). Many of the subtypes of the type I and III systems have been well studied, while little is known about type IV

systems, which lack the Cas1 and Cas2 adaptation machinery (37). However, type IV systems share similarities to Type I and III systems, using a unique variant of Cas6 to process crRNAs, which interact with a complex of at least 4 proteins (48). Type I systems, which are the most common type of CRISPR system, use the endonucleases Cas6 or Cas5d to cleave the CRISPR array transcript within the repeat sequences flanking each spacer (49-56). The Cas6 protein then transports the mature crRNA to a complex of Cas proteins called Cascade (CRISPR-associated complex for antiviral defense), which functions in interference, in some cases remaining attached to the crRNA and becoming a part of the interference complex (14, 49, 57-63). The Type I interference complex is comprised of four to five distinct Cas proteins, each with different stoichiometry (14, 49) (53). Six copies of Cas7, a protein with a ferredoxin fold that resembles an RNA Recognition Motif (RRM), form an RNA binding ridge that binds the crRNA, anchored by the other cas proteins at both ends of the Cas7 multimer (64-66). When the crRNA binds the target DNA, conformational changes result in the recruitment of the Cas3 endonuclease, which mediates target degradation and is the defining Cas protein of Type I systems (49) (16, 64).

Like the Type I systems, Type III systems use Cas6 for crRNA processing and form multi-protein complexes for target interference (55). However, the Cas proteins in the Type III complexes are different (17) (67). Cas10 is a component of Type III interference complexes and is the defining Cas protein of these systems (26). CryoEM structures of Type III systems demonstrate that the crRNA is positioned along a backbone of a Cas protein complex consisting of repeat units of Csm3 (III-A) and Cmr4 (III-B), much like the Cas7 repeats in Type I systems (17. Interestingly some, if not all Type III systems, are capable of targeting DNA and RNA {Hale, 2009 #63, 67-73). Interference by Type III-A systems is complex. First, RNA targets are recognized by the crRNA in the interference complex, resulting in even cleavage of the RNA

into even, 6 nucleotide intervals via the Csm3. Each identical subunit in the backbone individually cleaves the target to collectively fragment the invading nucleic acid into consistent and precisely sized sequences (17, 73). After recognition of an RNA target by the crRNA, two domains of Cas10 are activated, one to produce cyclic oligoadenylate (cOa) second messengers and another to degrade ssDNA in sites of active transcription, via the HD domain of Cas10 (73, 74). The cOa then activates non-specific RNA degradation by an effector outside of the complex, Csm6 (75, 76). These systems avoid autoimmunity with a more rigid interaction between Cas10 and self-RNAs that prevents activation of the ssDNA degradation domain (77).

Specificity of the crRNA for the target is enhanced through distinct mechanisms in different systems to avoid off-target effects that could occur because of binding of fully or partially complementary sequences, as mis-targeting of the host chromosome is likely lethal to the bacteria. Multiple systems, including Types I, II and V improve specificity through recognition of a specific nucleotide sequence adjacent to the target but on the complementary strand of DNA, called the PAM (protospacer adjacent motif) (8, 78) (79). PAM recognition facilitates Cas interference complex binding, DNA melting, and RNA:DNA heteroduplex formation (described for CRISPR-Cas9 in detail in the "in depth" section below), and prevents self-targeting of similar or identical sequences to the crRNA that lack a PAM. Interestingly, some Type III-A systems may avoid cleavage of sequences incorporated into the host genome through a unique transcription-dependent DNA targeting mechanism that enables tolerance of lysogenic phages while preventing lytic phage production (80).

Although there is overlap in the self vs non-self discrimination mechanism between some Class 1 and Class 2 systems, Class 2 systems use a single multi-domain effector protein. Of these, the best described systems and most commonly used in engineering is the Cas9

endonuclease from the Type II systems, and the Cas12a protein from the Type V system. The mechanism of Cas9 will be described in detail in the "in depth" section of this introduction and Cas12a will be discussed in Chapters 4+5. In brief, both type II and type V systems have RuvC-like endonuclease domains with an RNase H fold (81). Cas9 contains RuvC and HNH nuclease domains that each cleave a strand of the DNA to produce a blunt break (81). Cas9 crRNA forms a hairpin with a trans-activating crRNA (tracrRNA) that is also transcribed from the CRISPR locus and is required for crRNA processing (82). The crRNA and tracrRNA duplex is necessary for interaction with Cas9. Cas12a, which is the best characterized effector of the type V systems, works without tracrRNA and has a RuvC-like and putative Nuc nuclease domains that induce a staggered double-strand break (81). Unlike Cas9, Cas12a processes its own crRNAs, which have a small internal hairpin in the CRISPR repeats (81). Cas12a uses an AT rich PAM sequence to differentiate self from non-self, but the PAM is oriented on the other side of the spacer than the Cas9 PAM (GC rich) (81). Characterization of a different subtype from the type V systems, Cas12b (C2c1), highlights the variation between subtypes. Like Cas9, it also uses a tracrRNA but uses an AT rich PAM and produces staggered double strand breaks that depends on a RuvC-like domain (81). Although both of these systems are best characterized targeting double stranded DNA (dsDNA), some orthologs of Cas9 have been shown to be capable of directing cleavage of single stranded RNA (ssRNA) as well (83-85). This highlights how much remains to be discovered in the continually expanding mechanistic and functional landscape of these proteins.

Unlike the other Class 2 effectors, Cas13a and Cas13b of the type VI systems use a single effector protein to target RNA (81). Cas13a has been shown to process its own crRNA (81). Type VI effectors contain two higher eukaryote and prokaryote nucleotide-binding (HEPN)

domains which interestingly, are RNases similar to those found in many prokaryotic toxins (81). When the guide RNA recognizes a target, Cas13 cleaves RNA indiscriminately. When Cas13 is abundant, RNA degradation can lead to induction of physiological changes in the cell, such as dormancy, which has been hypothesized to come into play if the expression of the viral genes is too high (81).

Adaptive Immunity by CRISPR-Cas9 Systems

Adapted from: Ratner HK,* Sampson TR,* Weiss DS. Overview of CRISPR-Cas9 Biology. CRISPR-Cas: a Laboratory Manual. Chapter 1. Cold Spring Harbor Press. 2016.

crRNA maturation

Type II systems require a single Cas protein, the Cas9 endonuclease, to mediate crRNA maturation(82). The CRISPR array is first transcribed as a single, long transcript. Subsequently, this pre-crRNA transcript is processed into individual crRNAs, each specific for a different target (Fig 1 E, F). A single, matured, spacer-specific crRNA is then complexed with Cas9 along with tracrRNA. tracrRNA contains multiple stem-loop structures and a sequence with partial complementarity to the CRISPR repeat sequence, allowing binding to the crRNA to facilitate maturation and complex formation with Cas9 (27, 82, 86-88). The dsRNA endonuclease, RNase III, which is typically encoded distal from the CRISPR locus, is also required for crRNA maturation (82). RNase III recognizes the dsRNA structure created by the tracrRNA:crRNA duplex and cleaves both strands of RNA within the double stranded repeat region (82). The tracrRNA:crRNA duplex binds tightly to Cas9, and undergoes additional processing through an unknown mechanism that likely involves additional bacterial RNases (82). The dual RNA:Cas9 complex is then able to identify and cleave targets with sequence complementarity to the crRNA

spacer (Figure 1G, H) (82, 86-89). In some type II systems, notably those encoded by the pathogen *Neisseria meningitidis*, maturation of the crRNAs is independent of RNase III and tracrRNA (90). In this case, internal promoter sequences within each repeat sequence allow for transcription of individual crRNAs. These crRNAs still require tracrRNA in order to associate with Cas9, highlighting the importance of the RNA duplex for interactions with this protein (90).

Target interference

The mechanism of target interference by type II CRISPR-Cas systems has been greatly informed by solving the crystal structures of Cas9 alone and bound to DNA and RNA (82, 87-89) (91, 92). Similar to its role in crRNA maturation, Cas9 is the sole Type II Cas protein involved in target surveillance and interference (82, 87).

Cas9 has a two-lobed morphology, with a larger alpha-helical lobe and smaller nuclease lobe that together form a clam-like shape with a central channel to position the target (Figure 2 A, B) (91, 92). Cas9 first binds the crRNA:tracrRNA duplex via a positively charged arginine rich motif located on the inner surface of the α-helical lobe, where the two lobes come together at the end of the central cavity (91, 92). Upon RNA binding, Cas9 undergoes a first conformational change to create the channel that positions the nucleic acids along the length of the protein, by rotating the nuclease lobe around the nucleic acid binding pocket of the α-helical lobe (91, 92). This reorients the endonuclease domains to either side of the channel, into a favorable conformation for subsequent target cleavage (Figure 2 B,C) (91, 92).

Cas9 must then scan DNA to identify target sequences with a high degree of accuracy so as not to target its own chromosome. This is partially accomplished by the requirement for the PAM motif (typically ~3 base pairs) adjacent to the targeted region on the target DNA (Figure 1 and 2) (87-89). Cas9 associates and dissociates randomly along a DNA strand until encountering a

PAM sequence (93). Subsequently, the PAM-interacting domain of Cas9 (located in the carboxyl-terminus) binds tightly to the target DNA through two binding loops that interact with the major and minor grooves of the PAM (91, 92).Cas9 then undergoes a second conformational change, locking the DNA target into place along the length of the central cavity between the two lobes (91, 92). Interaction with the PAM leads to destabilization of adjacent double-stranded DNA and orients the target sequence to facilitate binding to the seed region of the crRNA (91, 92). If the target sequence has near-perfect complementarity in the PAM-proximal region of the spacer, melting along the DNA will occur as one strand of the target base pairs along the remainder of the complementary spacer, forming an RNA:DNA heteroduplex (91, 92, 94). This results in separation of the two DNA strands into distinct, metal ion-dependent endonuclease active sites (91, 92). The HNH endonuclease domain cleaves the DNA strand bound to the RNA three nucleotides upstream of the PAM, whereas the non-complementary strand is also bound by the nuclease lobe of Cas9 but cleaved by a separate RuvC domain (91, 92).

Spacer acquisition

In type II-A systems, all components of the CRISPR-Cas system form a complex that is required for adaptation (Cas1, Cas2, Cas9, Csn2, and tracrRNA) (46, 95). A similar mechanism is likely used by other Type II subtypes that contain these components, excluding Csn2, which is absent from Type II-C and is replaced by Cas4 in Type II-B subtypes (27, 86). Both Csn2 and Cas4 resemble RecB-like nucleases and may therefore play a similar role in adaptation, although their precise functions are not known (45). Csn2 and Cas4, as well as Cas1 and Cas2, are all dispensable for crRNA processing and target interference in Type II CRISPR-Cas systems(87). Interestingly, the Cas1 proteins present in Type II CRISPR-Cas systems cluster phylogenetically with those of Type I systems (27). This may suggest that the distinct functions of Type II

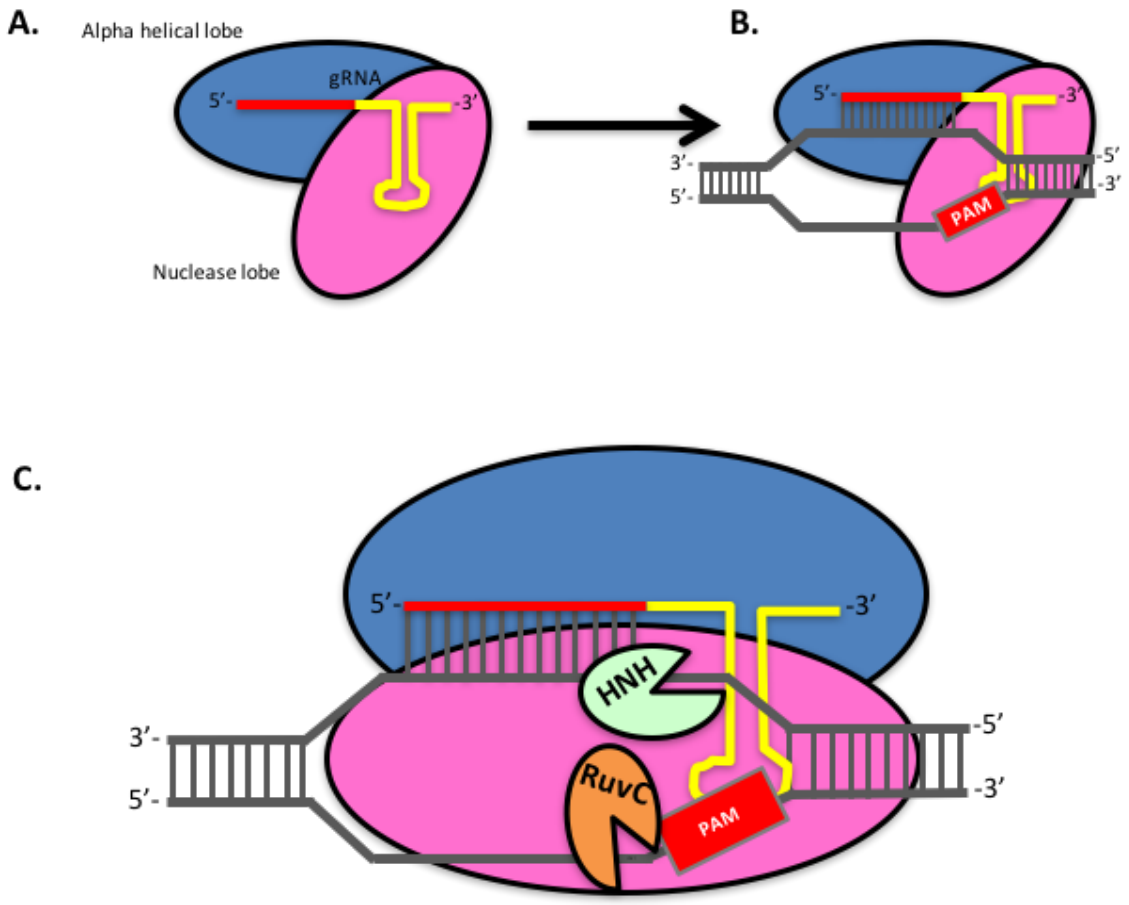


Figure 2. Schematic of Cas9:gRNA interactions. A. Upon association with a chimeric gRNA, consisting of a ssRNA targeting region similar to the crRNA (red) and a dsRNA structure similar to that created by the crRNA:tracrRNA complex (yellow), the alpha helical lobe (blue) and the nuclease lobe (pink) of Cas9 are opened into a conformation that reveals a channel for DNA targets to bind. B. When DNA containing a PAM sequence is identified by Cas9, and the targeting sequence of the gRNA (red) has significant sequence complementarity to the immediately adjacent DNA sequence, the DNA is melted and unwound generating a DNA:RNA hybrid. C. Cas9 then undergoes a conformational change, clamping its nuclease lobe across the targeted DNA, and positioning each strand into the HNH and RuvC active sites. The HNH and RuvC endonuclease domains then cleave the complementary and non-complementary strands, respectively, resulting in a double strand break in the target immediately adjacent to the PAM.

CRISPR systems arose via recombination events with Cas9 and other types of CRISPR-Cas systems, such as the Type I system (27).

Upon invasion by a foreign nucleic acid, CRISPR-Cas systems must select spacer sequences in a manner that prevents autoimmunity (46, 96). Type II systems accomplish this by requiring a specific PAM sequence adjacent to the one that will ultimately be integrated as the spacer (i.e. the protospacer) (46) (41, 97). In Type II-A systems, Cas9, in complex with Cas1, Cas2, and Csn1 and bound to tracrRNA, identifies PAMs on the invading DNA in order to facilitate spacer selection using the PAM-interacting domain (46, 91, 92, 95). There may be additional requirements for the selection of the spacer sequence, as there is an enrichment for certain spacer sequences that cannot be accounted for by the sequence of the PAM alone, however these requirements have yet to be identified (46).

Mutations in the PAM-interacting domain of Cas9 do not prevent spacer acquisition but instead result in incorporation of spacers that are not adjacent to a PAM in the target (46).The endonuclease activity of Cas9 is dispensable for acquisition, suggesting that the role for Cas9 is primarily to select spacers by binding to the PAM and protospacer sequence, whereas Cas1 (whose non-specific nuclease activity is required for adaptation) of the associated Cas1-Cas2-Csn1 complex cleaves the adjacent sequence, yielding a precisely selected spacer sequence (46). There are many unknowns in the mechanism of adaptation, but a general model has been developed (Figure 1A-D) (21, 41, 42, 46). Cas1-Cas2 together interact with the secondary structures of the CRISPR repeat sequences within the array, preferentially near the leader sequence, which also acts as a promoter (41, 42). A repeat sequence within the chromosomal array is then nicked at the 3’ end, allowing for ligation of the free hydroxyl to the spacer fragment (42). The spacer is inserted into the array, flanked by the single complementary strands

of the first CRISPR repeat (42). These are repaired into double stranded repeats by DNA polymerase, resulting in a new repeat-flanked spacer in the chromosome, to be transcribed and processed into a crRNA that can protect against future invasion by complementary, PAM-flanked sequences (42).

Francisella

Introduction to *Francisella spp*

The *Francisella* genus is comprised of three species of Gram-negative coccibacilli, *Francisella tularensis*, *Francisella novicida*, and *Francisella philomiragia* (98). *Francisella* are non-motile, encapsulated, facultative intracellular pathogens of many mammals including humans. Disease severity and ecology of *Francisella* varies by species. *F. novicida* is likely to reside in an environmental niche, as all environmental isolates have come from salt water (99). Analyses of brackish water and soil samples suggest that it may also reside there as well (100). In contrast, *F. tularensis* is a vector-borne zoonotic pathogen: it amplifies in a mammalian host, which is then fed on by arthropod vectors, and those vectors disseminate the bacterium to new hosts (99, 100).

Francisella tularensis is the most virulent species of the *Francisella* genus. It is the causative agent of tularemia, colloquially referred to as "rabbit fever." *F. tularensis* was first isolated in 1911 in Tulare county, California in 1911, after causing a plague-like disease in rodents (101). The namesake of the genus, Edward Francis, found that the same pathogen was responsible for "tularemia" in humans (102). There are three recognized subspecies of *F. tularensis*, *tularensis*, *holartica*, and *mediasiatica*. Of these, *tularensis* is found in the Northern

Hemisphere. *F.t. holartica* has been isolated in Europe and Asia, and *F. t. mediasiatica* has been found in central Asia (103).

The virulent *F. tularensis* is transmitted to humans by multiple routes including arthropod vectors and zoonotic transmission, ingestion of contaminated material, or aerosolized bacteria (98). Notably, inhalation of as few as 10 bacilli of the highly virulent SchuS4 strain can cause lethal infection (104, 105). The most common route of infection with *F. tularensis* is through exposure to infected animals such as rodents and rabbits, and there are no known cases of person-to-person spread (98, 106). The nature and severity of infection varies with route of transmission. After ~3-5 days, patients typically develop flu-like symptoms or a lesion at the infection site. Infection through the skin is the most common and typically results in ulceroglandular infection (106). Glandular infection can arise independently or from a cutaneous infection draining into the lymph nodes and causing systemic infection (106). Ocular, oropharyngeal and gastrointestinal infections are possible as well, and have diverse clinical presentations. Of the transmission routes, inhalation results in the most severe form of infection, respiratory tularemia (107). Treatment with antibiotics is typically effective.

The risk of naturally acquiring tularemia is quite low. In 2017, the CDC reported 239 cases. Never-the-less, because of the extremely low infectious dose and ease of aerosolization, *F. tularensis* is one of the most infectious bacterial pathogens known (99). This infectivity led to the development of *F. tularensis* as a bioterrorism agent by the U.S. and Soviet Union from 1940s-1960s (108). It is classified as a "category A Select Agent" by the CDC, making it one of 6 "most likely" biological threat agents (108). In the last few decades, *F. tularensis* has received renewed research interest, with a focus on elucidating the factors that enable its high level of virulence and evade host immune defenses. Much of this research has been conducted on

Francisella novicida, which is rarely causes disease in humans and can be studied in biosafety level 2 settings. *F. novicida* is often used as an alternative to *F. tularensis* in laboratory assays because it causes tularemia-like disease in mice and has similarities in intracellular lifecycle, although there are substantial differences as well (99, 106). Discoveries made during this resurgence of research attention has resulted in *Francisella*, and specifically the less pathogenic *Francisella novicida*, becoming a model for multiple fields of research, including Type VI secretion systems, intracellular infection, inflammasome activation, and Class II CRISPR-Cas biology.

Francisella **intracellular lifecycle**

The ability to survive and replicate intracellularly is a critical for *Francisella* virulence (98, 106). *Francisella* is taken up in the phagosome of the host cell, where it must escape phagosomal killing by host antimicrobials and prevent detection from the host innate immune system. From there, it escapes from the phagosome to replicate in the cytoplasm until the cell dies, allowing the released bacteria to go on to infect new cells.

Francisella can infect multiple host cell types (109). Entry into mammalian cells is not fully understood, and the receptors vary based on the opsonization state of the bacterium. For example, mannose is a receptor on macrophages for unopsonized bacteria, while opsonized bacteria are targeted to CR3 (complement receptor), although other receptors have also been shown to be important (109). Once inside the phagosome*, Francisella* encounters a smorgasbord of host-produced antimicrobials to restrict bacterial infection, including cationic antimicrobial peptides (CAMPS) and reactive oxygen stress (106). *Francisella* is unable to replicate intracellular in the phagosome, hence it must disrupt the membrane, escape and replicate in the

cytosol (98). It is not well established whether acidification is required for escape, which likely also depends on the opsonization state of the bacterium (98).

Phagosomes containing *Francisella* have been shown to have markers of the early and late endosome, but do not fuse with the lysosome, indicating that *Francisella* is able to delay, disrupt, or escape prior to maturation of the phagosome and toxification due to fusion (110). Phosphoinositides decorate the vacuolar compartments of the eukaryotic cells to direct their fate (111). In *F. novicida*, the levels of phosphatidylinositol-3-phosphate (PI(3)P drives phagosomal maturation (112). The type VI secretion system of *Francisella* secretes multiple effectors located outside of the *Francisella* Pathogenicity Island (FPI) where the type VI system is encoded, and has been shown to be required for phagosomal escape (113). A secreted protein, encoded in the genome outside of the FPI, OpiA, alters phagosomal processing by phosphorylating PI on the endosome into PI(3)P to a high level that stops maturation (112). The *in vivo* phenotype of OpiA is dependent on another secreted effector with unknown function, PdpC (112). It is not yet clear if there is a single route of escape to the cytosol, or the role of other genes that have been implicated in this process, such a pyrimidine biosynthesis genes (98). Upon escape into the cytosol, *Francisella* continues to modulate host pathways in order to replicate (98). Importantly, Francisella must also evade innate immune signaling in the cytosol to prevent detection (98).

Innate Immune Evasion and Survival

At each of the lifecycle stages described above, *Francisella* must prevent detection and killing by the host innate immune defenses. Extracellularly, *Francisella* must avoid host defenses that can bind to and kill the bacterium such as complement, antibodies, and cationic antimicrobial peptides (106). O antigen enables *Francisella* to inhibit the complement system and avoid opsonization by antibodies (106). To resist detection, killing, and the release of

bacterial components called PAMPs (pathogen associated molecular patterns) that could activate the innate immune response, *Francisella* also has a capsule and non-canonical LPS (lipopolysaccharide) (106).

During the intracellular lifecycle and passage through a eukaryotic host cell, *Francisella* must avoid recognition by pattern recognition receptors (PRRs) that recognize conserved molecular features of microbes in order to alert the innate immune system to control infection. PRRs such as Toll-like receptors located on the cell surface and on endosomes, and Nod-like receptors (NLRs) in the cytosol, recognize PAMPs and danger associated molecular patterns (DAMPs). Pathogen recognition triggers a proinflammatory response, inflammasome activation, and pathogen elimination (98).

TLRs are a family of PRRs that play a critical role in initiating host defenses by recognizing different bacterial PAMPS (106). Upon initiation of TLR signaling, it activates NFkB transcription factors, which lead to the production of proinflammatory cytokines and antimicrobial peptides to control bacteria dissemination and cause disease (106). *Francisella* has evolved specific modifications to the structure of LPS to subvert host defenses and resist cationic antimicrobials, and therefore does not efficiently activate TLR4 compared to LPS from other Gram-negative pathogens. *Francisella* species lipid A acyl chains are 2 to 6 carbons longer than those in *E. coli* LPS, it is tetra-acylated, and is missing both negatively charged 1 and 4 phosphate groups (106). In addition, the phosphate group at 1' position of lipid A is masked with positively charged sugar, GalN, leading to increased bacterial surface charge and repulsion of cationic antimicrobials peptides (cAMPs) to facilitate resistance to cationic antimicrobials (106).

Antimicrobials produced in the phagosome are a first line of defense against bacterial pathogens. cAMPs such as human cathelicidins, LL-37, can disrupt the negatively-charged

bacterial outer membrane, harm the bacterium, and promote the release of PAMPs that subsequently activate an inflammatory response (106). In addition to surface charge, *Francisella* can resist cAMPs though efflux systems (106). The oxidative bursts is another primary antimicrobial threats to *Francisella* survival in the phagosome, and as a result, Francisella has multiple ways to resist and prevent it (98). This includes interference with the assembly of ROS generating complexes on the phagocytic membrane and expression of detoxification genes such as superoxide dismutase (98).

TLR2, which recognizes bacterial lipoproteins (BLPs) and peptidoglycan, is the primary TLR that responds to *Francisella* infection (106). Discussed in more detail below (in the following section, Chapter 2, Chapter 6, and Appendix B), *Francisella novicida* represses a BLP using Cas9 in order to evade detection by this signaling pathway (114). FPI encoded genes and the type VI secretion system have also been implicated in TLR2 evasion and reducing TLR2 expression (106). Carbohydrates on the *Francisella* capsule may also help hide components of the outer membrane and cell wall from detection by TLR2 (106).

Once *Francisella* escapes to the cytosol and during replication, it must avoid recognition of PRRs and the absent in melanoma2 (AIM2) inflammasome, which detects dsDNA in the cytoplasm (106). When dsDNA is detected, the inflammasome molecular complex orchestrates caspase-1 activation and leads to cell death via pyroptosis to remove infected cells from the system (106). In doing so it induces an inflammatory response that recruits immune cells to the site of infection. Inflammasome activation is potentiated by TLR2 signaling, and depends on Type I interferon detection of *Francisella* in the cytosol (106). To avoid unnecessary DNA release, Francisella reduces lysis by increasing membrane integrity. BLP repression and other membrane protein regulation has been implicated in evasion of this pathway (115, 116).

Also in the cytosol, Francisella must avoid host defenses such as NLRs and autophagy. NLRs recognize bacterial components such as peptidoglycan, secretion systems, pore forming toxins, and flagellin. The role of these pathways and mechanisms of activation and evasion are not well established (106). *Francisella* has been detected in autophagosomes, and can replicate in the cytoplasm for long periods of time, indicating that it must avoid engulfment into the autophagosome and fusion to the lysosome, or control autophagosomal trafficking (98).

Model for CRISPR-Cas System Biology

As an intracellular pathogen, *Francisella* is in a constant tug-of-war with the host cells to prevent recognition and control of infection, and enable replication and escape. To understand the genes and pathways that underly *Francisella* virulence, multiple screens for genes involved in *Francisella novicida* fitness in different environments have been conducted, such as during macrophage infection. Interestingly, the gene *FTN_0757 (FTT_0584)* was identified as being essential for virulence and evasion of TLR2 during one of these screens (117). It was later identified that this gene was involved in the repression of a BLP (*FTN_1103*), which was essential for pathogenesis in a mouse model and evasion of TLR2 (114). Fascinatingly, *FTN_0757* encodes Cas9, providing some of the first evidence that CRISPR-Cas systems can have functions beyond foreign DNA defense (118).

Furthermore, *F. novicida* encodes a second CRISPR-Cas system, CRISPR-Cas12a (formerly Cpf1) (119). The simplicity of this system, described in detail in Chapter 5, has resulted in significant research attention to characterizing the mechanism of action and development for use as an alternative to Cas9 in genome engineering. As a result, *F. novicida* has become a model for studying the broader functions of CRISPR-Cas systems in prokaryotic

biology and more generally, the only system to study the biology of the two primary CRISPR-Cas systems used in genome engineering in their native bacterial context.

In this thesis, I propose that DNA defense likely represents just one of many functions of CRISPR-Cas systems, and in particular, Cas9- and Cas12a encoding systems. Chapter 2 outlines the early associations between CRISPR-Cas systems and roles in bacterial physiology beyond DNA defense. The molecular mechanism underlying the association between Cas9 and *F. novicida* virulence is elucidated in Chapter 3, which demonstrates that *F. novicida* Cas9 regulates endogenous transcription by binding to the DNA to enable virulence, but is also capable of DNA restriction. The different functions of *F. novicida* Cas9 are directed by competing small RNAs. Chapter 4 compares and characterizes the DNA targeting capabilities of the Cas9 and Cas12a systems of *F. novicida* in their native host. We synthesize the findings in Chapters 3 and 4 to show that crRNAs can direct both Cas9 and Cas12a to repress transcription or target DNA (Chapter 5). The function of these proteins is therefore determined by the associated crRNA, parameters we applied to reprogram Cas12a to repress endogenous targets (Chapter 5).

I conclude by providing new evidence that BLP repression by Cas9 contributes to virulence by protecting against killing by host antimicrobials, adding to an emerging body of evidence that links the bacterial envelope with sensitivity to reactive oxygen species. I discuss the ecological distributions of these strains and propose a hypothesis as to why their distribution doesn't track along the expected correlates like phylogeny. Then, I discuss how the discovery of the mechanism of natural Cas9 transcriptional repression, and the discovery that this capability co-exists with DNA defense, extends to another Class 2 CRISPR-Cas system, Cas12a. Collectively, this work contributes to growing evidence that CRISPR-Cas systems have diverse

functions in bacterial physiology. Finally, I outline how the multifunctional nature of these systems expands their capabilities as engineering tools, and how the continued study of non-immunity CRISPR-Cas functions is important for both engineering and biology.

Chapter 2. CRISPR-Cas Functions Beyond Adaptive Immunity

I can see CRISPR now, even when phage are gone: a view on alternative CRISPR-Cas functions from the prokaryotic envelope

Hannah K. Ratner[1,2,3,†] , Timothy R. Sampson[1,2,3*†], and David S. Weiss[2,3,4#]

[1]Department of Microbiology and Immunology, Microbiology and Molecular Genetics Program; [2]Emory Vaccine Center, [3]Yerkes National Primate Research Center, [4]Division of Infectious Diseases, Department of Medicine, Emory University School of Medicine, Atlanta, Georgia 30329.

† These authors contributed equally

Published in Current Opinion in Infectious Diseases, 2015.

Abstract

Purpose: CRISPR-Cas systems are prokaryotic immune systems against invading nucleic acids that adapt as new environmental threats arise. There are emerging examples of CRISPR-Cas functions in bacterial physiology beyond their role in adaptive immunity. This highlights the poorly understood, but potentially common, moonlighting functions of these abundant systems. We propose that these non-canonical CRISPR-Cas activities have evolved to respond to stresses at the cell envelope.

Recent findings: Here, we discuss recent literature describing the impact of the extracellular environment on the regulation of CRISPR-Cas systems, and the influence of CRISPR-Cas activity on bacterial physiology. The described non-canonical CRISPR-Cas functions allow the bacterial cell to respond to the extracellular environment, primarily through changes in envelope physiology.

Summary: This review discusses the expanding non-canonical functions of CRISPR-Cas systems, including their roles in virulence, focusing mainly on their relationship to the cell envelope. We first examine the effects of the extracellular environment on regulation of CRISPR-Cas components, and then discuss the impact of CRISPR-Cas systems on bacterial physiology, focusing on their roles in influencing interactions with the environment including host organisms.

Keywords: CRISPR-Cas, envelope stress, membrane composition, bacterial pathogenesis

Highlights

- CRISPR-Cas systems play roles in bacterial gene regulation.
- Regulatory roles of CRISPR-Cas systems center around actions at the bacterial envelope.
- The ability to respond to envelope stress may have driven the acquisition of CRISPR-Cas regulation

Introduction

Prokaryotic organisms have evolved unique, adaptive, nucleic acid restriction machineries to prevent the uptake of mobile genetic elements, such as those derived from bacteriophages and plasmids (1). Termed CRISPR (clustered, regularly interspaced, short, palindromic repeats) - Cas (CRISPR-associated) systems, these RNA-guided endonuclease machineries canonically act in a sequence-specific fashion to cleave foreign DNA or RNA targets (2-5). This protects cells from exposure to potentially harmful genetic elements (2-4). Beyond this well-established function, CRISPR-Cas systems have been observed to play alternative roles in physiology. These moonlighting functions of CRISPR-Cas systems include roles in oxidative stress tolerance, antibiotic resistance, extracellular structure formation, DNA repair, and host-microbe interactions.

The molecular mechanism of many alternative CRISPR-Cas functions has not yet been fully elucidated, but may utilize a similar activity to that used in canonical targeting of foreign nucleic acids (6, 7). The signature component of CRISPR-Cas systems is the CRISPR array, composed of short, repetitive, and often palindromic sequences (8). These repeats are interspaced by short, unique, spacer sequences that are complementary to different nucleic acid targets (2, 9, 10). In

most systems, the CRISPR array is transcribed as a single transcript (the pre-crRNA array) and is cleaved into small targeting RNAs (crRNAs)(11-14). These crRNAs form complexes with Cas proteins, which are encoded in adjacent, conserved operons (4). The complexes are capable of sequence-specific interaction with foreign nucleic acids (6). Upon hybridization of the crRNA to its target sequence, endonuclease activity of the associated Cas protein(s) is triggered, resulting in target cleavage (6). CRISPR-Cas systems are diverse and can be grouped into three main subtypes (types I, II, and III) defined by the unique Cas proteins used in crRNA processing and targeting/cleavage (1). While the type I and III systems use multimeric protein complexes for these processes, the type II system requires a single Cas protein, Cas9, as well as a unique accessory RNA, the trans-activating CRISPR RNA (tracrRNA) (1, 13, 15, 16). Uniquely, CRISPR-Cas systems can also acquire new spacer sequences within the CRISPR array as the nucleic acid threats (such as bacteriophages) in the environment change (2, 17).

Interestingly, many of the alternative activities (not involving the targeted degradation of foreign nucleic acid) of CRISPR-Cas systems are linked to processes occurring at the bacterial envelope. Herein, we present a CRISPR view of how CRISPR-Cas systems monitor and respond to stresses at the cell envelope, allowing bacteria to counteract not only bacteriophage infection, but also diverse insults such as antibiotics and host defenses. First, we discuss the transcriptional regulation of CRISPR-Cas systems in response to environmental changes signaled by the status of the bacterial envelope. We then describe the current understanding of how CRISPR-Cas systems regulate bacterial physiology, largely through changes at the cell surface, to promote resistance to environmental stresses. Finally, we highlight unanswered questions in the field of CRISPR-Cas biology, the exploration of which will provide insight into the evolution of

CRISPR-Cas systems and the origins of their increasingly broad functions in bacterial physiology.

Activation and function of CRISPR-Cas systems in response to envelope stress

Since CRISPR-Cas systems target nucleic acids that have entered the cell through the envelope, it is interesting to note that their transcriptional activation often occurs directly, and indirectly, in response to envelope stresses (Figure 1). The most explicit example of this occurs during bacteriophage infection. It is logical to think that upon bacteriophage adsorption and DNA injection the envelope is disrupted, resulting in an envelope stress response (18-20). Concomitantly, activation of CRISPR-Cas transcription has been observed, suggesting that the cell actively senses the status of the envelope in order to respond to invading threats (21, 22). Furthermore, it has been observed that membrane protein dysregulation is capable of inducing the increased expression of CRISPR-Cas systems. For instance, in *Escherichia coli*, the BaeSR extracytoplasmic stress response regulator acts to activate its CRISPR-Cas system when the bacterial envelope is perturbed (23). Furthermore, the transcriptional regulator H-NS is an inhibitor of CRISPR-Cas expression. Upon an envelope stress response, H-NS is inhibited, leading to an upregulation of a CRISPR-Cas system in *Salmonella enterica* and *E. coli* (24, 25). Additionally, high temperatures result in misfolding of membrane proteins and an envelope stress response leading to activation of heat shock protein G (HtpG) (26, 27). HtpG has subsequently been shown to activate transcription of CRISPR-Cas systems in *E. coli* (27). Thus, CRISPR-Cas systems can be primed by stress at the envelope, likely at least in part to counteract incoming foreign nucleic acids.

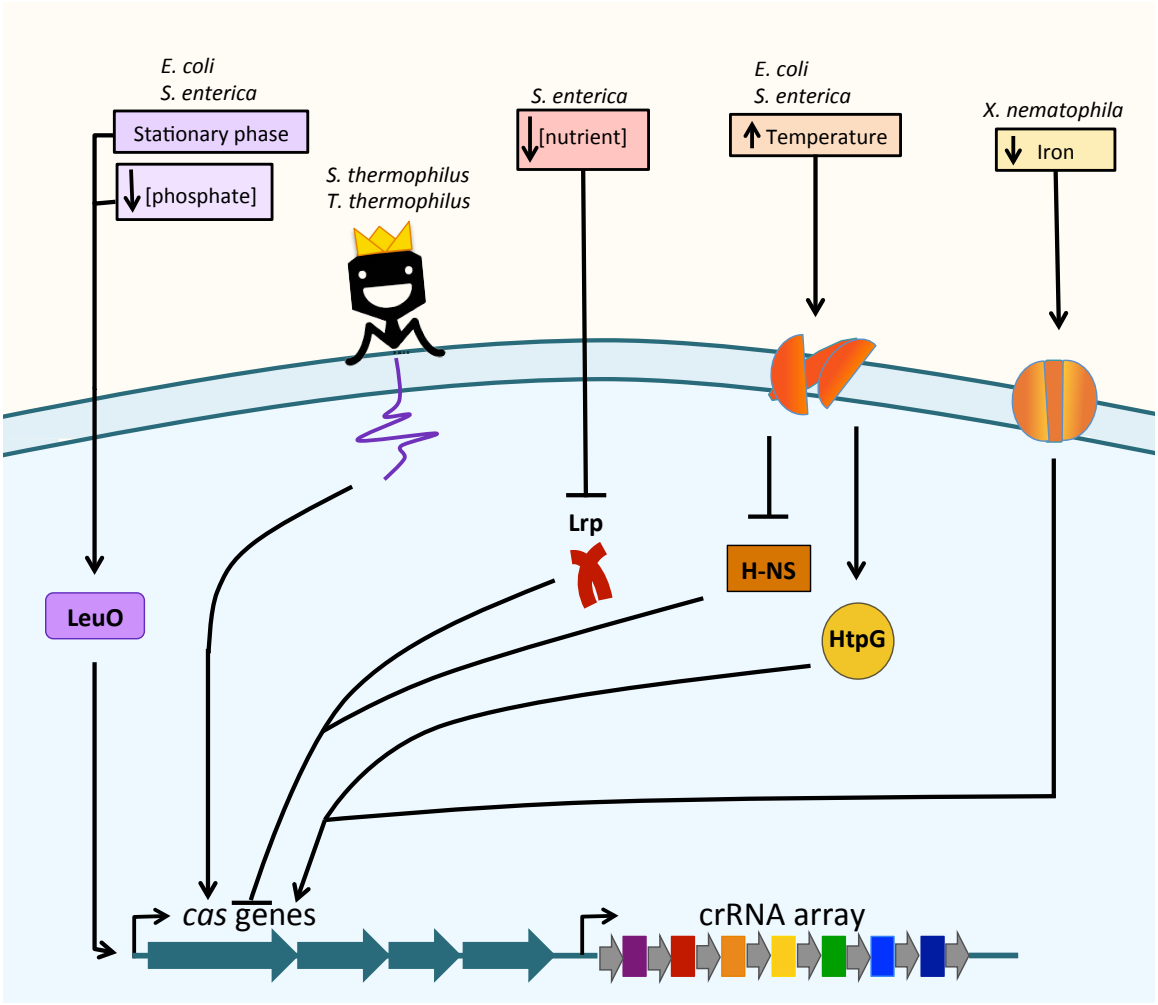


Figure 1. Activation of CRISPR-Cas systems in response to environmental changes. CRISPR-Cas systems can be activated in response to the broader environmental stressors of nutrient starvation, stationary phase growth, and iron limitation. Likewise, CRISPR-Cas systems can be activated directly in response to envelope stressors, such as phage infection and high temperature. These examples highlight the influence of the extracellular environment on the regulation of CRISPR-Cas systems.

In line with this idea, a recent study of *Streptococcus mutans*, a cause of tooth decay, revealed that expression of the Type II-A CRISPR-Cas system was negatively affected by the stress response regulator VicK/R two-component system, which also positively regulated the expression of its Type I-C system (28-30). Additionally, it was observed that both of these CRISPR-Cas systems play a role in temperature stress tolerance. CRISPR-Cas locus deletion mutants exhibited reduced survival after heat exposure, and surprisingly, double mutants in both loci had a greater sensitivity to high temperature than mutants from either locus alone, suggesting independent activity of each system (30). Furthermore, CRISPR-Cas mutants in the type II-A system, but not the Type I-C system, displayed reduced growth upon exposure to membrane stress (detergents) as well as oxidative stress (paraquat and hydrogen peroxide)(30). Together, these data directly link CRISPR-Cas function to envelope stresses, and further suggest that VicK/R may differentially regulate each CRISPR-Cas system under specific conditions. This raises the questions of whether these systems work together in nucleic acid defense as well, if they have distinct defense activities beyond adaptive immunity, or if they diverged in function to fulfill distinct regulatory roles, perhaps by altering the envelope. Exactly how these CRISPR-Cas systems regulate stress tolerance remains to be elucidated, and continued study of this phenomenon in diverse bacteria will be necessary to identify common themes. It is reasonable to postulate that this occurs through physiological changes at the envelope, which acts as the frontline to counteract environmental stressors.

CRISPR-Cas control of population behaviors

In addition to roles in the envelope stress response, CRISPR-Cas systems have been implicated in complex population behaviors that involve extensive envelope alterations, such as biofilm

formation and fruiting body development (Figure 2). Before CRISPR-Cas systems were identified, three genes encoded by the Gram-negative saprophytic bacterium *Myxococcus xanthus*, were found to be necessary for sporulation and fruiting body development (31-33). Interestingly, the three genes, *devT, devR,* and *devS*, respectively correspond to *cas8, cas7,* and *cas5* from a type I CRISPR-Cas system. In the absence of *devT* (*cas8*), *M. xanthus* displayed delayed cellular aggregation, sporulation, and chemotaxis, as well as decreased transcript levels for a fruiting body transcriptional activator (31). While the mechanism of regulation has not been fully elucidated, the *M. xanthus* CRISPR array encodes two spacers that have identity to endogenous sequences on the bacterial chromosome. One has identity to an integrase of a *Myxococcus* bacteriophage, while the other has identity to a *cas* gene in a different CRISPR-Cas locus, raising the intriguing possibility that the CRISPR-Cas system regulates endogenous targets (33). However, whether the CRISPR array itself is required for control of the aforementioned processes remains unknown.

M. xanthus regulation of fruiting body formation is further influenced by a type III-B CRISPR-Cas locus, which also regulates exopolysaccharide (EPS) production and type IV pili mediated chemotaxis (34). Not only is crRNA processing required for this regulatory activity, but the associated *cas* genes are as well (34). Further studies are needed to determine if and how the type I and III systems in *M. xanthus* interact to regulate fruiting body formation, as well as the mechanism of CRISPR-Cas mediated EPS regulation. It will be interesting to determine whether these functions evolved due to pressures to restrict mobile genetic elements, broader stresses at the envelope, or from entirely different environmental pressures.

Another population behavior involving extensive envelope changes, biofilm formation, is regulated by the type I CRISPR-Cas system in the opportunistic pathogen *Pseudomonas aeruginosa* (35, 36). A spacer within the *P. aeruginosa* CRISPR array has sequence similarity to a gene within a chromosomally integrated prophage (36). The CRISPR-Cas system interaction with this chromosomal element is necessary to repress swarming motility and biofilm formation (35, 36). While it is not known how repression occurs, it is established as a sequence-specific activity requiring all interference components of this CRISPR-Cas system (36, 37). Given the importance of biofilm formation to antibiotic resistance and pathogenesis in *P. aeruginosa*, it is likely that this CRISPR-Cas system plays an important role in mediating infection of eukaryotic hosts.

CRISPR-Cas mediated regulation of host-pathogen interactions

While all bacteria encounter numerous environmental stresses, those bacteria that interact with eukaryotes, particularly mammalian hosts, are subjected to a variety of microenvironments and stressors as they traffic through the host and encounter the immune system (Figure 2). It is therefore an exciting proposition that CRISPR-Cas systems may be utilized in response to these host-derived stresses and ultimately mediate host-microbe interactions.

Recently, it has been observed that CRISPR-Cas systems can modulate host immune evasion. The intracellular pathogen *Francisella novicida* upregulates its type II-B CRISPR-Cas system in the phagosome of host macrophages, a stressful environment containing a plethora of host defenses that attack the bacterial envelope (38). Components of this system (Cas9, tracrRNA, and a small CRISPR-Cas associated RNA [scaRNA]) regulate the production of an endogenous bacterial lipoprotein (BLP), a process necessary for strengthening the bacterial envelope (38, 39).

Loss of these components results in increased envelope permeability and subsequently increases susceptibility to membrane damaging compounds, such as those found in the macrophage phagosome (39). Furthermore, regulation of the BLP dramatically alters how *F. novicida* survives within its mammalian host. In fact, *cas9* mutants are attenuated in a mouse model by 10^3-10^4 fold compared to wild-type bacteria (38). Cas9 and its associated RNAs enable evasion of the host innate immune response through two distinct pathways, both of which originate due to changes at the membrane. In the absence of Cas9, the BLP transcript is de-repressed, and the bacteria are detected by the host pattern recognition receptor (PRR) Toll-like receptor 2 (TLR2), which initiates a proinflammatory response upon recognition of BLP (38). Additionally, repression of the BLP increases envelope integrity and reduces activation of the AIM2/ASC inflammasome, a protein complex involved in a programmed host cell death pathway that results in loss of *Francisella*'s replicative niche (39). This CRISPR-Cas mediated evasion of both TLR2 and the AIM2/ASC inflammasome is critical for the ability of *F. novicida* to cause disease.

Consistent with the idea that CRISPR-Cas systems have evolved functions to mediate interactions with eukaryotic hosts, *Neisseria meningitidis* Cas9 is necessary for intracellular survival in human epithelial cells (38). Further, *N. meningitidis* Cas9 is also required for attachment and entry into these cells, processes dependent on surface components, suggesting that it may regulate envelope structures in this bacterium (38). Cas9 is likewise necessary for attachment and intracellular survival of *Campylobacter jejuni*, a cause of diarrheal disease and Guillain-Barré syndrome, in epithelial cells (40). Furthermore, *C. jejuni cas9* mutants displayed increased surface antibody binding, as well as increased envelope permeability and antibiotic susceptibility, all potentially linking Cas9 to the regulation of envelope components (40).

Finally, it was observed bioinformatically that the presence of envelope sialylation correlates with a loss of the type II CRISPR-Cas system in multiple bacteria (including *N. meningitidis, C. jejuni*, and *Haemophilus parainfluenzae*)(40). Taken together, these data provide additional evidence for alternative functions of CRISPR-Cas systems in regulating envelope functions in response to environmental pressures.

Another example of a CRISPR-Cas system promoting host-microbe interactions is observed in the Gram-negative bacterium *Xenorhabdus nematophila*. Here, an orphan CRISPR RNA, termed NilD, is necessary for *X. nematophila* to colonize *Steinemema spp*. nematodes, a symbiotic relationship that facilitates the pathogenesis of these nematodes for their insect hosts (41). This is the first example of a CRISPR-Cas system modulating a mutualistic and tripartite interaction, and sheds light on the underexplored complexity of CRISPR-Cas functions in broader ecological niches. Interestingly, this CRISPR-Cas system is expressed at a higher level in iron limiting conditions, furthering the concept that these machineries respond to extracellular changes and to events that are tightly regulated at the bacterial envelope (41). Additionally, the role of the crRNA from this system in colonization is independent of the effector Cas3, suggesting that the NilD CRISPR RNA has a unique function not involving canonical CRISPR-Cas activity (41). Further studies to elucidate the molecular mechanism of NilD-mediated nematode colonization will shed light not only on envelope changes that facilitate colonization, but also on how orphan crRNAs can potentially function as regulatory elements.

Similar to NilD, it was observed that the *cas2* gene of the type II-B CRISPR-Cas system of *Legionella pneumophila* was required for intracellular survival within amoebae, and that *cas2*

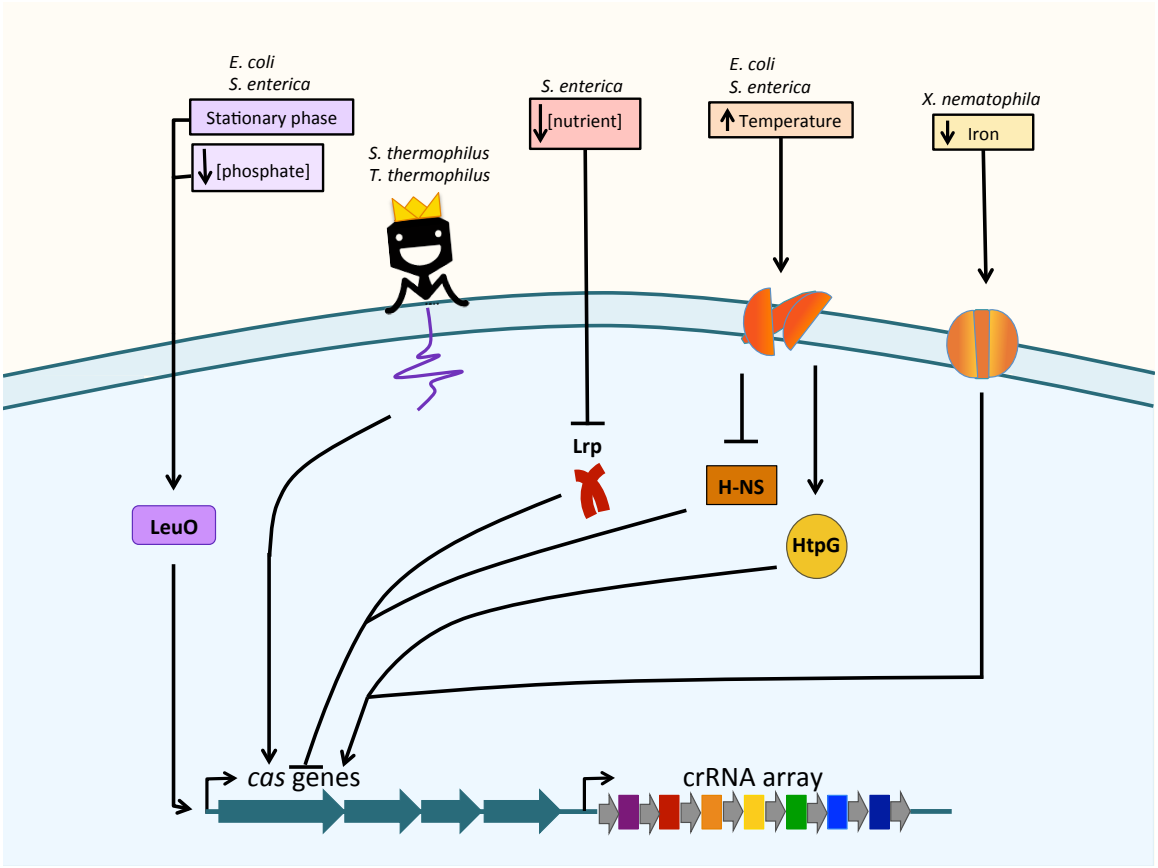

E. coli
S. enterica
Stationary phase
[phosphate]
S. thermophilus
T. thermophilus
S. enterica
[nutrient]
E. coli
S. enterica
Temperature
X. nematophila
Iron
Lrp
H-NS
HtpG
LeuO
cas genes
crRNA array

Figure 2. CRISPR-Cas mediated physiological changes. CRISPR-Cas systems influence bacterial physiology, altering population behavior and host-microbe interactions through events that are centered at the envelope. In *Francisella novicida*, Cas9, tracrRNA and scaRNA form a complex that represses a bacterial lipoprotein mRNA (BLP) (Correction of original publication by HKR 2019: Repression of the BLP mRNA levels occurs through transcriptional repression from the DNA, rather than interaction with the mRNA as depicted herein). Repression of the BLP increases membrane integrity, conferring resistance to membrane targeting antibiotics and enabling evasion of the host immune system, increasing virulence. Cas9 from *Neisseria meningitidis* and Cas2 from *Legionella pneumophila* type II systems increase host-cell attachment and intracellular survival. In *Xenorhabdus nematophila,* Cas6 and a CRISPR RNA (crRNA) of the type I-E system are required for host colonization. In *Myxococcus xanthus*, the type III CRISPR-Cas system regulates exopolysacchride production (EPS) to enable chemotaxis, while negatively effecting fruiting body formation. Conversely, Cas5, Cas7, and Cas8 of its type III CRISPR-Cas system are necessary for fruiting body formation and sporulation. Finally, in *Pseudomonas aeruginosa*, all interference components of the Type I CRISPR system are required for biofilm formation and swarming motility. These examples provide a framework for understanding the alternative functions of CRISPR-Cas systems from interactions at the prokaryotic envelope.

was upregulated during intra-amoeba growth (42). Interestingly, no other *cas* gene was required, and *cas2* was not required for growth in broth culture or intracellular infection of macrophages (42). Furthermore, expression of *cas2* in a *L. pneumophila* strain that lacks a CRISPR-Cas system increased the strain's ability to replicate within amoebae, further indicating that Cas2 can act independently of canonical CRISPR-Cas function (43). Cas2 orthologs have RNase and/or DNase activity, depending on the organism, and are involved in spacer acquisition (17, 44-47). Cas2 nuclease activity is dependent on a single catalytic residue, which is also required for *L. pneumophila* intra-ameobal survival (43). In. *L. pneumophila*, not only is Cas2 RNase activity more efficient than DNase activity, but each requires a different divalent ion (Mg^{2+} or Mn^{2+}, respectively)(43). Thus, preferred nuclease activity may change with shifts in the bacterial environment. It is unclear which nuclease activity promotes survival in amoebae, and a comparison of the ion concentrations in different growth environments may shed light on this difference. Likewise, the precise role of Cas2 in promoting intracellular survival is still unknown; it is tempting to consider that Cas2 has functions in mRNA regulation, particularly given that residues in its nuclease motif are essential for its role in intra-amoeba survival. Studies to observe which nucleic acids associate with Cas2 in different stages and contexts of *Legionella* growth, as well as determining the environmental cues governing the independent regulation of this Cas protein, will significantly enhance the understanding of CRISPR-Cas function as a regulator of intracellular survival.

Are CRISPR-Cas systems more broadly involved in stress responses?

Intriguingly, CRISPR-Cas systems are also regulated by a broad range of environmental conditions not necessarily linked to envelope stress (Figure 1). For instance, in nutrient rich

conditions, the leucine-responsive protein (Lrp) represses CRISPR-Cas expression in *Salmonella enterica* serovar Typhi (24). However, upon starvation, Lrp is inactivated and may de-repress CRISPR-Cas transcription (24). Additionally, the regulator LeuO is an activator of CRISPR-Cas expression in *S. enterica* and *Escherichia coli* (24, 25, 48). LeuO is active under low phosphate and stationary phase conditions, further suggesting that starvation responses can increase CRISPR-Cas expression (49, 50). It is interesting to speculate that expression of CRISPR-Cas systems may also be tied to nutrient conditions since prokaryotic organisms may actively seek out nucleic acids as a nutrient source (51). While starvation is a stress in itself, it can indirectly result in dysregulation of membrane composition, as well as serve as a signal for prophages to become lytic (52, 53). The same is true for oxidative and osmotic stress, which have been shown to activate CRISPR-Cas systems and cause broad stress to the cell, including at the membrane (54, 55). Therefore, it is unclear whether there is a universal link between induction of CRISPR-Cas systems and envelope stress, or if these machineries may more broadly be induced by diverse stresses. In total, these examples provide further links between CRISPR-Cas activation and the response to environmental cues, which may occur through either their canonical or alternative functions.

In addition, CRISPR-Cas systems may act to regulate the cell's response against other diverse environmental stresses (38, 41, 54-57). For example, in *E. coli*, both the CRISPR array and Cas1 can participate in mediating DNA repair, while in *Thermoproteus tenax*, a CRISPR-Cas system is activated in response to DNA damaging UV light (55, 56). Therefore, CRISPR-Cas systems may be responsible for alleviating the effects of stresses that damage the chromosome. In another example, the orphan CRISPR locus in *Listeria monocytogenes*, *rliB*, acts to upregulate the

production of the iron transport system *feoAB*, further demonstrating that CRISPR-Cas systems mediate physiological changes that are likely in response to environmental stress (57). Overall, these observations demonstrate that CRISPR-Cas systems may have evolved multiple functions to not only be activated in response to diverse environmental stress, but also to play active roles in preventing stress-promoted damage.

Conclusion

CRISPR-Cas systems are complex machineries that act to protect the cell against potentially harmful mobile genetic elements. As such, it would be efficient to regulate expression of these systems to times when the threat of such elements is imminent. Accordingly, there are now multiple examples of increased activation of CRISPR-Cas systems in response to envelope stress, such as bacteriophage binding and envelope disruption, ultimately enabling cells to activate defenses against potential genetic threats.

We have summarized numerous examples of CRISPR-Cas systems having functions beyond defense against foreign nucleic acids, many of which involve regulation of envelope physiology and how the cell interacts with its host and environment. It is interesting to consider how these non-canonical functions may have arisen. These observed roles could have appeared due to independent pressures, or stochastically due to accidental acquisition of spacers targeting self. Furthermore, the relationships between CRISPR-Cas system subtype and their non-canonical functions are poorly understood. Since some bacterial species encode multiple CRISPR-Cas subtypes within the same genome, each unique system may represent a fine-tuning of nucleic acid defense, perhaps based on niche and environmental cues. Alternatively, the presence of multiple systems may be linked to non-canonical functions, whereby some systems are

preferentially used for nucleic acid defense and others to regulate bacterial physiology, or multiple systems facilitate different non-canonical functions. We hypothesize that clues to these interactions lie at the envelope, and that by studying the non-canonical functions of CRISPR-Cas systems from this perspective, we will gain insight into the evolution of both commensal and pathogenic bacteria to defend against their own pathogens and survive within their diverse replicative niches.

Acknowledgments

Due to the rapidly expanding field, we have undoubtedly omitted some relevant studies. We apologize in advance to those authors whose work we did not cite.

Financial support and sponsorship

DSW is supported by NIH grant R01-AI110701 and a Burroughs Wellcome Fund Investigator in the Pathogenesis of Infectious Disease award.

Conflicts of interest

TRS and DSW have filed provisional patents based on CRISPR-Cas technological applications. HKR has no conflict of interest.

References

1. Makarova KS, Haft DH, Barrangou R, Brouns SJJ, Charpentier E, Horvath P, et al. Evolution and classification of the CRISPR–Cas systems. Nat Rev Micro. 2011;9(6):467-77.

2. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. Science. 2007;315(5819):1709-12.

3. Marraffini LA, Sontheimer EJ. CRISPR Interference Limits Horizontal Gene Transfer in Staphylococci by Targeting DNA. Science. 2008;322(5909):1843-5.

4. Wiedenheft B, Sternberg SH, Doudna JA. RNA-guided genetic silencing systems in bacteria and archaea. Nature. 2012;482(7385):331-8.

5. Hale CR, Zhao P, Olson S, Duff MO, Graveley BR, Wells L, et al. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. Cell. 2009;139(5):945-56.

*6. van der Oost J, Westra ER, Jackson RN, Wiedenheft B. Unravelling the structural and mechanistic basis of CRISPR-Cas systems. Nat Rev Micro. 2014;12(7):479-92.

A complete resource describing the molecular action of CRISPR-Cas systems.

7. Westra ER, Buckling A, Fineran PC. CRISPR-Cas systems: beyond adaptive immunity. Nat Rev Micro. 2014;12(5):317-26.

8. Makarova KS, Wolf YI, Koonin EV. Comparative genomics of defense systems in archaea and bacteria. Nucleic Acids Res. 2013;41(8):4360-77.

9. Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. Journal of molecular evolution. 2005;60(2):174-82.

10. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology (Reading, England). 2005;151(Pt 8):2551-61.

11. Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, Snijders AP, et al. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science. 2008;321(5891):960-4.

12. Pougach K, Semenova E, Bogdanova E, Datsenko KA, Djordjevic M, Wanner BL, et al. Transcription, processing and function of CRISPR cassettes in Escherichia coli. Mol Microbiol. 2010;77(6):1367-79.

13. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature. 2011;471(7340):602-7.

*14. Zhang Y, Heidrich N, Ampattu BJ, Gunderson CW, Seifert HS, Schoen C, et al. Processing-Independent CRISPR RNAs Limit Natural Transformation in Neisseria meningitidis. Molecular cell. 2013;50(4):488-503.

An elegant demonstration of a unique CRISPR system, whereby individual crRNAs contain their own promoter, rather than being processed from a single transcript.

*15. Chylinski K, Makarova KS, Charpentier E, Koonin EV. Classification and evolution of type II CRISPR-Cas systems. Nucleic Acids Res. 2014;42(10):6091-105.

A comprehensive resource and forward thinking analysis of type II CRISPR system.

**16. Chylinski K, Le Rhun A, Charpentier E. The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. RNA Biology. 2013;10(5):726-37.

A complete analysis of known type II systems and their molecular functions

*17. Heler R, Marraffini LA, Bikard D. Adapting to new threats: the generation of memory by CRISPR-Cas immune systems. Molecular Microbiology. 2014;93(1):1-9.

An inclusive resource for the molecular mechanism of CRISPR adaptation

18. Darwin AJ. Stress Relief during Host Infection: The Phage Shock Protein Response Supports Bacterial Virulence in Various Ways. PLoS Pathog. 2013;9(7):e1003388.

19. Fallico V, Ross RP, Fitzgerald GF, McAuliffe O. Genetic response to bacteriophage infection in Lactococcus lactis reveals a four-strand approach involving induction of membrane stress proteins, D-alanylation of the cell wall, maintenance of proton motive force, and energy conservation. Journal of virology. 2011;85(22):12032-42.

20. Raivio T. Identifying your enemies – could envelope stress trigger microbial immunity? Molecular Microbiology. 2011;79(3):557-61.

21. Young JC, Dill BD, Pan C, Hettich RL, Banfield JF, Shah M, et al. Phage-induced expression of CRISPR-associated proteins is revealed by shotgun proteomics in Streptococcus thermophilus. PloS one. 2012;7(5):e38077.

22. Agari Y, Sakamoto K, Tamakoshi M, Oshima T, Kuramitsu S, Shinkai A. Transcription profile of Thermus thermophilus CRISPR systems after phage infection. Journal of molecular biology. 2010;395(2):270-81.

23. Perez-Rodriguez R, Haitjema C, Huang Q, Nam KH, Bernardis S, Ke A, et al. Envelope stress is a trigger of CRISPR RNA-mediated DNA silencing in Escherichia coli. Mol Microbiol. 2011;79(3):584-99.

24. Medina-Aparicio L, Rebollar-Flores JE, Gallego-Hernandez AL, Vazquez A, Olvera L, Gutierrez-Rios RM, et al. 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. 2011;193(10):2396-407.

25. Westra ER, Pul U, Heidrich N, Jore MM, Lundgren M, Stratmann T, et al. H-NS-mediated repression of CRISPR-based immunity in Escherichia coli K12 can be relieved by the transcription activator LeuO. Mol Microbiol. 2010;77(6):1380-93.

26. Bardwell JC, Craig EA. Eukaryotic Mr 83,000 heat shock protein has a homologue in Escherichia coli. Proceedings of the National Academy of Sciences. 1987;84(15):5177-81.

27. Yosef I, Goren MG, Kiro R, Edgar R, Qimron U. High-temperature protein G is essential for activity of the Escherichia coli clustered regularly interspaced short palindromic repeats (CRISPR)/Cas system. Proceedings of the National Academy of Sciences. 2011;108(50):20136-41.

28. Loesche WJ. Role of Streptococcus mutans in human dental decay. Microbiological Reviews. 1986;50(4):353-80.

29. Senadheera D, Krastel K, Mair R, Persadmehr A, Abranches J, Burne RA, et al. Inactivation of VicK affects acid production and acid survival of Streptococcus mutans. J Bacteriol. 2009;191(20):6415-24.

**30. Serbanescu MA, Cordova M, Krastel K, Flick R, Beloglazova N, Latos A, et al. Role of the Streptococcus mutans CRISPR/Cas systems in immunity and cell physiology. J Bacteriol. 2014.

A new study which observed unique roles for the S. mutans CRISPR systems in tolerance to temperature, oxidative, and envelope stress

31. Boysen A, Ellehauge E, Julien B, Sogaard-Andersen L. The DevT protein stimulates synthesis of FruA, a signal transduction protein required for fruiting body morphogenesis in Myxococcus xanthus. J Bacteriol. 2002;184(6):1540-6.

32. Thony-Meyer L, Kaiser D. devRS, an autoregulated and essential genetic locus for fruiting body development in Myxococcus xanthus. J Bacteriol. 1993;175(22):7450-62.

33. Viswanathan P, Murphy K, Julien B, Garza AG, Kroos L. Regulation of dev, an operon that includes genes essential for Myxococcus xanthus development and CRISPR-associated genes and repeats. J Bacteriol. 2007;189(10):3738-50.

**34. Wallace RA, Black WP, Yang X, Yang Z. A CRISPR with roles in Myxococcus xanthus development and exopolysaccharide production. J Bacteriol. 2014;196(23):4036-43.

A recent study demonstrating a regulatory function of the M. xanthus CRISPR systems in the regulation of exopolysaccharide production, fruiting body formation, and motility.

35. Zegans ME, Wagner JC, Cady KC, Murphy DM, Hammond JH, O'Toole GA. Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of Pseudomonas aeruginosa. J Bacteriol. 2009;191(1):210-9.

36. Cady KC, O'Toole GA. Non-identity-mediated CRISPR-bacteriophage interaction mediated via the Csy and Cas3 proteins. J Bacteriol. 2011;193(14):3433-45.

37. Cady KC, Bondy-Denomy J, Heussler GE, Davidson AR, O'Toole GA. The CRISPR/Cas adaptive immune system of Pseudomonas aeruginosa mediates resistance to naturally occurring and engineered phages. J Bacteriol. 2012;194(21):5728-38.

38. Sampson TR, Saroj SD, Llewellyn AC, Tzeng Y-L, Weiss DS. A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. Nature. 2013;497(7448):254-7.

39. Sampson TR, Napier BA, Schroeder MR, Louwen R, Zhao J, Chin CY, et al. A CRISPR-Cas system enhances envelope integrity mediating antibiotic resistance and inflammasome evasion. Proceedings of the National Academy of Sciences of the United States of America. 2014;111(30):11163-8.

**40. Louwen R, Horst-Kreft D, de Boer AG, van der Graaf L, de Knegt G, Hamersma M, et al. A novel link between Campylobacter jejuni bacteriophage defence, virulence and

Guillain-Barre syndrome. European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology. 2013;32(2):207-26.

The first demonstration that components of a CRISPR system are important for traits involved in virulence.

**41. Veesenmeyer JL, Andersen AW, Lu X, Hussa EA, Murfin KE, Chaston JM, et al. NilD CRISPR RNA contributes to Xenorhabdus nematophila colonization of symbiotic host nematodes. Mol Microbiol. 2014;93(5):1026-42.

The first demonstration that components of a CRISPR system are involved in colonization of a symbiotic bacterium.

**42. Gunderson FF, Cianciotto NP. The CRISPR-associated gene cas2 of Legionella pneumophila is required for intracellular infection of amoebae. mBio. 2013;4(2):e00074-13.

An interesting observation that a component of the adaptive machinery of CRISPR systems is required for intracellular survival.

**43. Gunderson FF, Mallama CA, Fairbairn SG, Cianciotto NP. The Nuclease Activity of Legionella pneumophila Cas2 Promotes Intracellular Infection of Amoebal Host Cells. Infection and immunity. 2014.

A study describing the molceular action that Cas2 may require to mediate Legionella survival in amoeba.

44. Beloglazova N, Brown G, Zimmerman MD, Proudfoot M, Makarova KS, Kudritska M, et al. A novel family of sequence-specific endoribonucleases associated with the clustered regularly interspaced short palindromic repeats. The Journal of biological chemistry. 2008;283(29):20361-71.

45. Nam KH, Ding F, Haitjema C, Huang Q, DeLisa MP, Ke A. Double-stranded endonuclease activity in Bacillus halodurans clustered regularly interspaced short palindromic repeats (CRISPR)-associated Cas2 protein. The Journal of biological chemistry. 2012;287(43):35943-52.

46. Ka D, Kim D, Baek G, Bae E. Structural and functional characterization of Streptococcus pyogenes Cas2 protein under different pH conditions. Biochemical and biophysical research communications. 2014;451(1):152-7.

47. Yosef I, Goren MG, Qimron U. Proteins and DNA elements essential for the CRISPR adaptation process in Escherichia coli. Nucleic Acids Res. 2012;40(12):5569-76.

48. Hernandez-Lucas I, Gallego-Hernandez AL, Encarnacion S, Fernandez-Mora M, Martinez-Batallar AG, Salgado H, et al. The LysR-type transcriptional regulator LeuO controls expression of several genes in Salmonella enterica serovar Typhi. J Bacteriol. 2008;190(5):1658-70.

49. VanBogelen RA, Olson ER, Wanner BL, Neidhardt FC. Global analysis of proteins synthesized during phosphorus restriction in Escherichia coli. J Bacteriol. 1996;178(15):4344-66.

50. Stratmann T, Pul U, Wurm R, Wagner R, Schnetz K. RcsB-BglJ activates the Escherichia coli leuO gene, encoding an H-NS antagonist and pleiotropic regulator of virulence determinants. Mol Microbiol. 2012;83(6):1109-23.

51. Finkel SE, Kolter R. DNA as a Nutrient: Novel Role for Bacterial Competence Gene Homologs. Journal of Bacteriology. 2001;183(21):6288-93.

52. Death A, Notley L, Ferenci T. Derepression of LamB protein facilitates outer membrane permeation of carbohydrates into Escherichia coli under conditions of nutrient stress. J Bacteriol. 1993;175(5):1475-83.

53. Hansen AM, Gu Y, Li M, Andrykovitch M, Waugh DS, Jin DJ, et al. Structural basis for the function of stringent starvation protein a as a transcription factor. The Journal of biological chemistry. 2005;280(17):17380-91.

54. Strand K, Sun C, Li T, Jenney F, Jr., Schut G, Adams MW. Oxidative stress protection and the repair response to hydrogen peroxide in the hyperthermophilic archaeon Pyrococcus furiosus and in related species. Arch Microbiol. 2010;192(6):447-59.

55. Plagens A, Tjaden B, Hagemann A, Randau L, Hensel R. Characterization of the CRISPR/Cas subtype I-A system of the hyperthermophilic crenarchaeon Thermoproteus tenax. J Bacteriol. 2012;194(10):2491-500.

56. Babu M, Beloglazova N, Flick R, Graham C, Skarina T, Nocek B, et al. A dual function of the CRISPR–Cas system in bacterial antivirus immunity and DNA repair. Molecular Microbiology. 2011;79(2):484-502.

57. Mandin P, Repoila F, Vergassola M, Geissmann T, Cossart P. Identification of new noncoding RNAs in Listeria monocytogenes and prediction of mRNA targets. Nucleic Acids Res. 2007;35(3):962-74.

Chapter 3. Catalytically active Cas9 mediates transcriptional interference to facilitate bacterial virulence

Hannah K. Ratner[1,2,3], Andrés Escalera-Maurer[4,5,6], Anaïs Le Rhun[4,5,6], Siddharth Jaggavarapu[2,3,7], Jessie E. Wozniak[1,2,3], Emily K. Crispell[1,2,3], Emmanuelle Charpentier[4,5,6], David S. Weiss [1,2,3,7]

[1]Microbiology and Molecular Genetics Program, Emory University, Atlanta, GA 30329.

[2]Emory Vaccine Center, Emory University, Atlanta, GA 30329.

[3]Yerkes National Primate Research Center, Emory University, Atlanta, GA 30329.

[4]Max Planck Unit for the Science of Pathogens, D-10117 Berlin, Germany

[5]Helmholtz Centre for Infection Research, Department of Regulation in Infection Biology, D-38124, Braunschweig, Germany

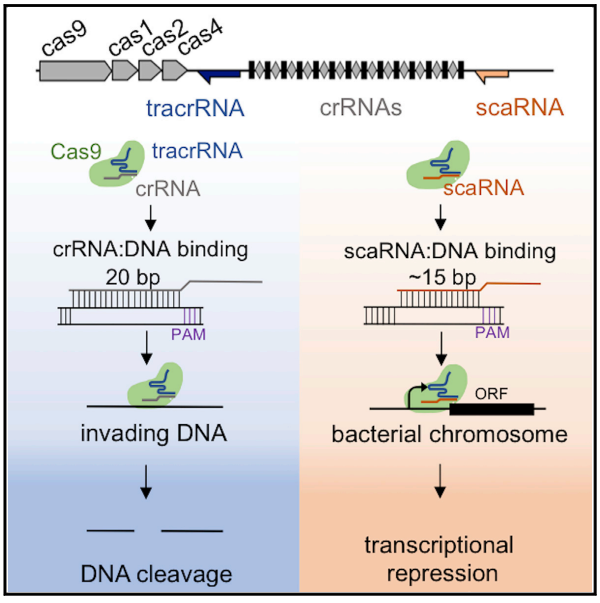
[6] Institute for Biology, Humboldt University, D-10115 Berlin, Germany

[7]Division of Infectious Diseases, Department of Medicine, Emory University School of Medicine, Atlanta, Georgia 30329.

Published in Molecular cell, 2019

Abstract

In addition to defense against foreign DNA, the CRISPR-Cas9 system of *Francisella novicida* represses expression of an endogenous immunostimulatory lipoprotein. We investigated the specificity and molecular mechanism of this regulation, demonstrating that Cas9 controls a highly specific regulon of four genes which must be repressed for bacterial virulence. Regulation occurs through a PAM-dependent interaction of Cas9 with its endogenous DNA targets, dependent on a non-canonical small RNA (scaRNA) and tracrRNA. The limited complementarity between scaRNA and the endogenous DNA targets precludes cleavage, highlighting the evolution of scaRNA to repress transcription without lethally targeting the chromosome. We show that scaRNA can be reprogrammed to repress other genes, and with engineered, extended complementarity to an exogenous target, the repurposed scaRNA:tracrRNA-FnoCas9 machinery can also direct DNA cleavage. Natural Cas9 transcriptional interference likely represents a broad paradigm of regulatory functionality, which is potentially critical to the physiology of numerous Cas9-encoding pathogenic and commensal organisms.



Highlights

- FnoCas9 (*Francisella novicida* Cas9) uses scaRNA to bind endogenous DNA and repress transcription
- The limited length of scaRNA:target complementarity prevents DNA cleavage
- Cleavage-competent FnoCas9 uses distinct RNAs for repression versus cleavage
- scaRNA can be reprogrammed to guide FnoCas9 to repress a new target

Introduction

CRISPR-Cas systems are prokaryotic adaptive immune systems that restrict infection by potentially harmful foreign genetic elements (Barrangou et al., 2007; Marraffini and Sontheimer, 2008). CRISPR-Cas9 systems use the single effector protein Cas9 for target recognition and cleavage (Gasiunas et al., 2012; Jinek et al., 2012). Cas9 forms a complex with a duplex of small RNAs, one being a crRNA that is transcribed and processed from a genomic CRISPR (clustered regularly interspaced short palindromic repeats) array (Deltcheva et al., 2011). The other is tracrRNA, a small RNA transcribed from the CRISPR-Cas9 locus that contains an inverted sequence with complementarity to the repeat sequence that is conserved in each crRNA derived from a CRISPR array (Deltcheva et al., 2011). Another portion of the crRNA, the spacer sequence, is often complementary to an exogenous DNA target (Bolotin et al., 2005; Mojica et al., 2005; Pourcel et al., 2005). Upon infection with a nucleic acid, the crRNA spacer binds to the complementary sequence on the incoming DNA (the protospacer), leading Cas9 to cleave the DNA target, resulting in a double strand break (Gasiunas et al., 2012; Jinek et al., 2012).

Before the crRNA spacer can interact with the complementary sequence on the DNA target, the Cas9 complex must first recognize a short nucleotide sequence on the opposite strand and adjacent to the protospacer, called a protospacer adjacent motif (PAM) (Anders et al., 2014; Jinek et al., 2012; Jinek et al., 2014; Nishimasu et al., 2014; Sapranauskas et al., 2011). This stage of target recognition is necessary to prevent cleavage of the genomic CRISPR array from which the crRNAs are transcribed (Gasiunas et al., 2012; Jinek et al., 2012). There are no PAM sequences next to the crRNA spacers in the genome (Marraffini and Sontheimer, 2010; Mojica et al., 2009).

Interestingly, CRISPR-Cas systems can have additional roles in bacterial physiology that extend beyond defense against foreign nucleic acids (Louwen et al., 2014; Ratner et al., 2015; Westra et al., 2014). Of particular interest are type II CRISPR-Cas systems, which include CRISPR-Cas9, because they are uniquely abundant in pathogenic and commensal organisms (Chylinski et al., 2013; Fonfara et al., 2014; Sampson et al., 2013). Mutants lacking *cas9* in *Streptococcus agalactiae, Campylobacter jejuni, Neisseria meningitidis,* and *Francisella novicida* are impaired in virulence processes such as attachment and invasion of host cells and intracellular survival (Louwen et al., 2013; Ma et al., 2018; Sampson et al., 2013). These defects correspond to reduced pathogenicity of *cas9* mutants in *S. agalactiae, C. jejuni*, and *F. novicida* during infection (Louwen et al., 2013; Ma et al., 2018; Sampson et al., 2013).

In *F. novicida*, which can cause human infections, the attenuation of the *cas9* mutant is due in part to *F. novicida* Cas9 (FnoCas9) regulation of the expression of an endogenous mRNA encoding a bacterial lipoprotein (BLP), *FTN_1103* (*1103*) (Jones et al., 2012). Regulation of *1103* by FnoCas9 is controlled by tracrRNA and scaRNA, which is a distinct small RNA transcribed from an independent promoter near the CRISPR locus (Chylinski et al., 2013; Postic et al., 2010; Sampson et al., 2013). Interestingly, FnoCas9, tracrRNA, and scaRNA together enable robust

repression of *1103* transcript levels, which occurs in the presence and absence of crRNAs (Sampson et al., 2013). Since BLPs are ligands for mammalian innate immune proteins, repression of *1103* helps facilitate the evasion of these sensors by *F. novicida* (Jones et al., 2012; Sampson et al., 2014; Sampson et al., 2013). However, reduction of *1103* levels alone is not sufficient to completely restore the virulence of a *cas9* mutant, suggesting that additional factors are involved (Jones et al., 2012; Sampson et al., 2013).

To comprehensively characterize the endogenous regulatory role of FnoCas9, we performed a genome-wide expression analysis which revealed that FnoCas9 has a specific regulon of just two transcripts encoding four genes, including *1103*. Regulation is PAM-dependent and uses catalytically active FnoCas9 and two RNAs, scaRNA and tracrRNA, that likely form an RNA duplex (scaRNA:tracrRNA) that is distinct from the duplex used for targeting foreign DNA (crRNA:tracrRNA). scaRNA is complementary to the template strand of the 5' UTRs of the two transcripts, thus targeting FnoCas9 to specific sites on the endogenous genomic DNA to repress transcript levels. Repression of all four genes contributes to the virulence of *F. novicida*. These findings show for the first time that a cleavage-competent Cas9 complex can exist in two distinct states in the bacterium to mediate two different functions: binding to endogenous DNA as a transcriptional repressor and cleaving foreign DNA to prevent infection. We further demonstrate that the scaRNA can be reprogrammed to guide FnoCas9 to repress other genes in *F. novicida*, highlighting the potential utility of this system in the control of gene expression. Taken together, these findings likely represent a broader paradigm in the way CRISPR-Cas9 systems mediate non-canonical functions that are distinct from DNA cleavage and contribute to bacterial physiology.

Results

FnoCas9 has a highly specific regulon

To identify regulatory targets of FnoCas9, we performed a genome-wide analysis of mRNA levels in a *cas9* deletion mutant compared to wild-type (WT) *F. novicida* (excluding small RNAs). Only 4 out of 1,782 genes (0.22%) were significantly up-regulated in the *cas9* mutant by more than 2-fold, including *FTN_1103* (*1103*) encoding a bacterial lipoprotein (BLP), which we previously demonstrated to be regulated by FnoCas9 (Figure 1A, S1A, C-F) (Jones et al., 2012; Sampson et al., 2013). Similar experiments in tracrRNA and scaRNA deletion strains revealed the identical regulon of only four genes (*1103*; *FTN_1104*, "*1104*"; *FTN_1102*, "*1102*"; *FTN_1101*, "*1101*") (Figure 1A, S1A). These results, validated by Northern blots of *1104-1101* in Δ*cas9,* Δ*tracrRNA*, Δ*caRNA*, Δ*crRNA*, and the complemented strains, indicated that similar to *1103*, repression of *1104*, *1102*, and *1101* transcripts was dependent on FnoCas9, tracrRNA, and scaRNA (Figure 1C-F, S1A). The upregulation of the *1104-1101* transcripts upon deletion of FnoCas9, scaRNA, or tracrRNA did not occur in the *crRNA* mutant (Figure 1C-F, S1A). Interestingly, the four genes are encoded at the same genomic region (Figure 1B, S1A). Northern blots (Figure 1C-F) and PCR amplification over the gene junctions between *1104-1101* (Figure S1B) revealed that they are encoded on two separate transcripts, one containing the operon *1104-1102* and another containing *1101* (Figure 1B-F). These data highlight that the FnoCas9 machinery targets a regulon that is highly specific to a region of the genome.

FnoCas9 represses transcript levels by targeting the 5' UTR of target genes

We next investigated whether the two repressed transcripts contained regions that could serve as targeting sites by the scaRNA:tracrRNA-FnoCas9 machinery. We identified

A

sequenced strain:		Δ*cas9*		Δ*tracrRNA*		Δ*scaRNA*	
Gene	**NCBI locus tag**	**log2FC**	**padj**	**log2FC**	**padj**	**log2FC**	**padj**
1101	FTN_RS05655	2.50	1.28E-24	2.95	7.58E-57	2.19	1.72E-10
1102	FTN_RS05660	*	*	2.50	2.32E-20	2.21	1.69E-08
1103	FTN_RS05665	3.15	3.48E-35	3.84	5.36E-69	3.71	9.55E-34
1104	FTN_RS05670	3.64	5.19E-51	4.07	4.70E-81	3.53	6.68E-27

*expression levels determined by qRT-PCR in Figure S1E

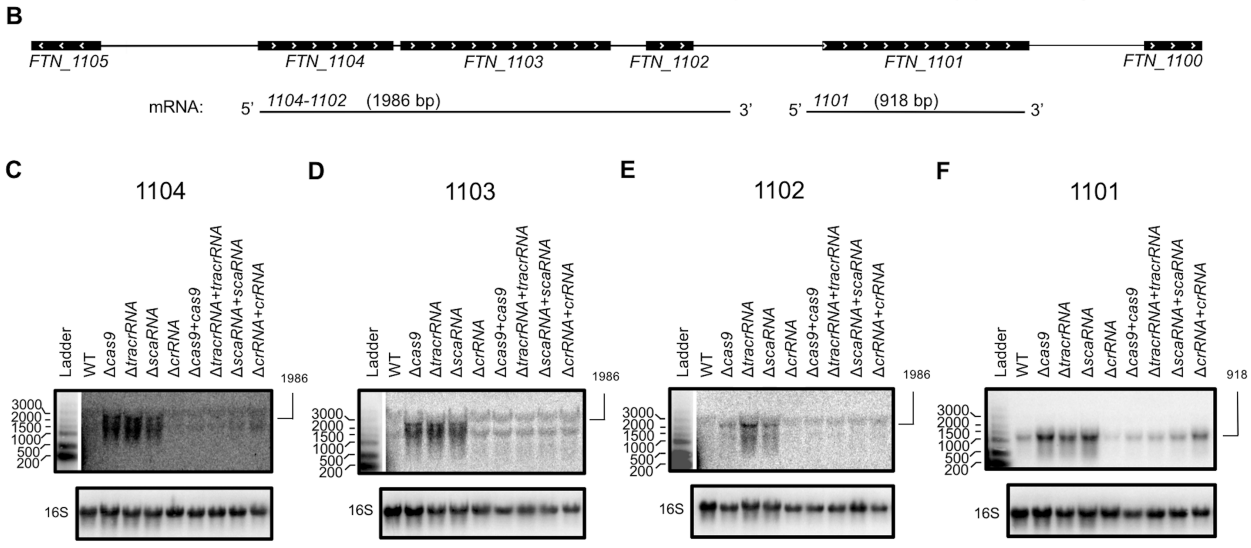


Figure 1. FnoCas9, scaRNA, and tracrRNA regulate transcript levels in a specific genomic region of F. novicida. A) Differential transcript expression in the indicated *F. novicida* deletion mutants (Δ*cas9*, Δ*tracrRNA*, and Δ*scaRNA*) relative to expression in wild-type (WT). Table represents all genes with a log fold change (log2FC) > 1 and an adjusted p-value (padj) <0.05 in each strain compared to WT. The names used for these genes throughout the paper are *1104*, *1103*, *1102*, and *1101*. The genome of *F. novicida* has recently been re-annotated in NCBI and the new locus tags are indicated. B) Schematic of the chromosomal locus encoding the four FnoCas9-regulated genes (*1104*, *1103*, *1102*, and *1101*) and the two mRNA products of the locus. C-F) Northern blots for (C) *1104*, (D) *1103*, (E) *1102*, (F) *1101*, in wild-type (WT) and deletion mutants for each component of the *F. novicida* CRISPR-Cas9 system (**Δ***cas9*, **Δ***tracrRNA*, **Δ***scaRNA*, **Δ***crRNA*) and their respective complementation strains. The ladders in C-E were from a different exposure, indicated by the white separating line.

a 17 bp region in both the *1104* and *1101* 5' UTRs with 100% sequence identity as a potential site of repression (Figure 2A, S2A). To test whether the genomic regions encoding the 5' UTRs can confer FnoCas9-dependent repression, we constructed a series of chromosomal promoter fusions driving expression of a non-native sequence *gfp**, replacing the *1104-1101* locus in the genome, in strain backgrounds with Cas9 (Cas9+) and without Cas9 (Cas9-) (Figure 2B, S3A). *gfp** was placed directly downstream of the *1104* promoter and the *1104* 5' UTR. Expression of *gfp** was greatly enhanced in the *cas9* mutant as compared to WT, consistent with a role for FnoCas9 in mediating repression (Figure 2C). No such repression of *gfp** was observed in the WT strain from a similar reporter construct lacking the *1104* 5' UTR (Figure 2C). These data indicate that the genomic region encoding the *1104* 5' UTR can direct a non-native transcript to be under scaRNA:tracrRNA-FnoCas9-dependent repression.

To test whether the *1104* 5' UTR could promote FnoCas9-dependent repression downstream of non-native promoters, we generated chromosomal reporter constructs in which the *1104* promoter was replaced with either p146, a synthetic promoter with constitutive expression in *F. novicida,* or the broad host range T5 promoter from *Eschericia coli* containing a *lac* operator (Bryksin and Matsumura, 2010; McWhinnie and Nano, 2014). Irrespective of the promoter used, the *1104* 5' UTR conferred FnoCas9-dependent repression of *gfp** (Figure 2C). Furthermore, the genomic region encoding the *1101* 5' UTR was also capable of conferring FnoCas9-dependent repression of *gfp** when located downstream of the native *1101*, synthetic, or *lac* promoters (Figure 2D). Taken together, these data demonstrate that the *1104* or *1101* 5' UTRs can direct a transcript to be under FnoCas9 regulatory control independent of the promoter used.

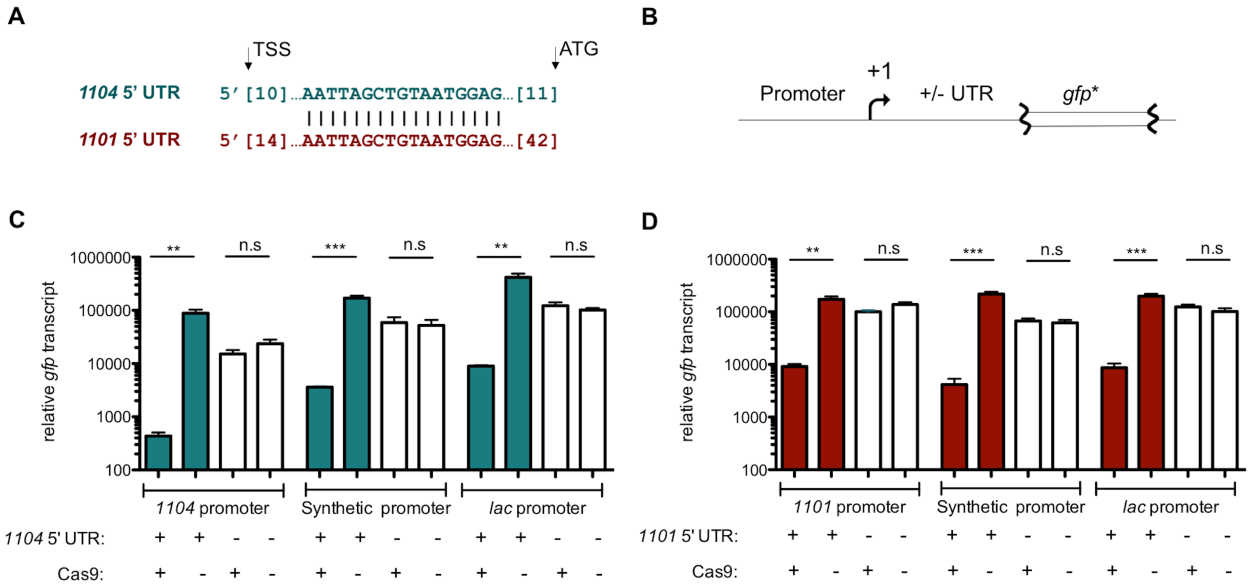


Figure 2. FnoCas9 targets sequences coding for 5' untranslated regions (UTRs) leading to transcriptional interference. A) Alignment of *1104* (teal) and *1101* (red) 5' UTRs with the location of the transcriptional start site (TSS) and start codon (ATG) highlighted. Brackets indicate the number of nucleotides in each UTR flanking the aligned sequence. B) Schematic indicating the design of the fusion constructs used to interrogate the roles of the promoters and 5' UTRs in transcriptional repression by FnoCas9. C-D) Relative *gfp** transcript level measured by quantitative real time PCR (qRT-PCR) from constructs with either the native promoter ("*1104*" in C, and "*1101*" in D), p146 synthetic promoter, or *lac* promoter, and with or without the native 5' UTR (*1104* in C, and *1101* in D), in a wild-type (Cas9+) or *cas9* mutant (Cas9-) background. (n=3, error bars represent s.e.). **p≤0.005; ***p≤0.001.

scaRNA has complementarity to the *1104* and *1101* 5' UTRs

To determine how the FnoCas9 machinery might target the 5' UTRs of *1104* and *1101*, we bioinformatically searched for predicted interactions between the scaRNA or tracrRNA and these regions. We precisely defined the sequences of the scaRNA and tracrRNA using a small RNAseq (sRNA-seq) analysis of WT *F. novicida* (Figure 3A, S2E-F) (Chylinski et al., 2013). The 56 base pair scaRNA contains the degenerated repeat predicted previously, however, the tail of the scaRNA extends to the 3' of the repeat (Sampson et al., 2013). With these new data, we identified a sequence of 11 bases of perfect complementarity between the template strand of the *1104* 5' UTR and the tail of the scaRNA (which was included in a larger region of 15 bases of complementarity over 19 bases in the scaRNA tail; Figure 3A). This site overlapped with the 17 bp stretch of homology between the *1104* and *1101* 5' UTRs (Figure S2A), with the template strand of the *1101* 5' UTR also having 11 bases of perfect complementarity to the scaRNA tail (and 12 bases of complementarity over a region of 15 bases; Figure 3B). This suggests that scaRNA may direct FnoCas9 to the DNA encoding the *1104* and *1101* 5' UTRs since the sequence of the template strand is not encoded on the mRNA.

FnoCas9 uses a PAM to interact with target 5' UTR DNA

The Cas9 complex requires recognition of a PAM sequence in the dsDNA target before the crRNA spacer can interact with its complementary sequence in the DNA. We identified an NGG PAM sequence on the coding strand adjacent to the predicted scaRNA binding site on the template strand of both 5' UTRs (Figure 3A-B). To test if the PAM was required for FnoCas9-dependent repression, we mutated the PAM of the *1104* 5' UTR from TGG to TAA in the *gfp** reporter strain with the *1104* promoter. FnoCas9 lost the ability to repress *gfp** expression upon mutation of the PAM, consistent with the possibility that FnoCas9 binds the *1104* 5' UTR DNA (Figure 3C).

Similarly, we constructed the TAA PAM mutation in the *1101* promoter and 5' UTR reporter strain and observed a loss of FnoCas9-dependent repression (Figure 3D).

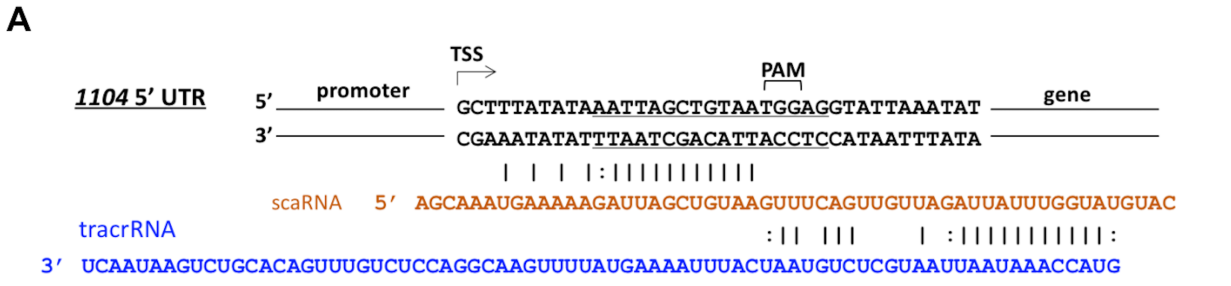
To determine whether scaRNA mediates FnoCas9 DNA binding as suggested by the previous data, and to further investigate whether a PAM is required, we conducted electrophoretic mobility shift assays (EMSAs) with scaRNA:tracrRNA, FnoCas9, and DNA oligonucleotide targets containing the 11 bases of identity to the scaRNA tail shared between the *1104* and *1101* 5' UTRs, and either a WT TGG PAM or a mutated TAA PAM (Figure 3E). FnoCas9 binding to DNA was dependent on scaRNA:tracrRNA as well as the WT PAM, since there was no binding observed to DNA encoding the mutated PAM (Figure 3E). These data suggest that scaRNA interacts with DNA in a PAM-dependent manner.

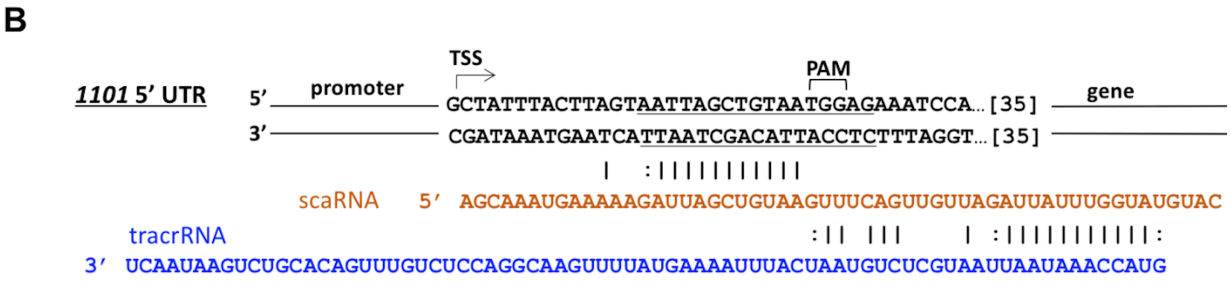
To determine whether FnoCas9 interacts with the *1104* 5' UTR DNA in *F. novicida*, we added a crosslinking reagent to intact cells expressing either a FLAG epitope-tagged WT Cas9 or a FLAG-tagged point mutant (Cas9:R59A-FLAG) that is unable to interact with any of the FnoCas9-associated RNAs due to a mutation in the RNA binding domain. We next immunoprecipitated Cas9-FLAG from bacterial lysates, fragmented and isolated crosslinked DNA, and amplified the *1104* 5' UTR region by qPCR. *1104* was enriched in the WT Cas9-FLAG pulldown relative to the Cas9:R59A-FLAG mutant, demonstrating that FnoCas9 does indeed interact with DNA (Figure S2B). Furthermore, this interaction was dependent on the scaRNA since Cas9-FLAG in a *scaRNA* deletion strain behaved similarly to Cas9:R59A-FLAG in the WT (Figure S2B). *1101* followed the same trend of enrichment in the WT *cas9*-FLAG pulldown compared to the *cas9*:R59A-FLAG and Δ*scaRNA cas9*-FLAG strains (Figure S2C). A control gene, *FTN_0544*, that is not regulated by FnoCas9 (Figure 7B-C) was not enriched in the Cas9-FLAG and Cas9:R59A-FLAG pulldowns compared to WT (Figure S2D). The overall

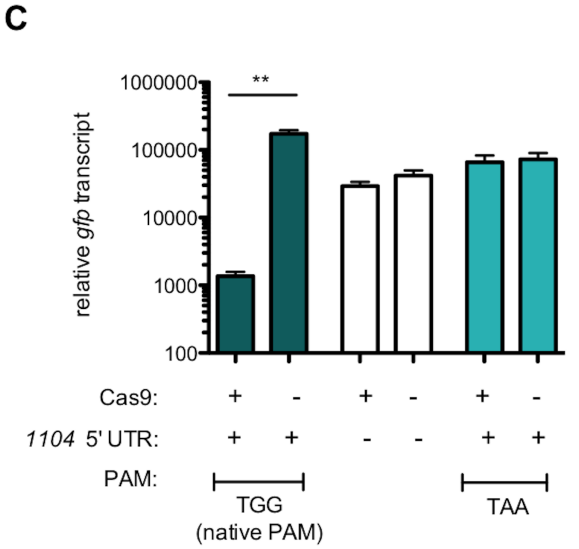
reduced rate of *1101* DNA enrichment compared to *1104* is likely the result of fewer complementary between scaRNA and the *1101* 5' UTR compared to the *1104* 5' UTR, which may reduce the affinity of the pulldown as well as the level of transcriptional repression (Figure S2B-C). Collectively, these data indicate that scaRNA mediates the interaction of FnoCas9 with *1104* and *1101* DNA, representing the first described natural example of Cas9-mediated transcriptional interference.

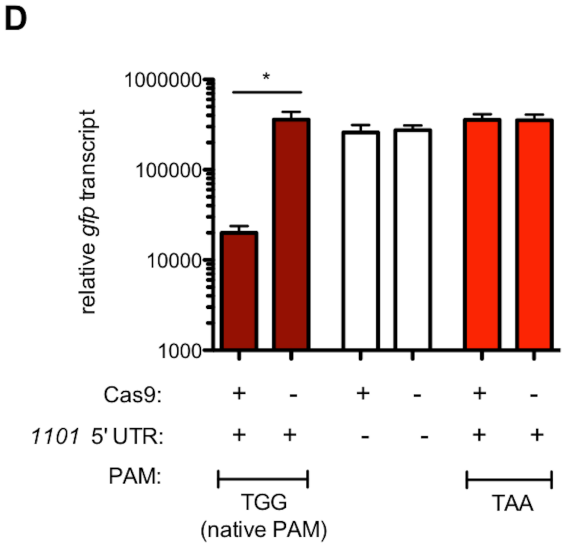
Extent of complementarity to scaRNA modulates transcriptional interference

We next evaluated which factors control the level of transcriptional interference exhibited by FnoCas9. The importance of complementarity between the scaRNA and *1104* 5' UTR was tested by measuring *gfp** transcript level from chromosomal fusion constructs with either 0, 8, 11, or 15 bases of complementarity to scaRNA on the template strand, followed by a PAM, as well as from the equivalent plasmid-based fusion constructs (Figure S3A). Compared to the strain without any complementarity between scaRNA and the *1104* 5' UTR, the highest level of *gfp** repression was observed in the strain with 15 bp of complementarity to scaRNA, with lesser repression in the strain with 11 bp of complementarity (Figure 4A, S4). Repression of *gfp** was alleviated in the strain with only 8 bp of scaRNA complementarity (Figure 4B, S4). These data indicate that 11 to 15 bp of identity between the scaRNA and its target is sufficient for repression, with 15 bp providing the strongest effect (Figure 4A, S4). We mutated the PAM in the construct with 15 bp complementarity to scaRNA, and found that this mutation restored *gfp** levels to that of a construct with 0 bp complementarity to scaRNA (Figure S3B).


A
1104 5′ UTR
5′
promoter
TSS
PAM
GCTTTATATAAATTAGCTGTAATGGAGGTATTAAATAT
gene
3′
CGAAATATATTTAATCGACATTACCTCCATAATTTATA
scaRNA 5′ AGCAAAUGAAAAAGAUUAGCUGUAAGUUUCAGUUGUUAGAUUAUUUGGUAUGUAC
tracrRNA
3′ UCAAUAAGUCUGCACAGUUUGUCUCCAGGCAAGUUUUAUGAAAAUUUACUAAUGUCUCGUAAUUAAUAAACCAUG


B
1101 5′ UTR
5′
promoter
TSS
PAM
GCTATTTACTTAGTAATTAGCTGTAATGGAGAAATCCA...[35]
gene
3′
CGATAAATGAATCATTAATCGACATTACCTCTTTAGGT...[35]
scaRNA 5′ AGCAAAUGAAAAAGAUUAGCUGUAAGUUUCAGUUGUUAGAUUAUUUGGUAUGUAC
tracrRNA
3′ UCAAUAAGUCUGCACAGUUUGUCUCCAGGCAAGUUUUAUGAAAAUUUACUAAUGUCUCGUAAUUAAUAAACCAUG


C
relative gfp transcript
1000000
100000
10000
1000
100
**
Cas9: + - + - + -
1104 5′ UTR: + + - - + +
PAM:
TGG
(native PAM)
TAA


D
relative gfp transcript
1000000
100000
10000
1000
*
Cas9: + - + - + -
1101 5′ UTR: + + - - + +
PAM:
TGG
(native PAM)
TAA

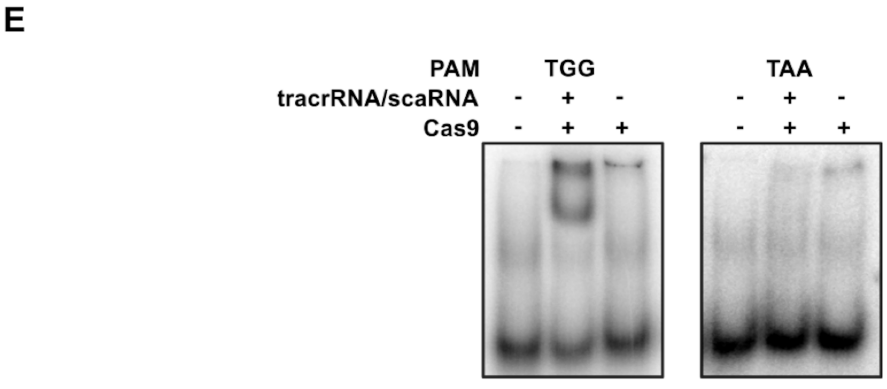

E
PAM
TGG
TAA
tracrRNA/scaRNA - + - - + -
Cas9 - + + - + +

Figure 3. A PAM motif is required for FnoCas9 transcriptional interference. A-B) Schematic of predicted interactions between scaRNA (orange), tracrRNA (blue), and A) the *1104* 5' UTR or B) the *1101* 5' UTR. The underlined region represents the identical sequence conserved between the *1104* and *1101* UTRs. The transcriptional start site (TSS) and PAM are shown on the coding strand of the UTRs. C-D) Relative *gfp** transcript levels were measured by qRT-PCR. (C) Expression of *gfp** is driven by the *1104* promoter and *1104* 5' UTRs containing either a WT TGG PAM sequence or a TAA mutation to the PAM, in strains WT (Cas9+) or Δ*cas9* (Cas9-) strains (n=3, error bars represent s.e., **p≤0.005). (D) Expression of *gfp** is driven by the *1101* promoter and 5' UTRs containing either a WT TGG PAM sequence or a TAA mutation to the PAM, in strains WT (Cas9+) or Δ*cas9* (Cas9-) strains (n=3, error bars represent s.e., *p≤0.05). E) Electromobility shift assays (EMSA) with FnoCas9 (150 nM) and scaRNA:tracrRNA duplex (300 nM), as indicated. A DNA target (1 nM) containing 11 bases of complementarity to scaRNA adjacent to a WT PAM (TGG, 9029/9030; DNA sequences in Table S2) or a mutant PAM (TAA, 9041/9042) was used.

The degree of transcriptional interference correlated with the affinity of FnoCas9 binding. EMSAs revealed that the affinity of the scaRNA:tracrRNA-FnoCas9 complex for the DNA increased with the amount of scaRNA complementarity (8, 11, 15 bp) (Figure 4C-E). The FnoCas9 complex did not bind to DNA with no scaRNA complementarity or in the absence of the scaRNA:tracrRNA duplex (Figure 4B).

To determine if either strand of the DNA could be targeted by FnoCas9 for repression, the reverse complement of the construct with 15 bp scaRNA complementarity was placed between the promoter and *gfp**, resulting in a fusion construct with the PAM and 15 bp of complementarity to scaRNA on the coding strand (Figure S3A). FnoCas9 repressed transcription equally when targeted to the coding and template strands of the DNA (Figure S3C-D).

We also attempted to investigate the ability of scaRNA to bind and repress a construct to which it had 20 bases of complementarity. However, we observed a significantly reduced number of transformants while attempting to make such a strain, and the transformants had mutations in the scaRNA target (Figure S3E). We reasoned that this might be due to lethal targeting by cleavage of chromosomal DNA of recombinants by the scaRNA:tracrRNA-FnoCas9 complex, and set out to test this. First, we validated that the scaRNA:tracrRNA-FnoCas9 complex was indeed able to bind a 20 bp target *in vitro* (Figure 4F). We next used a transformation inhibition model to directly test whether modulating the amount of complementarity between scaRNA and artificial exogenous targets (both a linear allelic exchange fragment and a plasmid) harboring the 5' UTR region with the PAM was sufficient to block transformation. As expected, transformation of constructs containing 8, 11, and 15 bases of complementarity to scaRNA into WT and *cas9* mutant strains resulted in an equal number of transformants. However, when the region of complementarity between the target plasmid and the scaRNA tail was artificially extended to 20 bases, we observed

a significantly reduced number of plasmid transformants in the WT but not the *cas9* mutant strain (Figures 4G). We then constructed mutants with 20 bp of complementarity to scaRNA between a promoter and *gfp** in a d*cas9* strain, in which FnoCas9 has point mutations in the active sites used for DNA cleavage. The mutations present in d*cas9* did not alter the ability of FnoCas9 to repress transcription of *1104* (Figure S3F). dCas9 was able to repress transcription of *gfp** in strains with 20 bases of complementarity to scaRNA, and repression was more efficient than the native repression from the scaRNA interaction with the *1104* 5' UTR (Figure S3G). Finally, we constructed a d*cas9* strain with 20 bases of complementarity to a crRNA. dCas9 repressed transcription using 20 bases of crRNA complementarity at the same level as WT FnoCas9 repressed transcription from the *1104* 5' UTR using scaRNA. Further, there was less repression from both of these constructs than by scaRNA from a 20 bp target with dCas9 (Figure S3G). These results suggest that scaRNA is capable of guiding the FnoCas9 complex to cleave a DNA target only when the extent of complementarity is sufficient, and that repression via binding occurs independently of the cleavage active sites. The viability of *F. novicida* indicates that the native complementarity between scaRNA and the *1104* and *1101* 5' UTRs is not sufficient to cause cleavage of the genomic DNA.

Proximity of the scaRNA binding site to the TSS is required for transcriptional interference

We next investigated the importance of proximity between the scaRNA target in the 5' UTR and the transcriptional start site (TSS). To do this, we measured *gfp** transcript levels from fusion constructs with either 0, 5, 10, or 20 bases between the TSS and the *1104* 5' UTR region

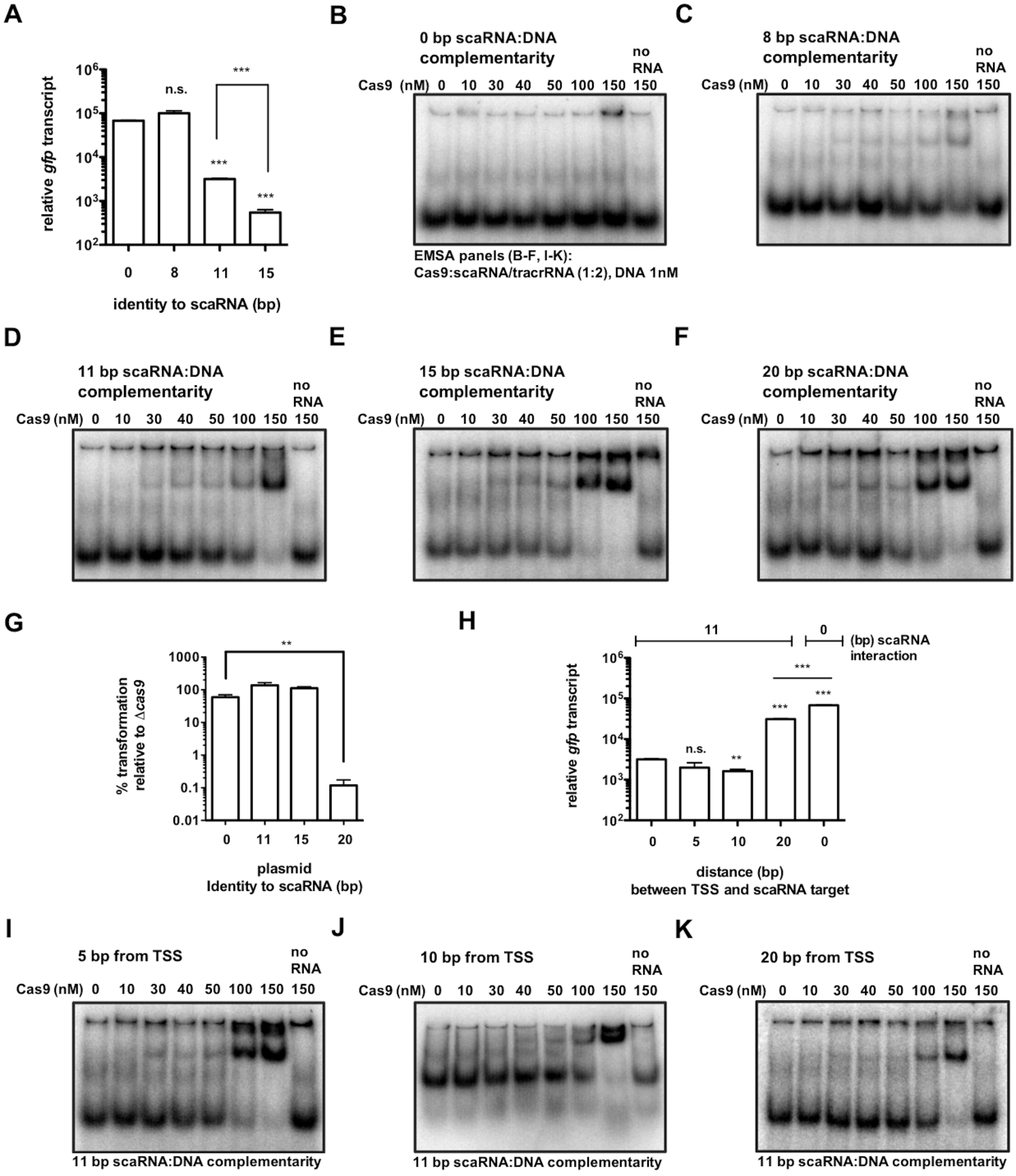

A
relative gfp transcript
n.s.
***
identity to scaRNA (bp)
B
0 bp scaRNA:DNA complementarity
no RNA
Cas9 (nM) 0 10 30 40 50 100 150 150
EMSA panels (B-F, I-K):
Cas9:scaRNA/tracrRNA (1:2), DNA 1nM
C
8 bp scaRNA:DNA complementarity
D
11 bp scaRNA:DNA complementarity
E
15 bp scaRNA:DNA complementarity
F
20 bp scaRNA:DNA complementarity
G
% transformation relative to Δcas9
plasmid
Identity to scaRNA (bp)
H
(bp) scaRNA interaction
relative gfp transcript
distance (bp)
between TSS and scaRNA target
I
5 bp from TSS
11 bp scaRNA:DNA complementarity
J
10 bp from TSS
11 bp scaRNA:DNA complementarity
K
20 bp from TSS
11 bp scaRNA:DNA complementarity

Figure 4. FnoCas9 transcriptional interference is controlled by degree of scaRNA complementarity and target proximity to the TSS. A) Relative *gfp** transcript levels were measured by qRT-PCR from constructs containing sequences with different lengths of complementarity to the scaRNA tail (0, 8, 11, 15 bp) between a synthetic promoter and *gfp** sequence in the chromosome (*n*=3, error bars represent s.e., ***p≤0.001). B-F) Electromobility shift assays (EMSA) with a FnoCas9:scaRNA/tracrRNA complex (1:2 molar ratio FnoCas9:RNA duplex) and DNA target (1 nM) containing different extents of complementarity to scaRNA: (B) 0 bp (9025/9026; DNA sequences in Table S2), (C) 8 bp (9027/9028), (D) 11 bp (9029/9030), (E) 15 bp (9031/9032), (F) 20 bp (9033/9034). G) Plasmid inhibition assay of WT *F. novicida* and Δ*cas9* with plasmids containing PAM-adjacent target sequences with 0, 11, 15, and 20 bases of complementarity to scaRNA between a synthetic promoter and *gfp**. Results are presented as percent transformation into WT relative to Δ*cas9* (*n*=3, error bars represent s.e., **p≤0.005). H) Relative *gfp** transcript levels were measured by qRT-PCR in strains with varying numbers of additional bases (0, 5, 10, 20 bp) placed between the TSS of a synthetic promoter and a sequence with 11 bp of complementarity to the scaRNA, followed by *gfp**. A strain with *gfp** placed downstream of the synthetic promoter and 0 bp of complementarity to scaRNA was used as a control (*n*=3, error bars represent s.e., **p≤0.005; ***p≤0.001). I-K) EMSAs with a FnoCas9:scaRNA/tracrRNA complex (1:2 molar ratio FnoCas9:RNA duplex) and DNA targets (1 nM) with 5, 10, or 20 bp between the TSS and the 11 bp region of complementarity to scaRNA: (I) 5 bp from TSS (9035/9036; DNA sequences in Table S2), (J) 10 bp from TSS (9037/9038), and (K) 20 bp from TSS (9039/9040).

with complementarity to scaRNA (Figure S3A). We observed that constructs with 0, 5, and 10 bases between the TSS and the scaRNA complementarity region effectively repressed *gfp** (Figure 4H). However, the construct with 20 bp between the TSS and the scaRNA complementarity region exhibited significantly reduced repression (Figure 4H). To test if this was the result of altered binding, we conducted *in vitro* DNA binding assays with constructs that have 0, 5, 10, or 20 bases between the TSS and an 11 bp scaRNA target region. All of the constructs had comparable binding affinities to FnoCas9 in complex with scaRNA:tracrRNA, suggesting that the difference in transcriptional inhibition observed when the FnoCas9 binding site is moved further from the TSS is not the result of decreased binding affinity (Figure 4D,I-K). These data highlight that the scaRNA complementarity region must be in close proximity to the TSS for effective FnoCas9-dependent transcriptional interference to occur. Together, these results indicate that binding affinity through complementarity to scaRNA can modulate the level of transcriptional interference nearby a TSS, until the number of bases alters Cas9 function from DNA binding to cleavage.

Cleavage-capable FnoCas9 binds competing RNAs to form two distinct complexes with different functions

We demonstrated that the presence of crRNA does not contribute to the repression of *1104-1101* by scaRNA:tracrRNA-FnoCas9, and that scaRNA with artificially extended complementarity to a target could inhibit transformation with a target-containing plasmid (Figure 1C-F, S3E). To test whether DNA targeting by the CRISPR array was independent of scaRNA, we transformed WT *F. novicida* and deletion mutants of crRNA, scaRNA, tracrRNA, and *cas9* with a plasmid containing a target of an endogenous crRNA (with and without a PAM), and evaluated the ability of each strain to restrict transformation relative to WT. All strains were transformed with the plasmid lacking a PAM at the same frequency. However, *cas9*, *tracrRNA*,

and *crRNA* mutants were unable to restrict transformation with a target plasmid containing a PAM, while the *scaRNA* mutant inhibited transformation similarly to WT (Figure S5A). Together, these results led us to hypothesize that FnoCas9 forms two distinct cleavage-capable RNA complexes in *F. novicida*, one with scaRNA:tracrRNA, and another with crRNA:tracrRNA (Figure S5B). To test this, we measured the interaction between purified FnoCas9 and preformed complexes with either scaRNA:tracrRNA or crRNA:tracrRNA. This analysis revealed that FnoCas9 can indeed interact with both small RNA pairs (Figures S5C-D), and raised the question of how these distinct small RNAs may affect each other.

We performed Northern blot analysis for tracrRNA, crRNA, and scaRNA from a panel of mutant and complemented strains. As expected, the abundance of each small RNA was dependent on the presence of FnoCas9 since each was undetected in a *cas9* mutant but restored in the complemented strain. In addition, analysis of the *tracrRNA* mutant and complemented strains indicated that the presence of crRNA and scaRNA was dependent on tracrRNA. tracrRNA processing was retained upon deletion of the crRNAs, suggesting that scaRNA can guide the processing of tracrRNA, likely through a similar duplex interaction with the tracrRNA anti-repeat as crRNA. Interestingly, deletion of scaRNA did not significantly alter the levels of crRNA, while deletion of crRNA led to an increase in scaRNA levels that was recovered to WT levels in a crRNA complement (Figures 5A-C). To investigate the impact of the increased abundance of scaRNA in a *crRNA* mutant, we measured the level of *1104-1101* in a *crRNA* mutant compared to WT by qRT-PCR (quantitative reverse transcriptase PCR). In a *crRNA* mutant, *1104-1101* were expressed at lower levels than in WT *F. novicida*, consistent with

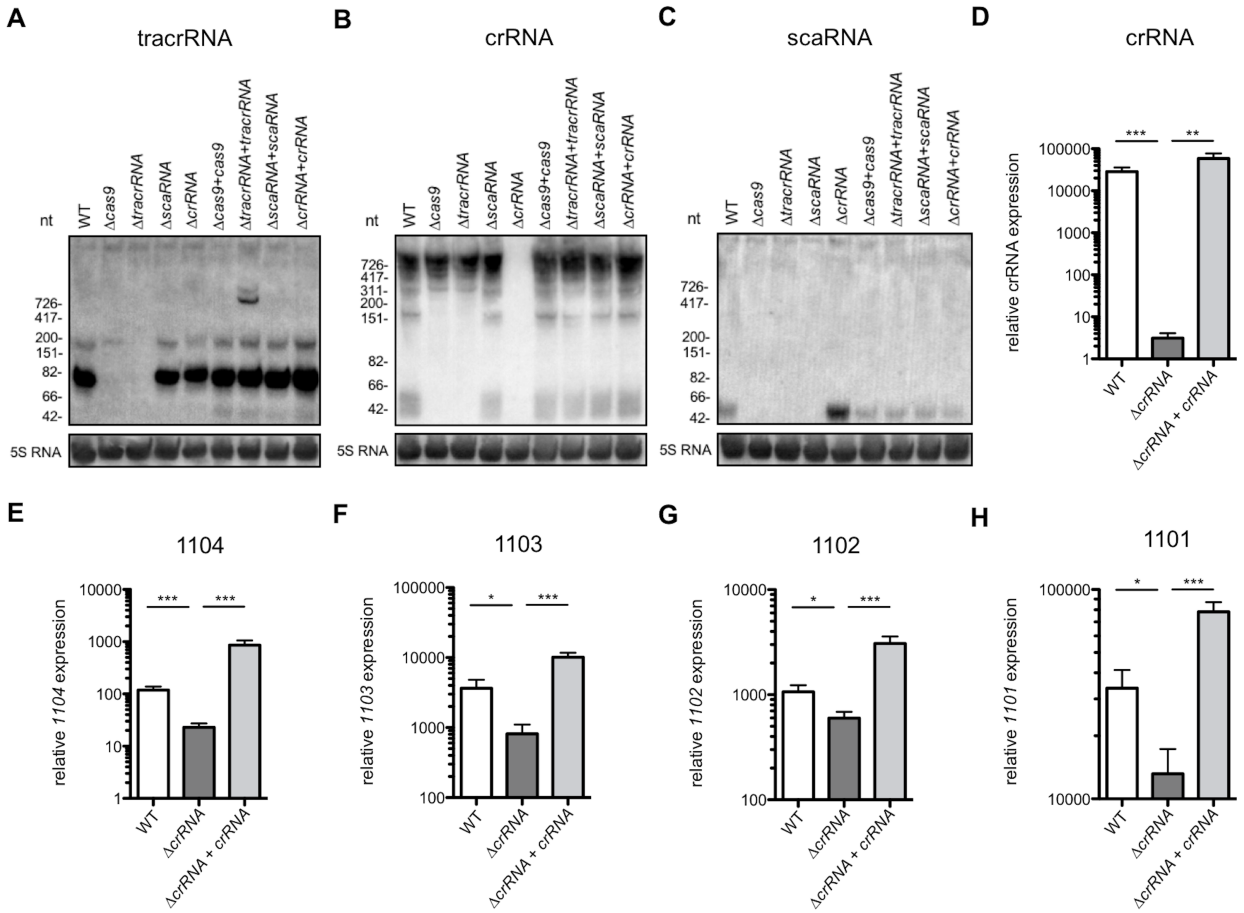


Figure 5. FnoCas9 forms complexes with two different RNA duplexes. (A-C) Northern blots for (A) tracrRNA, (B) crRNA, and (C) scaRNA in wild-type (WT), mutants for each component of the FnoCas9 complexes (Δ*cas9*, Δ*tracrRNA*, Δ*crRNA* and Δ*scaRNA*), and the complementation strains. D-H) qRT-PCR for transcript levels of (D) crRNA, (E) *1104*, (F) *1103*, (G) *1102*, and (H) *1101* in WT, Δ*crRNA*, and Δ*crRNA* + crRNA complemented strains ((D-G) n=7-9, (H) n=4-6, error bars represent s.e., *p≤0.05; **p≤0.005; ***p≤0.001).

enhanced transcriptional repression (Figure 5E-H, 1C-F). Complementation of the *crRNA* mutant restored wild-type levels of crRNA expression and repression of *1104-1101* (Figure 5D-H). Taken together, these results suggest that there is competition for FnoCas9 between the scaRNA and crRNA.

Repression of each gene in the FnoCas9 regulon contributes to virulence

It was unclear whether repression of each gene in the FnoCas9 regulon contributes to *F. novicida* virulence. To test this, we infected mice subcutaneously with *cas9* double mutants also lacking one of the repressed genes. This revealed that *1104* expression plays a major role in the attenuation of the *cas9* mutant, similar to *1103* (Figure S6B). Deletion of *1104* from a *cas9* mutant led to a greater than 4-log enhancement in the levels of bacteria recovered after infection as compared to the *cas9* mutant (Figure S6B). *1102* and *1101* played more minor roles, but still contributed to attenuation of *F. novicida* when expressed, reducing virulence by 1-2 logs (Figure S6C-E). To test if virulence could be completely restored to the *cas9* mutant by deletion of the entire *1104-1101* locus, we infected mice with a Δ*cas9*Δ*1104-1101* strain and evaluated the bacterial burden in the spleen 48 hours post-infection. We found that unlike the Δ*cas9*Δ*1103* mutant, virulence was restored to WT levels by deletion of the four genes in the FnoCas9 regulon in a Δ*cas9* strain (Figure 6). These data demonstrate that *F. novicida* Cas9 represses these four genes whose expression would otherwise lead to attenuation of bacterial virulence.

scaRNA can be reprogrammed to guide FnoCas9 to repress non-native targets

We next sought to determine whether scaRNA can be reprogrammed to repress new targets and whether promoter regions can be targeted for repression similar to the 5’ UTRs. We

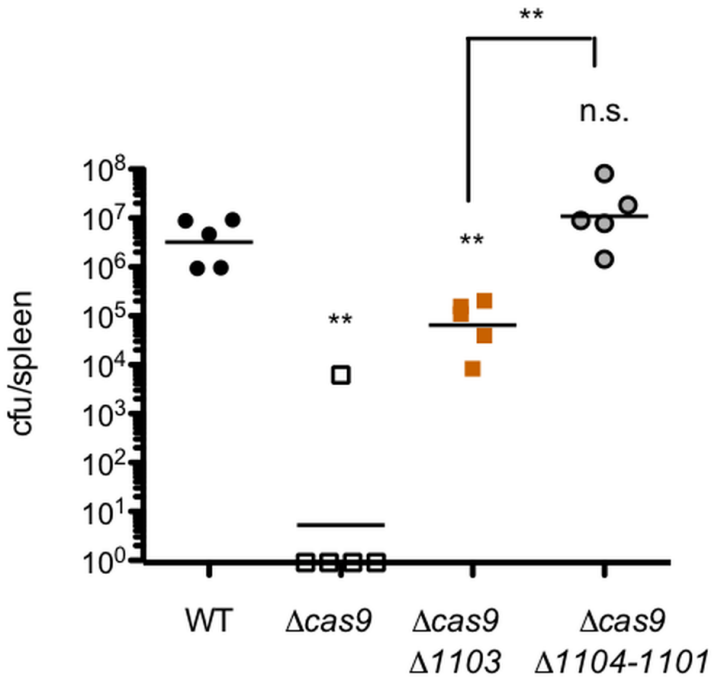


Figure 6. Deletion of 1104-1101 restores virulence of a cas9 mutant. Mice were subcutaneously infected with either WT *F. novicida*, Δ*cas9*, Δ*cas9*Δ*1103*, or Δ*cas9*Δ*1104-1101*. Spleens were homogenized and plated for enumeration of colony forming units (CFU) at 48 hr post-infection (*n*=5, bar represents geometric mean, **p≤0.005).

replaced the *1104-1101* targeting portion of the scaRNA tail with a 16 bp sequence complementary to a portion of the 98 bp intergenic region between *FTN_0544* (0544; *naxD*, new NCBI locus tag: *FTN_RS02820*) and *FTN_0545* (0545; *flmF2*, new NCBI locus tag: *FTN_RS02825*), located upstream of the TSS for each gene (Figure S7). These two genes are required for the modification of outer membrane lipid A that leads to resistance to the antibiotic polymyxin B; they are transcribed in opposite directions, and are not regulated by FnoCas9 (Figure 7A-C, S7A) (Kanistanon et al., 2008; Llewellyn et al., 2012). In the scaRNA reprogrammed strain, we observed a significant reduction in transcript levels of both *0544* and *0545* compared to WT (Figures 7B, C). This suggests that the FnoCas9 CRISPR-Cas9 system can be engineered to repress the expression of new targets, and that repression is independent of the strand targeted. Furthermore, this repression was dependent on FnoCas9, since a *cas9* mutant harboring the reprogrammed scaRNA exhibited WT levels of *0544* and *0545* (Figures 7B, C). In addition, when the natural scaRNA was reprogrammed for *0544-0545*, the strain lost the ability to repress *1104* (Figure 7D). Reprogrammed scaRNA repressed transcription of *0544* at the same efficiency in strains harboring WT Cas9 or dCas9 (Figure S7B). The repression of *0544* and *0545* in the scaRNA reprogrammed strain led to an increased susceptibility to polymyxin B of almost 100-fold, similar to a *0544* deletion strain (Figure 7E). The susceptibility was reversed by deletion of *cas9* from the reprogrammed *scaRNA* strain (Figure 7E). These results indicate that FnoCas9 can be reprogrammed to repress expression from new targets in a scaRNA:tracrRNA-dependent manner, highlighting the scaRNA:tracrRNA-FnoCas9 machinery as a potential new tool to control gene expression and modulate bacterial physiology.

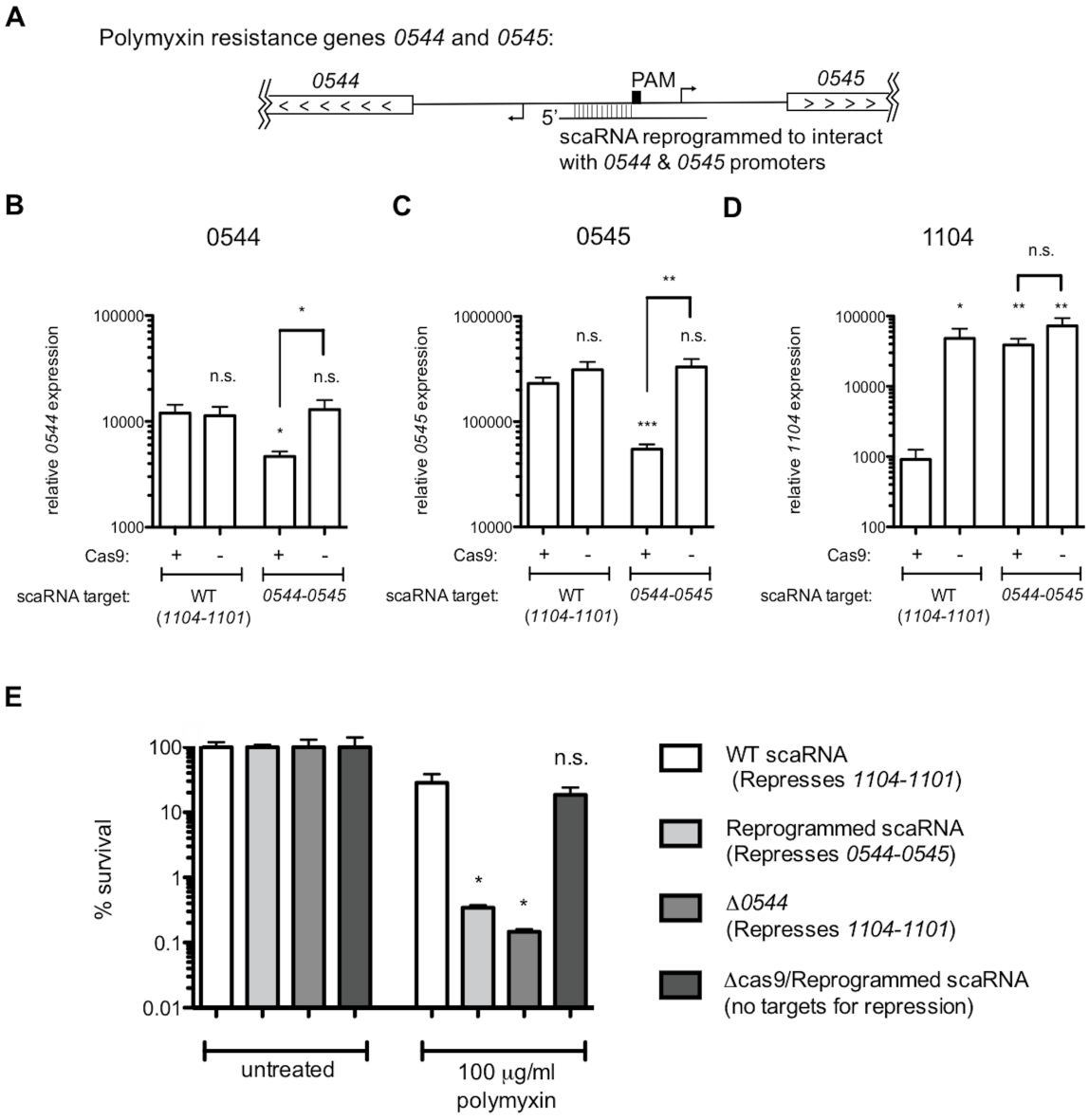

A
Polymyxin resistance genes 0544 and 0545:
0544
PAM
0545
5'
scaRNA reprogrammed to interact
with 0544 & 0545 promoters
B
0544
relative 0544 expression
100000
10000
1000
n.s.
n.s.
*
*
Cas9: + - + -
scaRNA target: WT (1104-1101) 0544-0545
C
0545
relative 0545 expression
1000000
100000
10000
n.s.
n.s.
**
***
Cas9: + - + -
scaRNA target: WT (1104-1101) 0544-0545
D
1104
relative 1104 expression
100000
10000
1000
100
n.s.
*
**
**
Cas9: + - + -
scaRNA target: WT (1104-1101) 0544-0545
E
% survival
100
10
1
0.1
0.01
n.s.
*
*
untreated
100 μg/ml polymyxin
WT scaRNA (Represses 1104-1101)
Reprogrammed scaRNA (Represses 0544-0545)
Δ0544 (Represses 1104-1101)
Δcas9/Reprogrammed scaRNA (no targets for repression)

Figure 7. scaRNA can be reprogrammed to repress new targets. A) Schematic of the scaRNA target site when reprogrammed to interact with the intergenic region between two *F. novicida* genes, *0544* and *0545*, that are transcribed in opposite directions (TSS indicated with arrows). *0544* and *0545* ORFs are 98 bp apart in the genome and were targeted upstream of the two TSSs. B-D) qRT-PCR for transcript levels of (B) *0544*, (C) *0545*, and (D) *1104* in WT, Δ*cas9,* reprogrammed scaRNA (WT with the scaRNA tail reprogrammed to target *0544-0545*), and Δ*cas9*+reprogrammed scaRNA (Δ*cas9* with the scaRNA tail reprogrammed to target *0544-0545*) (*n*=6, error bars represent s.e., *p≤0.05; **p≤0.005; ***p≤0.001). E) Percent survival of WT, reprogrammed scaRNA, Δ0544, and Δ*cas9*+reprogrammed scaRNA strains 6 hours after polymyxin treatment (100 ¦ g/mL) relative to untreated strains (*n*=3, error bars represent s.e., *p≤0.05).

Discussion

Using an analysis of the native FnoCas9 transcriptome to elucidate the specificity of endogenous gene regulation, we located the site of interaction between the FnoCas9 complex and the DNA of the 5' UTR of each transcript in its regulon. FnoCas9 uses scaRNA to interact with the template strand of the 5' UTR by recognition of a PAM and a scaRNA-complementary sequence on the DNA target. By targeting the 5' UTR DNA, FnoCas9 functions as a transcription factor, repressing gene expression. Through this interaction, FnoCas9 regulates the expression of four endogenous genes with remarkably high specificity. We determined that repression is dependent on a PAM in the 5' UTR, and that the sensitivity of natural FnoCas9 regulation could be modulated by the length of the RNA-target interaction and proximity of the scaRNA binding site to the TSS. Further, transcriptional repression by FnoCas9 could be achieved through the targeting of either strand. We demonstrated that the extent of complementarity between scaRNA and the DNA target alters the binding affinity of the dual-RNA-FnoCas9 complex to the DNA. We also observed that the distance of the TSS from the scaRNA target region does not affect the binding affinity of the complex to the DNA. Using this knowledge of scaRNA:tracrRNA-FnoCas9 interaction with DNA, we reprogrammed scaRNA such that FnoCas9 targeted the promoters of desired genes to repress transcription, highlighting the potential use of scaRNA:tracrRNA-FnoCas9 in the control of gene expression.

Previous work from our lab suggested a different model for FnoCas9-mediated regulation of *1103* mRNA that depended on its direct interaction with tracrRNA. We therefore closely re-examined the experiments that suggested RNA targeting in Sampson et al., 2013, and have either been unable to reproduce critical experiments or identified flaws that led to misinterpretation of the previous results, as summarized below. While we cannot rule out the possibility of a low

level of direct targeting of *1103* mRNA by the scaRNA:tracrRNA-FnoCas9 complex or an alternative mechanism of FnoCas9-mediated gene regulation, the data presented here clearly demonstrate that transcriptional interference is the dominant contributor. We sincerely apologize for the previous misleading data and the impact they may have had on others in the field.

We have obtained different results when trying to reproduce the RNA degradation experiment (Figure 2a) from Sampson et al 2013 and no longer observe differences in the rates of *1103* mRNA stability between the wild-type and *cas9* mutant strain (Correction of Sampson et al., 2013, Nature). We do not have an explanation for this discrepancy, but think it may in part be due to the complications derived from measuring *1103* transcript stability in strains with vastly different baseline levels of this mRNA. We have also been unable to replicate the immunoprecipitation experiments in Figure 2g of Sampson et al., 2013 (Correction of Sampson et al., 2013, Nature). The inherent noise in the RNA pulldowns, which did not incorporate a cross-linking step like the DNA pulldowns reported herein, may have contributed to the previous results. Further, the pulldown of *1103* mRNA may have been caused by binding of FnoCas9 to the *1104* UTR DNA where a low level of transcriptional read-through could have occurred, enabling the enrichment of nascent mRNA that are undergoing transcription and are attached to the DNA. However, consistent with Sampson et al., 2013, here we present robust data indicating scaRNA and tracrRNA together with FnoCas9 are required for FnoCas9-dependent repression of RNA levels (Figure 1A, 1C-F, 3E, 5A-H, 7B-D, S2B-D, S5C-D) (Sampson et al., 2013). Finally, in Figure 2h of Sampson et al, 2013, a *tracrRNA* mutant containing substitutions in the putative *1103* interaction site was used to suggest that *1103* was repressed via interaction with tracrRNA. However, due to the location of the tracrRNA mutation in a stem loop that interacts with FnoCas9, we feel that this data can no longer be used to draw a conclusion as to the role of the

mutated bases in the repression of *1103*, and therefore we no longer have evidence to support a role of tracrRNA in direct interaction with *1103* (Correction Sampson et al., 2013, Nature). Therefore, we no longer have conclusive evidence to support a role for RNA degradation in FnoCas9-mediated regulation of *1103*. Although we cannot definitively conclude that the native FnoCas9 system cannot target RNA in some conditions, the data in this manuscript clearly indicate that DNA targeting is the predominant mechanism of endogenous gene regulation.

In some systems, an artificial FnoCas9 complex has been engineered to target RNA. We showed that an FnoCas9 complex with an engineered guide RNA led to a reduction in hepatitis C virus (HCV) levels in human cells (Price et al., 2015). Since HCV is an RNA virus whose genome is only in the form of RNA and never DNA, the reduction in HCV levels by FnoCas9 were due to targeting of RNA, the result of either repression of viral genome replication, translational inhibition, or RNA degradation. In this system, we employed an FnoCas9 catalytic domain point mutant to show that FnoCas9 catalytic activity was not required for repression of HCV levels, ruling out direct RNA degradation via the RuvC and HNH motifs (Price et al., 2015). Another group has also shown that FnoCas9 can repress the levels of tobacco mosaic virus, another RNA virus that also lacks a DNA stage in its lifecycle, during infection of plants (Zhang et al., 2018). However, these mechanisms of RNA targeting are distinct from the regulation of *1104-1101* at the DNA level by the native FnoCas9 system. Whereas the native system uses scaRNA and tracrRNA to target DNA, RNA targeting by the artificial system utilizes a single-guide RNA with tracrRNA modifications. Furthermore, a PAM is not required for HCV repression, while a PAM is involved in FnoCas9 repression of endogenous gene transcription from the DNA.

The mechanism of gene regulation via transcriptional inhibition we have described here is also unique compared to other examples of repression mediated by Cas9 orthologs. *N. meningitidis*

(NmeCas9), *Staphylococcus aureus* (SauCas9) and *C. jejuni* (CjeCas9) have recently been shown to degrade ssRNA in a PAM-independent manner that requires the HNH catalytic domain and perfect or near perfect complementarity with a crRNA spacer (Dugar et al., 2018; Rousseau et al., 2018; Strutt et al., 2018). ssRNA cleavage by Cas9 has been proposed to have roles in endogenous gene regulation (CjeCas9) and foreign nucleic acid defense (SauCas9) (Dugar et al., 2018; Strutt et al., 2018). However, at least *in vitro*, FnoCas9 is not capable of this mechanism of ssRNA targeting (Strutt et al., 2018). FnoCas9-mediated transcriptional repression of *1104-1101* through interaction with the DNA is also distinct from the engineered targeting of RNA by SpyCas9, which is mediated by Cas9 interaction with the RNA target and requires supplementation of a short PAMmer sequence (O'Connell et al., 2014).

It is not clear why FnoCas9 evolved to repress endogenous gene expression. One explanation for the ability of FnoCas9 to repress transcription from the DNA is as an expansion of the phage defense toolkit, to not only block replication of phage DNA, but also repress the transcription of harmful phage genes. It is hypothesized that scaRNA evolved from a degenerated CRISPR array that contains repeats with impaired complementarity to the inverted repeat of tracrRNA (Chylinski et al., 2013). We found that scaRNA is still capable of interacting with tracrRNA, likely through the repeat similar to the interaction between crRNA and tracrRNA. Transcriptional repression of *1104-1101* by FnoCas9 could be the result of genome shuffling events in the area of the CRISPR arrays, leading to the evolution of the self-targeting scaRNA through environment-specific fitness advantages of *1104-1101* repression. Alternatively, a spacer might have been acquired from the bacterial genome and degenerated during this process, avoiding self-cleavage. In a later step, a promoter may have evolved upstream of this spacer, allowing the bacterium to control its expression independently of the array.

It is particularly interesting that in spite of the degeneration of its repeat sequence, scaRNA has retained the ability to direct DNA cleavage. When *F. novicida* is transformed with an artificial target containing 20 bases of identity to the scaRNA, FnoCas9 restricts transformation. However, the 11 consecutive bases of perfect complementarity between scaRNA and the native *1104* and *1101* 5' UTRs is sufficient for robust transcriptional repression, which we hypothesize is due to the inability of FnoCas9 to enter a cleavage-favorable conformation with a partial scaRNA-DNA target interaction. If so, this would be similar to what has been observed with shortened crRNA spacers, which guide Cas9 to bind but not cleave a DNA target (Bikard et al., 2013; Sternberg et al., 2015). Thus, modification of the length of the targeting sequence of the guiding scaRNA:tracrRNA duplex determines whether FnoCas9 represses transcription or cleaves its DNA target. We utilized this knowledge to reprogram scaRNA to target genes involved in polymyxin resistance. This led to efficient repression of the targeted genes and greatly increased sensitivity to polymyxin.

The use of a catalytically active Cas9 for gene repression makes this system unique compared to engineered CRISPRi technologies that use catalytically inactive mutants of the protein (dCas9) in complex with an RNA guide containing 20 bases of complementarity to its target (Bikard et al., 2013; Larson et al., 2013; Qi et al., 2013). This highlights the potential for scaRNA:tracrRNA-FnoCas9 complexes to be used to control gene expression, and especially for applications that seek to multiplex DNA cleavage and transcriptional control. Furthermore, within *F. novicida*, FnoCas9 is able to prevent transformation and regulate gene expression simultaneously, suggesting that, at least in our experimental conditions, sufficient FnoCas9 molecules are bound to both duplexes (tracrRNA:crRNA and tracrRNA:scaRNA) to fulfill each

function with minimal effect on the other. This would be analogous to the multiple spacers protecting against different phages in parallel without out-competing each other.

The limited scaRNA:target complementarity required for FnoCas9 transcriptional repression, as compared to 20 bp crRNA or guide RNA complementarity to cleavage targets, could increase the risk of off-target effects. Typically, analyses to identify off-target cleavage sites are performed with full length (20 bp) target sequences, however, we propose that for any use of Cas9, such analyses should include an examination of potential off-target transcriptional effects as well. Similarly, potential endogenous regulatory functions of native CRISPR-Cas9 systems could be identified by decreasing the stringency of self-targeting spacer identification.

FnoCas9 transcriptional repression is critical for *F. novicida* virulence and we found that repression of each gene in the FnoCas9 regulon has a distinct contribution to virulence in a mouse model (Figure S6B). We find that the attenuation of a *cas9* mutant of *F. novicida* can be reversed by deletion of the *1104-1101* locus (Figure 6). *1104*, *1103*, and *1101* all exhibit conserved features of Gram-negative bacterial lipoproteins, while *1102* has some but not all of these features (Figure S6A).

Together, these results highlight a novel role for CRISPR-Cas9 systems in endogenous gene regulation and provide a mechanistic explanation of the role Cas9 plays in the virulence of *F. novicida*. Interestingly, *F. novicida* utilizes two distinct RNA duplexes for foreign DNA restriction and transcriptional repression, although both are capable of DNA restriction. The prevalence of these systems and the minimal base pair requirements needed for a shift between DNA cleavage and transcriptional interference suggest that a role of Cas9 as a transcriptional regulator may be a broader phenomenon in bacterial physiology than previously expected.

Supplemental Figures.

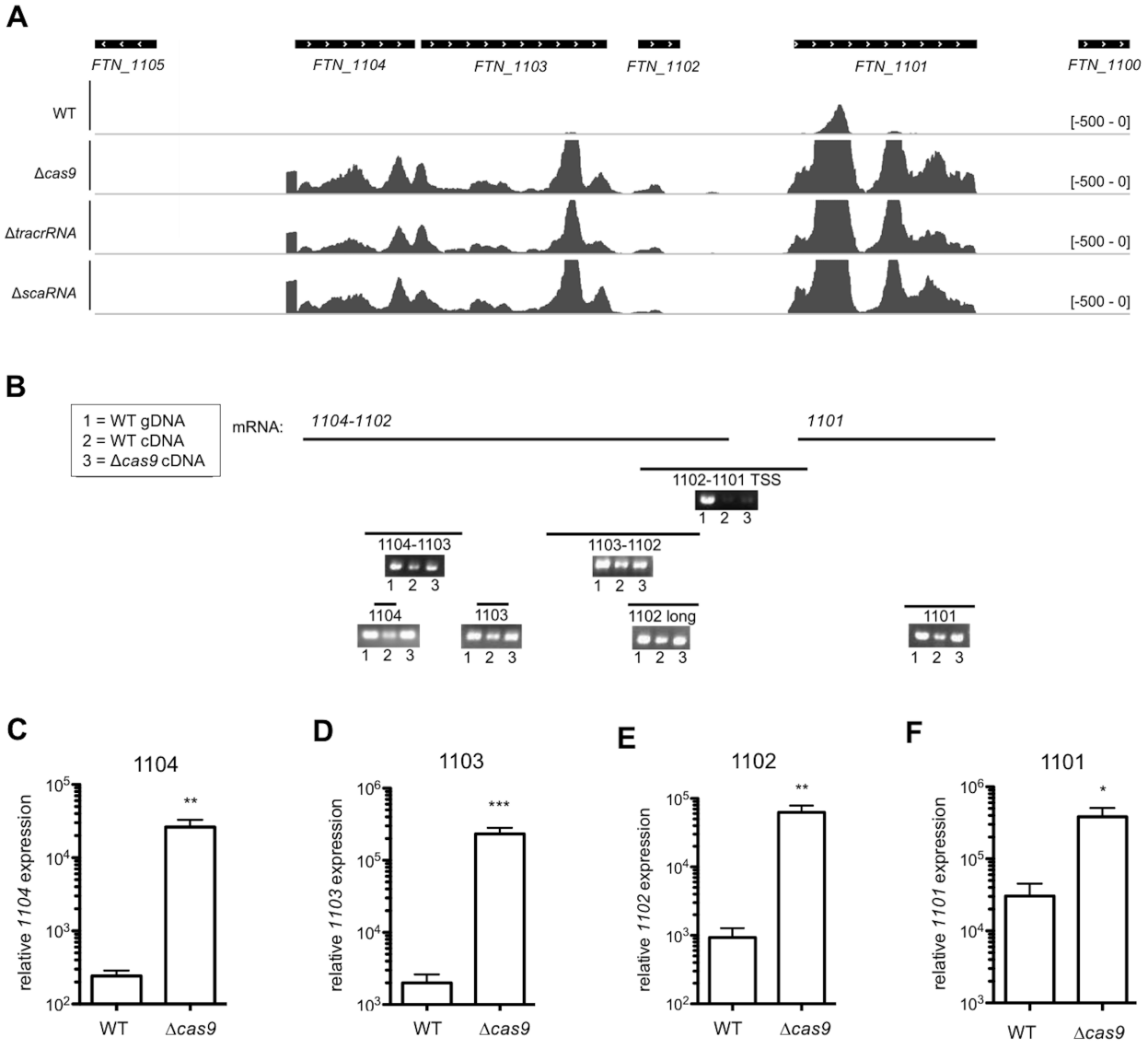


Figure S1. Validation of the 1104-1103-1102 and 1101 Transcripts, Related to Figure 1. A) RNA sequencing coverage of the FnoCas9-regulated region in WT, Δ*cas9*, Δ*tracrRNA* and Δ*scaRNA*. B) PCR amplification over gene junctions. The location of the amplified regions is indicated in relation to the transcribed RNAs determined by RNAseq. Each primer set was used to amplify from "1" WT U112 gDNA as a control, "2" WT U112 cDNA, and "3" Δ*cas9* cDNA. C-E) qRT-PCR for transcript levels of each gene in the FnoCas9 regulon in WT and Δ*cas9* strains; (C) *1104*, (D) *1103*, (E) *1102*, (F) *1101* (n=9, error bars represent s.e.; *p≤0.05; **p≤0.005; ***p≤0.001).

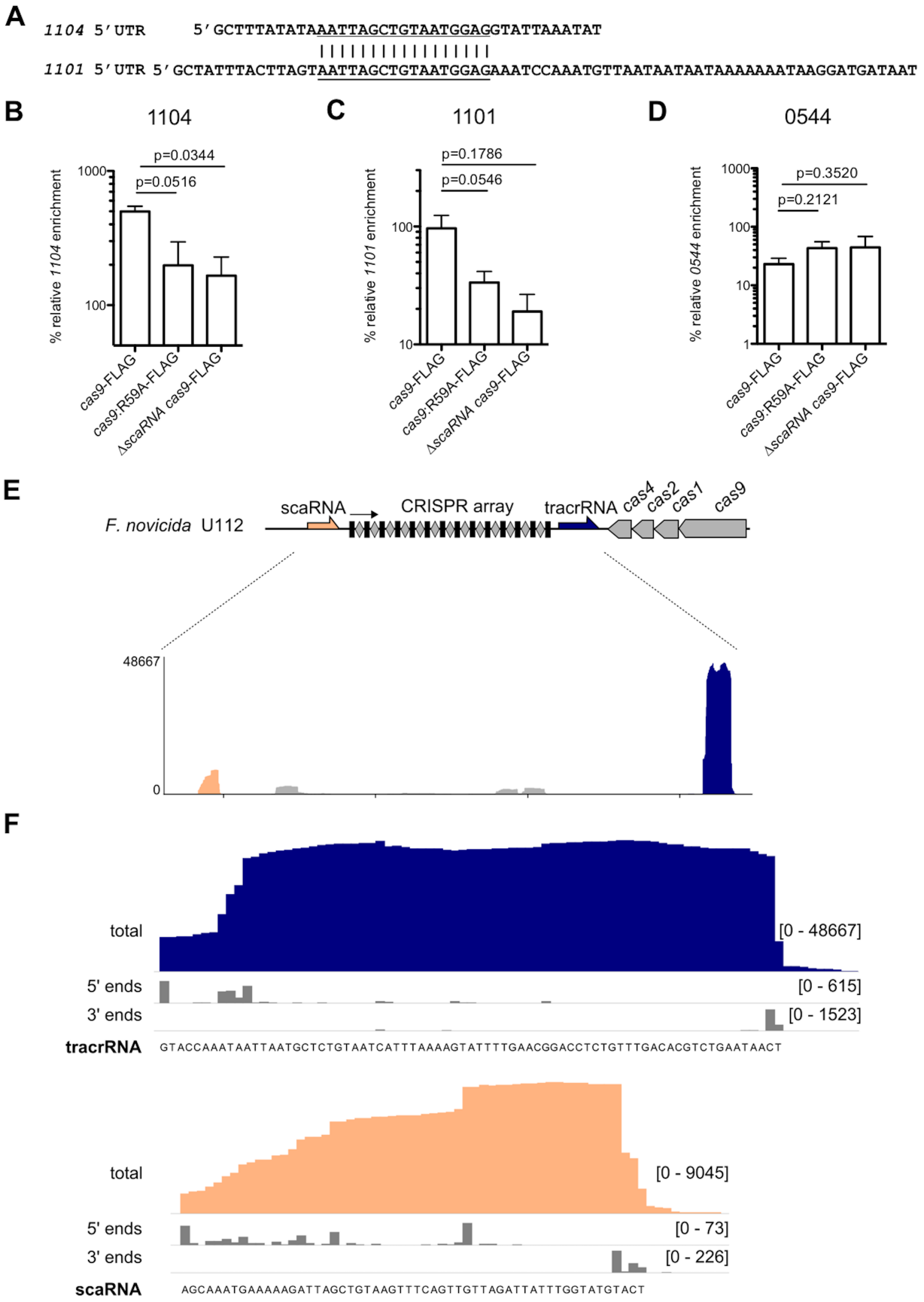

A
1104 5′UTR 5′GCTTTATATAAATTAGCTGTAATGGAGGTATTAAATAT
1101 5′UTR 5′GCTATTTACTTAGTAATTAGCTGTAATGGAGAAATCCAAATGTTAATAATAATAAAAAAATAAGGATGATAAT
B
1104
% relative 1104 enrichment
p=0.0344
p=0.0516
cas9-FLAG
cas9:R59A-FLAG
ΔscaRNA cas9-FLAG
C
1101
% relative 1101 enrichment
p=0.1786
p=0.0546
D
0544
% relative 0544 enrichment
p=0.3520
p=0.2121
E
F. novicida U112
scaRNA
CRISPR array
tracrRNA
cas4
cas2
cas1
cas9
48667
0
F
total
[0 - 48667]
5' ends
[0 - 615]
3' ends
[0 - 1523]
tracrRNA
GTACCAAATAATTAATGCTCTGTAATCATTTAAAAGTATTTTGAACGGACCTCTGTTTGACACGTCTGAATAACT
total
[0 - 9045]
5' ends
[0 - 73]
3' ends
[0 - 226]
scaRNA
AGCAAATGAAAAAGATTAGCTGTAAGTTTCAGTTGTTAGATTATTTGGTATGTACT

Figure S2. scaRNA is Required for FnoCas9 Interaction with 1104 and 1101 DNA, Related to Figures 2 and 3. A) Alignment of the coding strands of the *1104* and *1101* 5' UTRs. B-D) *cas9*-FLAG, *cas9*:R59A-FLAG, and Δ*scaRNA cas9*-FLAG strains were crosslinked and Cas9-FLAG was immunoprecipitated from bacterial lysates. Quantitative PCR was performed on immunoprecipitated DNA and the enrichment of (B) *1104* DNA, (C) *1101* DNA, and (D) *0544* DNA, was determined relative to a housekeeping gene (n=3-6, error bars represent s.e.). E-F) tracrRNA and scaRNA detected by small RNA sequencing. (E) RNAseq of tracrRNA (blue) and scaRNA (orange) in the context of the CRISPR locus. (F) The DNA sequence, total coverage, and read coverage of the 5' and 3' ends are indicated.

A

Fusion construct — Promoter — UTR — PAM — GFP

Distance from TSS — Synthetic — +1 — 0, 5, 10, or 20 bp + 11 bp identity to scaRNA — +PAM

0 bp taaataccaattgacatatataaatgattctgatataaattagataagggattagctgtaatgggccaacacttgtcactactctgacg…
5 bp taaataccaattgacatatataaatgattctgatataaattagataagggttagtattagctgtaatgggccaacacttgtcactactctgacg…
10 bp taaataccaattgacatatataaatgattctgatataaattagataagggtttacttagtattagctgtaatgggccaacacttgtcactactctgacg…
20 bp taaataccaattgacatatataaatgattctgatataaattagataagggagaaatccaatttacttagtattagctgtaatgggccaacacttgtcactactctgacg…

Identity to scaRNA — Synthetic — 0 bp + 0, 8, 11, 15, or 20 bp identity to scaRNA — +PAM

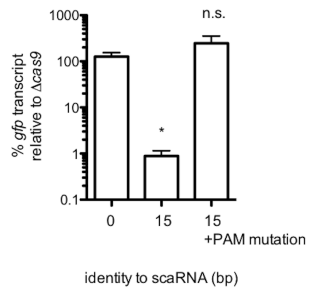
0 bp taaataccaattgacatatataaatgattctgatataaattagataagggaggccaacacttgtcactactctgacg…
8 bp taaataccaattgacatatataaatgattctgatataaattagataagggagctgtaatgggccaacacttgtcactactctgacg…
11 bp taaataccaattgacatatataaatgattctgatataaattagataagggattagctgtaatgggccaacacttgtcactactctgacg…
15 bp taaataccaattgacatatataaatgattctgatataaattagataagggaaagattagctgtaatgggccaacacttgtcactactctgacg…
20 bp taaataccaattgacatatataaatgattctgatataaattagataagggatgaaaaagattagctgtaatgggccaacacttgtcactactctgacg…

PAM mutant:
15 bp taaataccaattgacatatataaatgattctgatataaattagataagggaaagattagctgtaataaggccaacacttgtcactactctgacg…

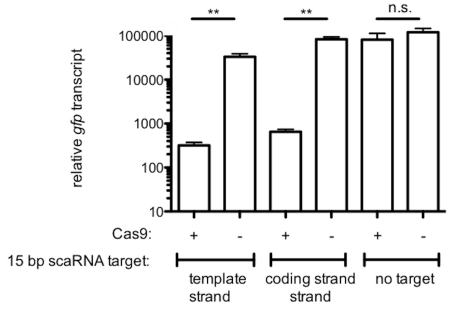
Reverse complement:
15 bp taaataccaattgacatatataaatgattctgatataaattagataagggccattacagctaatctttggccaacacttgtcactactctgacg…

Identity to crRNA
20 bp taaataccaattgacatatataaatgattctgatataaattagataaggggattacttttatagtgtgattgggccaacacttgtcactactctgacg…

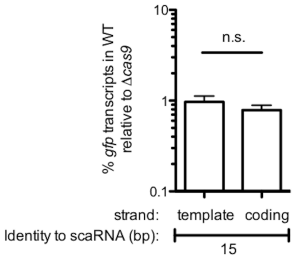
B



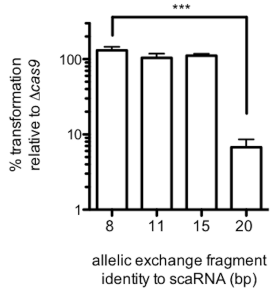
C



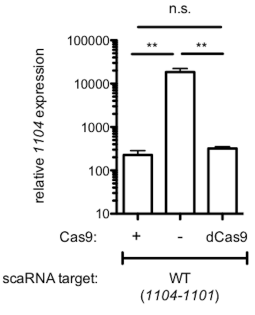
D



E



F



G

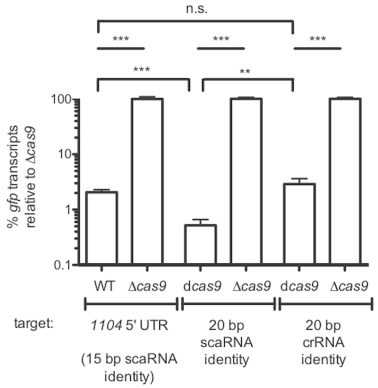


Figure S3. Parameters Governing FnoCas9-mediated Transcriptional Repression and Transformation Restriction, Related to Figure 4. A) All fusion constructs driving expression of *gfp** were made in the chromosome in place of the *1101-1104* locus, using a synthetic constitutive promoter. Between the promoter and the PAM-*gfp**, the distance of the scaRNA targeting site from the TSS or amount of scaRNA complementarity to the target were modulated to evaluate their effect on transcriptional repression. Constructs with a PAM mutation, the target and PAM on the opposite strand, and a crRNA target were also tested. Sequences as indicated. B) Relative *gfp** transcript levels were measured by qRT-PCR from constructs containing 15 bp complementarity to the scaRNA tail followed by a WT TGG PAM or a TAA PAM mutation, between a synthetic promoter and *gfp** sequence. Results are presented as % *gfp** transcript with WT FnoCas9 relative to Δ*cas9* (*n*=3, error bars represent s.e.; *p≤0.05). C) Relative *gfp** transcript levels were measured by qRT-PCR in WT (Cas9+) or Δ*cas9* (Cas9-) strains, from fusions constructs with 15 bp complementarity to the scaRNA and adjacent PAM located on either the template or coding strand. A strain with *gfp** placed downstream of the synthetic promoter and 0 bp of complementarity to scaRNA was used as a "no target" control (*n*=3, error bars represent s.e.; **p≤0.005). D) Results from "C" are presented as % *gfp** transcript in WT relative to a Δ*cas9* strain. E) Linear DNA fragments of the fusion constructs with 8, 11, 15, or 20 bases of complementarity to scaRNA between the promoter and *gfp** were transformed into WT *F. novicida* and a Δ*cas9* mutant, and recombinants were selected. Results are presented as the percent transformation efficiency into WT relative to a Δ*cas9* mutant for each DNA fragment. (*n*=4-8, error bars represent s.e.; ***p≤0.001). F) Relative *1104* transcript levels were measured by qRT-PCR in WT, Δ*cas9*, and d*cas9* strains (*n*=3-4, error bars represent s.e.; **p≤0.005). G) Relative *gfp** transcript levels were measured by qRT-PCR in WT (Cas9+) and Δ*cas9* (Cas9-) strains from a fusion construct with the native *1104* 5' UTR between the synthetic promoter and *gfp**, and in d*cas9* and Δ*cas9* strains from fusion constructs containing either 20 bp identity to scaRNA or 20 bp identity to crRNA. Results are presented as percent transformation into WT relative to Δ*cas9* (*n*=6, error bars represent s.e.; **p≤0.005; ***p≤0.001).

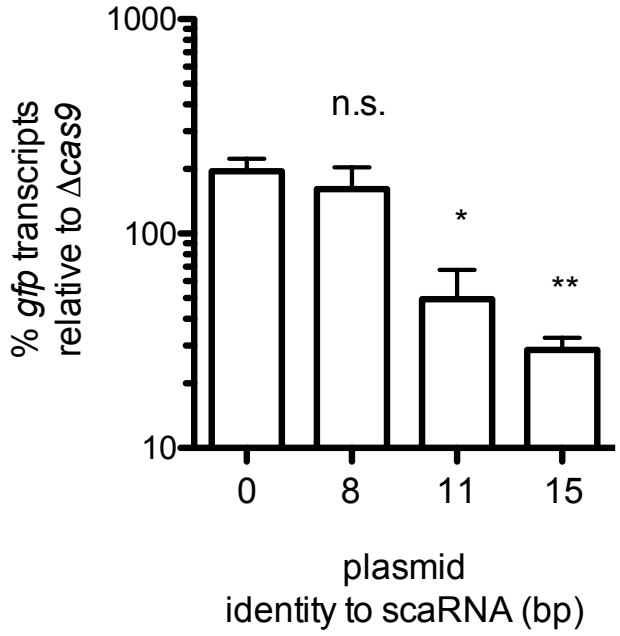


Figure S4. Identity to scaRNA Determines Repression Level From a Plasmid-based Reporter Construct, Related to Figure 4. pBAV-derived plasmid constructs with 0, 8, 11, or 15 bases of complementarity to scaRNA, inserted between a synthetic promoter and *gfp**, were transformed into WT *F. novicida* and a Δ*cas9* mutant. *gfp** transcript levels were measured by qRT-PCR (*n*=3, error bars represent s.e.; *p≤0.05; **p≤0.005).

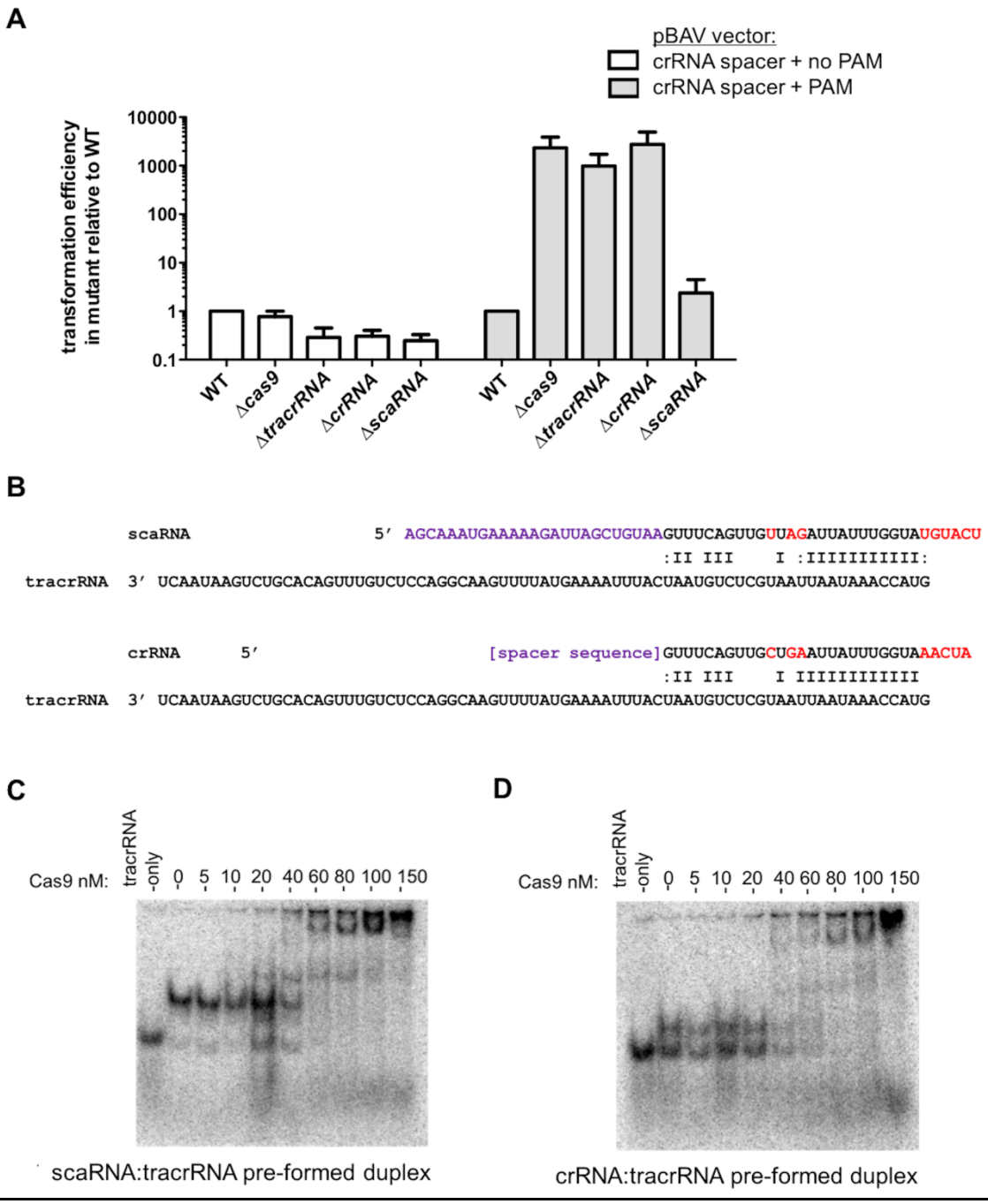

A
pBAV vector:
crRNA spacer + no PAM
crRNA spacer + PAM
transformation efficiency in mutant relative to WT
10000
1000
100
10
1
0.1
WT
Δcas9
ΔtracrRNA
ΔcrRNA
ΔscaRNA
WT
Δcas9
ΔtracrRNA
ΔcrRNA
ΔscaRNA
B
scaRNA 5′ AGCAAAUGAAAAAGAUUAGCUGUAAGUUUCAGUUGUUAGAUUAUUUGGUAUGUACU
:II III I :IIIIIIIIIII:
tracrRNA 3′ UCAAUAAGUCUGCACAGUUUGUCUCCAGGCAAGUUUUAUGAAAAUUUACUAAUGUCUCGUAAUUAAUAAACCAUG
crRNA 5′ [spacer sequence]GUUUCAGUUGCUGAAUUAUUUGGUAAACUA
:II III I IIIIIIIIIIII
tracrRNA 3′ UCAAUAAGUCUGCACAGUUUGUCUCCAGGCAAGUUUUAUGAAAAUUUACUAAUGUCUCGUAAUUAAUAAACCAUG
C
Cas9 nM:
tracrRNA -only
0 5 10 20 40 60 80 100 150
scaRNA:tracrRNA pre-formed duplex
D
Cas9 nM:
tracrRNA -only
0 5 10 20 40 60 80 100 150
crRNA:tracrRNA pre-formed duplex

Figure S5. crRNA Restriction of Foreign DNA is scaRNA-independent, Related to Figure 5. A) pBav-derived plasmids containing a spacer sequence from the CRISPR-Cas9 locus, with and without an adjacent PAM sequence, were transformed into wild-type (WT) and deletion mutants for each component of the *F. novicida* CRISPR-Cas9 system (**Δ***cas9*, **Δ***tracrRNA*, **Δ***scaRNA*, **Δ***crRNA*). Transformation efficiency of the plasmids into each strain was determined relative to WT (*n*=3, error bars represent s.e.). B) Schematic of predicted scaRNA:tracrRNA and crRNA:tracrRNA duplexes. Differences in the scaRNA and crRNA repeat interactions with tracrRNA are highlighted in red and the scaRNA tail and crRNA spacer sequence location are shown in purple. C-D) EMSA showing that FnoCas9 binds scaRNA:tracrRNA and crRNA:tracrRNA pre-formed duplexes. Increasing concentrations of FnoCas9 were incubated with a stable concentration of the RNA duplex: (C) scaRNA:tracrRNA or (D) crRNA:tracrRNA.

A

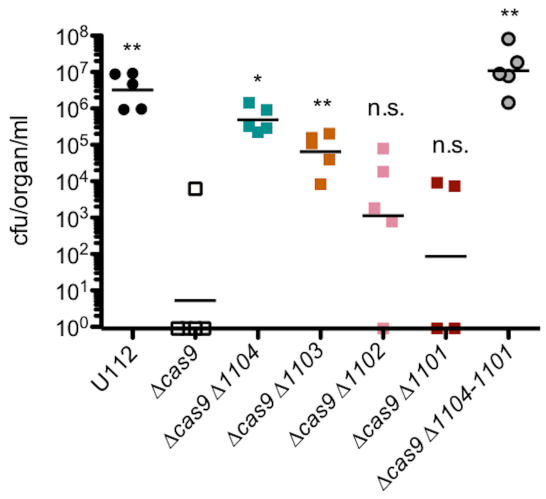
FTN_1104
Lipopredict SpI score=19.6362
MRCNGLIMKDKLLIGVAIISSAALLEGCGKTET

FTN_1103
Lipopredict SpI score=18.4244
MKYGNLMMTKKKLLIGMVTISGIVILGSCGKTET

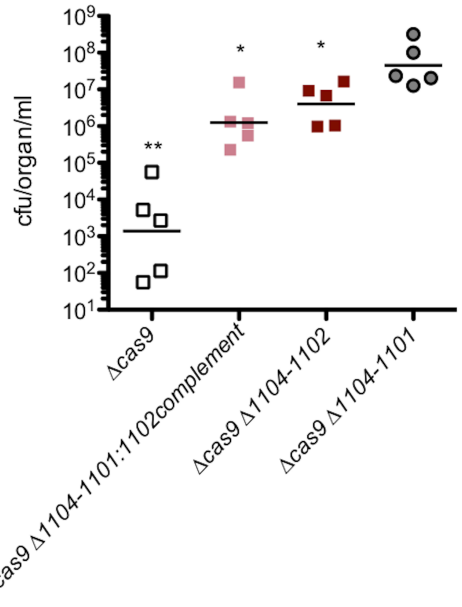
FTN_1102
Lipopredict SpI score=26.2134
MITLFYYAYINSSLRTVTCQYSNT

FTN_1101
Lipopredict SpI score=18.1717
MKITKKKVVIGAAIISGTVLLASCGKTTT

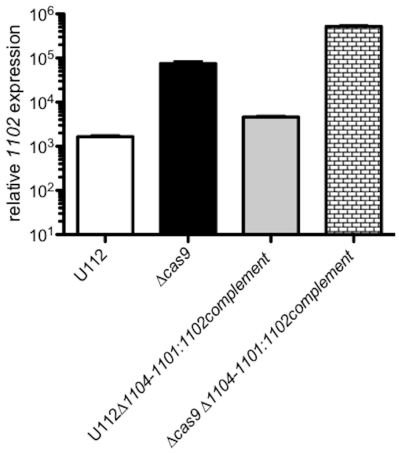
B



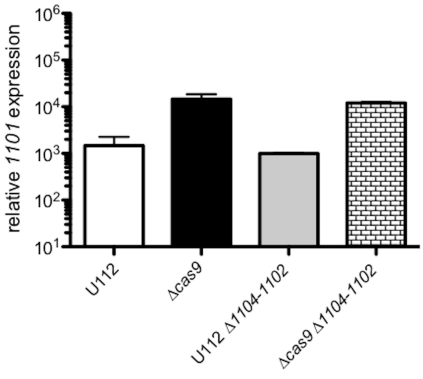
C



D



E



Δ*cas9* Δ*1104-1101:1102 complement* expresses *1102* only
Δ*cas9* Δ*1104-1102* expresses *1101* only

Figure S6. Repression of Each Gene in the FnoCas9 Regulon Contributes to Virulence of F. novicida, Related to Figure 6. A) Lipoprotein prediction scores for the FnoCas9-regulated proteins were generated using the Gram-negative lipopredict software as well as direct analysis of the sequence. The N-terminal amino acid sequences of 1104, 1103, 1102, and 1101 were annotated for canonical lipoprotein features: charged N-terminal region (red), hydrophobic region (blue), and the lipobox (bolded). The cysteine at which signal sequence cleavage occurs is underlined. B) CFU burden in mouse spleens 48 hr post-infection with U112 (WT), **Δ***cas9***,** **Δ***cas9***Δ***1104*, **Δ***cas9***Δ***1103*, **Δ***cas9***Δ***1102*, **Δ***cas9***Δ***1101, or* **Δ***cas9***Δ***1101-1104*. (*n*=4-5, bar represents geometric mean; *p≤0.05; **p≤0.005). C) CFU burden in mouse spleens 48 hours post-infection with **Δ***cas9,* **Δ***cas9***Δ***1104-1102* (expresses 1101), **Δ***cas9***Δ***1104-1101:1102* complemented strain (expresses *1102*), or **Δ***cas9***Δ***1104-1101* (does not express FnoCas9-regulated bacterial lipoproteins) were measured. (*n*=5, line represents geometric mean; *p≤0.05; **p≤0.005). D-E) qRT-PCR of relative (D) *1102* and (E) *1101* expression in U112 (WT) and **Δ***cas9* strains compared to **Δ***1104-1102* and **Δ***1104-1101:1102* complemented strains +/- *cas9* (*n*=3, error bars represent s.e.).

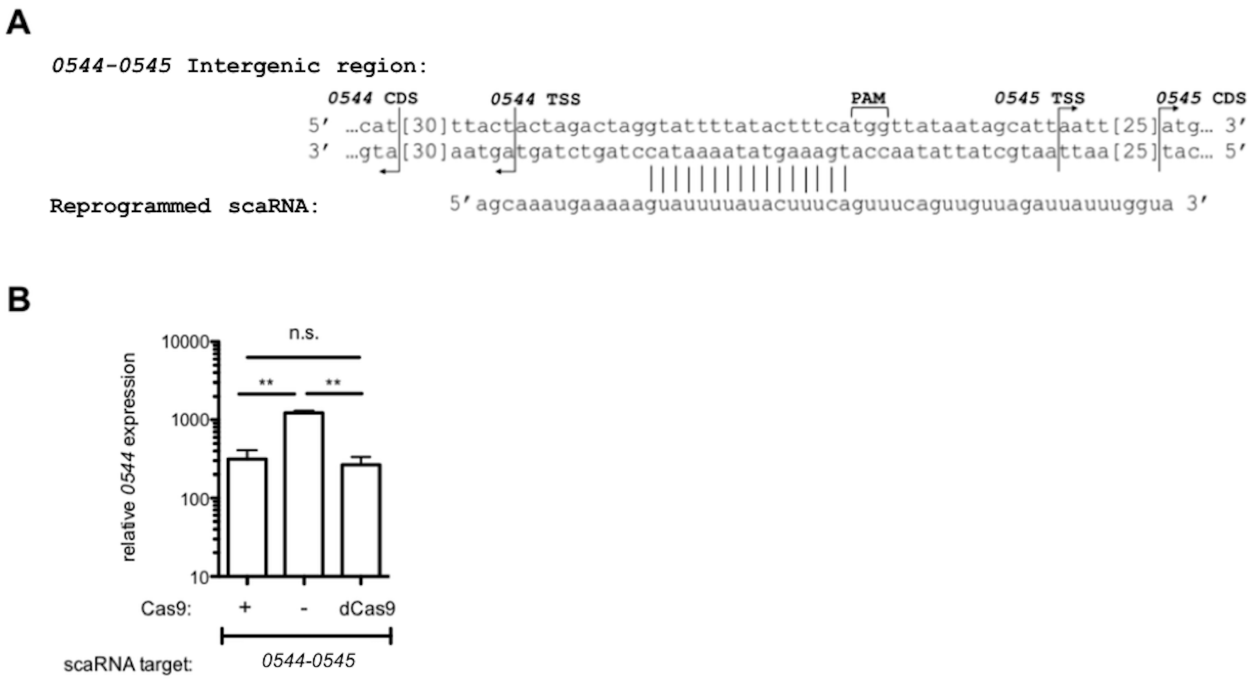


Figure S7. Reprogrammed scaRNA Represses Transcription by Binding Promoter Regions, Related to Figure 7. A) Alignment of the intergenic region between *0544* and *0545* targeted by reprogrammed scaRNA. The locations of transcriptional start sites (TSS) and coding sequences (CDS) are indicated with directional arrows, as is the location of the PAM. B) qRT-PCR of relative *0544* expression in WT *cas9*, Δ*cas9*, and d*cas9* strains with reprogrammed scaRNA (*n*=3, error bars represent s.e.; **p≤0.005).

STAR Methods

KEY RESOURCES TABLE

See published version of this chapter in Ratner et al., Molecular Cell, 2019

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, David Weiss (david.weiss@emory.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

In vivo animal work

Specific-pathogen free mice were kept in filter-top cages at Yerkes National Primate Center, and provided food and water *ad libitum* (Sampson et al., 2013). Emory University Institutional Animal Care and Use Committee (protocol #YER-2000573-061314BN) approved all procedures (Sampson et al., 2013). Female C57BL/6J mice (Jackson) between 8 and 12 weeks were used for all experiments.

Bacterial strains and growth conditions

For information about the strains used and growth conditions, please refer to the strain list (Table S1) and the "METHOD DETAILS" for each experiment. See published version of this chapter in Ratner et al., Molecular Cell, 2019 for Table S1 and S2.

In vitro studies

The DH5α derivative of *E. coli* K12 (NEB 5-alpha) was used for plasmid isolation and cells were grown at 37°C in LB medium supplemented with 30 μg/mL kanamycin. FnoCas9 was purified from *E. coli* NiCo21(DE3) (NEB) expressing Cas9 from pEC657. Cultures for Cas9 purification

were grown at 37°C to OD_{600} of 0.7–0.8, Cas9 expression was induced with IPTG (5mM) and cultures were grown overnight at 13°C, as previously described (Fonfara et al, 2014).

METHOD DETAILS

Construction of cDNA libraries for total RNA analysis

Biological triplicates of *F. novicida* U112 WT, Δ*cas9*, ΔtracrRNA and ΔscaRNA were grown overnight on LB plates (37°C), cultured into LB medium (37°C shaking), and grown until $OD_{620\ nm}$= 0.1. Twenty-five mL of bacterial culture were mixed with 25 mL of 1:1 acetone/ethanol and total RNAs were extracted using TRIzol (Ambion) and treated using turbo DNAse (Ambion). RNA integrity was checked using a bioanalyzer (RIN > 8). cDNA libraries were prepared at the HZI genome analytics platform in Braunschweig, Germany as described with some modifications (Dötsch et al., 2012). Briefly, rRNAs were removed using the MICROB*Express* kit (Ambion) and samples were treated with TAP (tobacco acid phosphatase). The RNAs were fragmented using sonication (Covaris) to fragments of 200 nucleotides. T4 Polynucleotide kinase (Fermentas) was used to phosphorylate 5' ends and remove the 3' phosphate. Successively, 3' and 5' adapters were added using T4 RNA ligase. Reverse transcription was performed using SuperScript II (Invitrogen) followed by 15 cycles of PCR with Phusion (New England Biolabs) and agarose gel purification. The sequencing was performed using 50 nucleotide single end reads (HiSeq2500). cDNA libraries for small RNA analysis were generated as in Chylinski et al., 2013 (Chylinski et al., 2013).

Analysis of total RNA sequencing

The raw data files were demultiplexed using the specific barcode sequences. The reads were trimmed from adapter sequences and the read quality was assessed using fastQC. The reads were

mapped using STAR to the *F. novicida* U112 reference genome (NC_008601.1) (120). We retrieved from 201525 to 518855 of uniquely mapped reads. The number of reads for each gene were counted using HTseq and the differential expression analysis was done for each deletion mutant compared to the WT using DESeq2 (94, 121).

Northern blot analysis

After RNA extraction (see above), RNAs were resolved on 1% agarose containing 0.74% formaldehyde and transferred by capillarity to a nylon membrane (Hybond™ N+, GE healthcare). Forty pmol of oligonucleotide probes were 5′ radiolabeled using [γ-^{32}P] ATP (Hartmann Analytics) and T4 polynucleotide kinase (Fermentas) and purified over Microspin™ G-25 columns (GE Healthcare) following the manufacturers' instructions. The probes were hybridized at 55°C using Rapid-hyb buffer (GE healthcare). The radioactive signal was visualized using a Thyphoon FLA-9500 phosphorimager (GE healthcare) and the transcript sizes were determined using a 3'-end radiolabeled RiboRuler high range RNA ladder (Thermo Scientific^{TM}). 16S rRNA was used as a loading control.

Electrophoretic mobility shift assay

tracrRNA, crRNA and scaRNA were *in vitro* transcribed from annealed oligonucleotides or PCR fragments (Table S2, see published version of this chapter in Ratner et al., Molecular Cell, 2019) using the AmpliScribe™ T7-Flash™ Transcription Kit (Epicenter). scaRNA:tracrRNA or crRNA:tracrRNA duplexes were hybridized by heating to 95°C and cooling to room temperature in hybridization buffer (200 mM NaCl, 20 mM HEPES pH 7.5). DNA substrates were generated by annealing complementary DNA oligonucleotides (Table S2) in a similar manner. Annealed

oligos were 5′ radiolabeled using [γ-^{32}P] ATP (Hartmann Analytics) and T4 polynucleotide kinase (Fermentas) according to the manufacturers' instructions, and purified using Microspin™ G-50 Columns (GE Healthcare). Cas9 was pre-incubated at 37°C with two-fold molar excess of prehybridized scaRNA:tracrRNA duplex for 15 min in DNA-binding buffer (20 mM Tris HCl pH 7.4, 100 mM KCl, 5 mM $CaCl_2$, 1 mM DTT, 5% (w/v) glycerol, 40 ng/µl poly(dI-dC)), then 1 nM labelled DNA substrate was added. For duplex RNA EMSAs, tracrRNA was dephosphorylated using FastAP Alkaline Phosphatase (Thermo Scientific™) and then 5′ radiolabeled as described above, before hybridization with scaRNA or crRNA. The reaction was performed as described above using 100 ng/ul tRNA as competitor. Binding reactions were incubated for 1 h. The samples were loaded on a native 5% polyacrylamide gel, which was run in 0.5X TBE supplemented with 5 mM $CaCl_2$. The gels were exposed on autoradiography films and visualized by phosphorimaging.

Francisella novicida strain construction and growth conditions

All strains are listed in Table S1. *Francisella novicida* U112 derived strains were grown in Tryptic Soy Broth (VWR International) supplemented with 0.2% cysteine (BD Biosciences) at 37 °C with shaking. Strains were plated on Tryptic Soy Agar (TSA, VWR International) plates supplemented with 0.1% cysteine. Deletion and fusion construct mutants were constructed by allelic exchange (primers listed in Table S3). Promoter fusions were inserted in place of the *1104-1101* locus in the chromosome in **Δ***1104-1101* background strains. Chromosomal promoter and 5' UTR fusions were made using a fragment of *gfp* derived from the pBav-kGFP vector, *gfp** (bases 170-499). CRISPR-Cas9 system mutants **Δ***cas9,* **Δ***scaRNA,* **Δ***crRNA,* **Δ***tracrRNA*, their respective complemented strains, and **Δ***cas9***Δ***1103, cas9*-FLAG, *cas9*:R59A-FLAG, and **Δ***scaRNA cas9*-FLAG were described previously (Table S1) (Jones et al., 2012; Sampson et al., 2013; Weiss et al., 2007).

Phusion high-fidelity DNA polymerase (New England Biolabs) was used to amplify homologous sequences (500-1000 bp) flanking the region of interest from genomic DNA isolated with DNeasy Blood and Tissue Kit (Qiagen). Overlapping PCR was used to construct the allelic exchange substrate by inserting a kanamycin selectable marker containing Flp recombinase target sites (FRT) between the flanking sequences (Llewellyn et al., 2011). Allelic exchange substrates were transformed into chemically competent *F. novicida (Llewellyn et al., 2011)*. Mutants were selected on media supplemented with kanamycin sulfate (30 μg/ml, Fisher Scientific). Mutants were confirmed by PCR amplification from outside of the recombined region followed by sequencing (Genewiz) using "seq" primers (Table S3). The selection cassette was removed from the mutants using a temperature sensitive suicide vector at 30 °C, pFFlp, carrying the Flp recombinase in *trans* (Gallagher et al., 2008). pFFlp was selected for on TSA plates with 1% cysteine and 15 μg/ml tetracycline (Alfa Aesar). Following unmarking the strains were moved to 37 °C to remove the plasmid, as described previously (Gallagher et al., 2008).

Quantitative real-time PCR

RNA was isolated from bacterial cultures at OD600 nm of 0.8-1.0 using TRI-reagent and a Direct-zol RNA MiniPrep Kit (Zymo Research). DNA was removed with Turbo DNaseI (Ambion Biosciences). qRT-PCR (quantitative reverse transcriptase PCR) was performed with biological triplicates using the primers indicated in Table S3 and Power Sybr Green RNA-to-CT one-step kit (Applied Biosystems). *CT* values for each gene were normalized to the *Francisella novicida* housekeeping gene DNA helicase II (*uvrD*, *FTN_1594*) to determine $2^{-\Delta\Delta}$ct for each condition (Sampson et al., 2013). Results are plotted as relative transcript levels of percent transcript in WT (intact Cas9) compared to the transcript level in a *cas9* mutant.

Cas9-FLAG crosslinking and immunoprecipitation

DNA was crosslinked and immunoprecipitated as described previously, with the indicated modifications to optimize for *F. novicida* and Cas9-FLAG (Jaggavarapu and O'Brian, 2014). Cultures grown to OD_{600} nm of 0.6-0.8 were crosslinked by adding formaldehyde to a final concentration of 1% in 10 mM PO_4 buffer, and shaking for 10 min at RT. Reactions were quenched using 1:10 volume of 100 mg/ml glycine and shaken for 30 min at 4°C. Cells were pelleted, washed 2X in PBS, and concentrated 20X in lysis buffer (100 mM Tris-HCl pH 8.0, 300 mM NaCl, 10 mM EDTA) with protease inhibitor cocktail (Bacterial ProteaseArrest, G-Biosciences). Samples were further lysed and DNA was fractionated to 500-3000 bp fragments by sonication (15x 10 sec pulses with 15 sec pauses). Lysates were cleared by centrifugation and Cas9-FLAG was pulled-down from the supernatant using anti-Flag M2 agarose beads (Sigma). Following elution from the anti-FLAG beads, DNA was uncrosslinked from Cas9 by adding NaCl (final concentration 0.2 M) and incubating overnight at 65°C. DNA was purified (Qiagen PCR purification kit), and used as a template for qPCR. Results were normalized to the input DNA levels and a housekeeping gene (see qRT-PCR methods).

PCR amplification from cDNA

RNA and extracted from WT and Δ*cas9* (see above) and converted to cDNA using High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Segments were PCR amplified with

Phusion high-fidelity DNA polymerase (New England Biolabs) from the cDNA of each strain and compared to amplification from WT gDNA (Table S3).

Plasmid construction

Plasmids were constructed using the primers indicated in Table S3. The broad host range vector pBAV1K-T5-GFP (pBAV) was used as the control plasmid and backbone for the plasmids in all assays (Bryksin and Matsumura, 2010). Plasmids that were used to measure repression and transformation inhibition were made by replacing the promoter and RBS driving expression of *gfp* in the pBav vector with the synthetic constitutive promoter (p146) followed by different amounts of complementarity to the scaRNA tail and a PAM sequence directly upstream of *gfp* (McWhinnie and Nano, 2014). Plasmids were constructed using a Gibson Assembly Cloning Kit (NEB) and transformed and isolated from competent *E. coli* (NEB 5-alpha) using Zyppy Mini, Midi and Maxi prep kits (Zymo Research). Cas9 target plasmids containing a Cas9 crRNA spacer with and without a PAM were previously described (Price et al., 2015).

Transformation Assays

Competent cells of *F. novicida* were made by concentrating cultures at an $OD_{600\ nm}$ of 0.8-1.0 10X in 4°C chemical transformation buffer (CTB) (Llewellyn et al., 2011). For the transformations, DNA was added to 100 µl competent cells, and transformations were incubated by shaking at 37°C for 20 minutes. 1 ml of recovery media (TSB+0.2% cysteine) was then added and transformations were incubated for another 2 hours (shaking, 37°C). Transformations were plated on TSA plates with kanamycin selection and incubated at 37°C overnight. For transformations with plasmid vectors, 500 ng of plasmid was used. For transformation inhibition experiments, transformants per

100 ng plasmid were enumerated and compared between strains to determine transformation efficiency. To measure Cas9 repression of plasmid *gfp* expression, plasmids used to transform WT and *cas9* mutants. Transformants from each strain were isolated and grown in TSB+0.2% cysteine+kanamycin 30 μg/ml selection to an $OD_{600\ nm}$ of 0.8-1.0, RNA was isolated and *gfp* transcript level was measured (see qRT-PCR methods) and normalized to the kanamycin resistance cassette on pBAV to account for variations in plasmid copy number. For transformation assays with allelic exchange fragments, DNA of purified allelic exchange fragments were normalized by concentration and transformed. Transformation efficiency after 24 hours was measured by comparing the number of transformants into WT and a *cas9* mutant.

5' RACE

RNA was isolated from WT cultures as described above. 5' ends of 0544 and 0545 mRNA were mapped using the reagents and protocols from the 5'/3' RACE Kit, 2nd Generation (Roche) unless otherwise noted (primers listed in Table S3). cDNA was synthesized using the primer "GSP1" for each gene and purified with QIAquick PCR purification kit (Qiagen). PolyA-tails were added with terminal transferase and the 5' ends were amplified using "0544_GSP2" or "0545_GSP3", and "oligo dT anchor primers" (Table S3 and kit) by PCR as described above. Products were isolated by agarose gel electrophoresis, extracted using a QIAquick Gel Extraction Kit (Qiagen), and 5' ends sequenced using "0544_GSP3" or "0545_GSP2" primers (Table S3, Genewiz Sanger Sequencing).

Polymyxin susceptibility assay

Overnight cultures of WT, *scaRNA_0544/0545*, Δ*0544* and Δ*cas9*+*scaRNA_0544/0545* strains of *F. novicida* were prepared as described previously (Llewellyn et al., 2012). Strains were incubated

with 100 ❡ g/ml polymyxin B (Tokyo Chemical Industries Japan) shaking at 37°C for 6 hours and then plated to enumerate CFU surviving bacteria from each condition. Results are presented as % survival of each strain treated with polymyxin relative to untreated.

Mouse infections

Mice were infected subcutaneously with ~2 x 10^5 cfu bacteria (Weiss et al., 2007). Spleens were harvested at 48 hours post infection, homogenized in PBS, and the bacterial burden per organ was determined.

QUANTIFICATION AND STATISTICAL ANALYSIS

Prism 5 Graphpad Software was used for statistical analyses. The significance of the bacterial and culture experiments (qPCR of RNA and DNA, killing assays, transformation assays) was determined using a two-tailed student's *t*-test, for data with normal distribution. Significance was determined using the Mann-Whitney test for the mouse infections, as not all data was normally distributed. Biological replicate number and error are indicated in the figure legends. Representative gel images are shown for *in vitro* experiments.

DATA AND SOFTWARE AVAILABILITY

The RNA sequencing data reported in this paper have been deposited in the National Center for Biotechnology Information Sequence Read Archive database, http://www.ncbi.nlm.nih.gov/sra (accession no. SRP148943). The unprocessed gel images have been published at Mendeley under http://dx.doi.org/10.17632/jtjvh9m7zk.1.

Table S3. RT-PCR, Cloning, and Other Primers, Related to the STAR Methods.

See published version of this chapter in Ratner et al., Molecular Cell, 2019.

Acknowledgements

We thank Mikael Huss (SciLifeLab, Uppsala), Davide Chiarugi and Rina Ahmed-Begrich from the Charpentier lab for support with RNAseq data analysis. We are grateful to Katja Schmidt from the Charpentier lab for the purification of FnoCas9. We thank the members of the Weiss and Charpentier labs for their feedback and assistance, and Charles Moran and Phil Rather at Emory University for their experimental insight. This work was supported by National Institutes of Health (NIH) grants U54-AI057157 from the Southeastern Regional Center of Excellence for Emerging Infections and Biodefense, as well as R01-AI110701 to David S. Weiss, who is also supported by a Burroughs Wellcome Fund Investigator in the Pathogenesis of Infectious Disease award. We also thank the Alexander von Humboldt Foundation (Alexander von Humboldt Professorship to E.C.), the German Research Foundation (Leibniz Prize to E.C.), and the Max Planck Society for financially supporting this research study in the Charpentier lab.

Author contributions

H.K.R. A.E.M, A.L.R., E.C. and D.S.W conceived and designed this study. H.K.R, A.E.M, and A.L.R conducted the majority of the experiments. S.J. designed and assisted with the Cas9 pulldown assays and made the Δ*1104-1101* mutant. J.E.W. conducted the polymyxin killing assay. E.K.C. constructed the Δ*cas9*Δ*1104* mutant. H.K.R., A.E.M., A.L.R., E.C., and D.S.W. wrote the manuscript.

Declaration of Interests

The authors have filed a provisional patent application related to this work and declare no other competing interests.

References

Anders, C., Niewoehner, O., Duerst, A., and Jinek, M. (2014). Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. Nature *513*, 569-573.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics (Oxford, England) *31*, 166-169.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science *315*, 1709-1712.

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 research *41*, 7429-7437.

Bolotin, A., Quinquis, B., Sorokin, A., and Ehrlich, S.D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology (Reading, England) *151*, 2551-2561.

Bryksin, A.V., and Matsumura, I. (2010). Rational design of a plasmid origin that replicates efficiently in both gram-positive and gram-negative bacteria. PloS one *5*, e13244.

Chylinski, K., Le Rhun, A., and Charpentier, E. (2013). The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. RNA Biology *10*, 726-737.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature *471*, 602-607.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Dötsch, A., Eckweiler, D., Schniederjans, M., Zimmermann, A., Jensen, V., Scharfe, M., Geffers, R., and Häussler, S. (2012). The *Pseudomonas aeruginosa* Transcriptome in Planktonic Cultures and Static Biofilms Using RNA Sequencing. PloS one *7*, e31092.

Dugar, G., Leenay, R.T., Eisenbart, S.K., Bischler, T., Aul, B.U., Beisel, C.L., and Sharma, C.M. (2018). CRISPR RNA-Dependent Binding and Cleavage of Endogenous RNAs by the *Campylobacter jejuni* Cas9. Molecular Cell *69*, 893-905.e897.

Fonfara, I., Le Rhun, A., Chylinski, K., Makarova, K.S., Lecrivain, A.L., Bzdrenga, J., Koonin, E.V., and Charpentier, E. (2014). Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. Nucleic acids research *42*, 2577-2590.

Gallagher, L.A., McKevitt, M., Ramage, E.R., and Manoil, C. (2008). Genetic Dissection of the *Francisella novicida* Restriction Barrier. Journal of Bacteriology *190*, 7830-7837.

Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proceedings of the National Academy of Sciences *109*, E2579–E2586.

Jaggavarapu, S., and O'Brian, M.R. (2014). Differential control of *Bradyrhizobium japonicum* iron stimulon genes through variable affinity of the iron response regulator (Irr) for target gene promoters and selective loss of activator function. Molecular microbiology *92*, 609-624.

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.

Jinek, M., Jiang, F., Taylor, D.W., Sternberg, S.H., Kaya, E., Ma, E., Anders, C., Hauer, M., Zhou, K., Lin, S.*, et al.* (2014). Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. Science *343*, 1247997.

Jones, C.L., Sampson, T.R., Nakaya, H.I., Pulendran, B., and Weiss, D.S. (2012). Repression of bacterial lipoprotein production by *Francisella novicida* facilitates evasion of innate immune recognition. Cellular microbiology *14*, 1531-1543.

Kanistanon, D., Hajjar, A.M., Pelletier, M.R., Gallagher, L.A., Kalhorn, T., Shaffer, S.A., Goodlett, D.R., Rohmer, L., Brittnacher, M.J., Skerrett, S.J.*, et al.* (2008). A *Francisella* mutant in lipid A carbohydrate modification elicits protective immunity. PLoS pathogens *4*, e24.

Larson, M.H., Gilbert, L.A., Wang, X., Lim, W.A., Weissman, J.S., and Qi, L.S. (2013). CRISPR interference (CRISPRi) for sequence-specific control of gene expression. Nature protocols *8*, 2180-2196.

Llewellyn, A.C., Jones, C.L., Napier, B.A., Bina, J.E., and Weiss, D.S. (2011). Macrophage Replication Screen Identifies a Novel *Francisella* Hydroperoxide Resistance Protein Involved in Virulence. PloS one *6*, e24201.

Llewellyn, A.C., Zhao, J., Song, F., Parvathareddy, J., Xu, Q., Napier, B.A., Laroui, H., Merlin, D., Bina, J.E., Cotter, P.A.*, et al.* (2012). NaxD is a deacetylase required for lipid A modification and *Francisella* pathogenesis. Molecular microbiology *86*, 611-627.

Louwen, R., Horst-Kreft, D., de Boer, A.G., van der Graaf, L., de Knegt, G., Hamersma, M., Heikema, A.P., Timms, A.R., Jacobs, B.C., Wagenaar, J.A.*, et al.* (2013). A novel link between Campylobacter jejuni bacteriophage defence, virulence and Guillain-Barre syndrome. European journal of clinical microbiology & infectious diseases: official publication of the European Society of Clinical Microbiology *32*, 207-226.

Louwen, R., Staals, R.H., Endtz, H.P., van Baarlen, P., and van der Oost, J. (2014). The role of CRISPR-Cas systems in virulence of pathogenic bacteria. Microbiology and molecular biology reviews : MMBR *78*, 74-88.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology *15*, 550.

Ma, K., Cao, Q., Luo, S., Wang, Z., Liu, G., Lu, C., and Liu, Y. (2018). cas9 Enhances Bacterial Virulence by Repressing the regR Transcriptional Regulator in Streptococcus agalactiae. Infection and immunity *86*.

Marraffini, L.A., and Sontheimer, E.J. (2008). CRISPR Interference Limits Horizontal Gene Transfer in Staphylococci by Targeting DNA. Science *322*, 1843-1845.

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

McWhinnie, R.L., and Nano, F.E. (2014). Synthetic promoters functional in Francisella novicida and Escherichia coli. Applied and environmental microbiology *80*, 226-234.

Mojica, F.J., Diez-Villasenor, C., Garcia-Martinez, J., and Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology (Reading, England) *155*, 733-740.

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. Journal of molecular evolution *60*, 174-182.

Nishimasu, H., Ran, F.A., Hsu, P.D., Konermann, S., Shehata, S.I., Dohmae, N., Ishitani, R., Zhang, F., and Nureki, O. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell *156*, 935-949.

O'Connell, M.R., Oakes, B.L., Sternberg, S.H., East-Seletsky, A., Kaplan, M., and Doudna, J.A. (2014). Programmable RNA recognition and cleavage by CRISPR/Cas9. Nature *516*, 263-266.

Postic, G., Frapy, E., Dupuis, M., Dubail, I., Livny, J., Charbit, A., and Meibom, K.L. (2010). Identification of small RNAs in Francisella tularensis. BMC genomics *11*, 625.

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 (Reading, England) *151*, 653-663.

Price, A.A., Sampson, T.R., Ratner, H.K., Grakoui, A., and Weiss, D.S. (2015). Cas9-mediated targeting of viral RNA in eukaryotic cells. Proceedings of the National Academy of Sciences of the United States of America *112*, 6164-6169.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell *152*, 1173-1183.

Ratner, H.K., Sampson, T.R., and Weiss, D.S. (2015). I can see CRISPR now, even when phage are gone: a view on alternative CRISPR-Cas functions from the prokaryotic envelope. Current opinion in infectious diseases *28*, 267-274.

Rousseau, B.A., Hou, Z., Gramelspacher, M.J., and Zhang, Y. (2018). Programmable RNA Cleavage and Recognition by a Natural CRISPR-Cas9 System from Neisseria meningitidis. Mol Cell *69*, 906-914.e904.

Sampson, T.R., Napier, B.A., Schroeder, M.R., Louwen, R., Zhao, J., Chin, C.Y., Ratner, H.K., Llewellyn, A.C., Jones, C.L., Laroui, H.*, et al.* (2014). A CRISPR-Cas system enhances envelope integrity mediating antibiotic resistance and inflammasome evasion. Proceedings of the National Academy of Sciences of the United States of America *111*, 11163-11168.

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.

Sapranauskas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., and Siksnys, V. (2011). The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. Nucleic acids research *39*, 9275-9282.

Sternberg, S.H., LaFrance, B., Kaplan, M., and Doudna, J.A. (2015). Conformational control of DNA target cleavage by CRISPR-Cas9. Nature *527*, 110-113.

Strutt, S.C., Torrez, R.M., Kaya, E., Negrete, O.A., and Doudna, J.A. (2018). RNA-dependent RNA targeting by CRISPR-Cas9. eLife *7*.

Weiss, D.S., Brotcke, A., Henry, T., Margolis, J.J., Chan, K., and Monack, D.M. (2007). In vivo negative selection screen identifies genes required for Francisella virulence. Proceedings of the National Academy of Sciences of the United States of America *104*, 6037-6042.

Westra, E.R., Buckling, A., and Fineran, P.C. (2014). CRISPR-Cas systems: beyond adaptive immunity. Nat Rev Micro *12*, 317-326.

Zhang, T., Zheng, Q., Yi, X., An, H., Zhao, Y., Ma, S., and Zhou, G. (2018). Establishing RNA virus resistance in plants by harnessing CRISPR immune system. Plant biotechnology journal *16*, 1415-1423.

Chapter 4. *F. novicida* CRISPR-Cas systems can functionally complement each other in DNA defense while providing target flexibility

Hannah K. Ratner[1,2,3] and David S. Weiss,[2,3,4,*]

[1]Microbiology and Molecular Genetics Program, Emory University, Atlanta, GA 30329.

[2]Emory Vaccine Center, Emory University, Atlanta, GA 30329.

[3]Yerkes National Primate Research Center, Emory University, Atlanta, GA 30329.

[4]Division of Infectious Diseases, Department of Medicine, Emory University School of Medicine, Atlanta, Georgia 30329.

*Corresponding author:

David S. Weiss

Emory Vaccine Center

954 Gatewood Rd, Room 2028

Atlanta, GA 30329

Tel: (404) 727-8214

Abstract

CRISPR-Cas systems are prokaryotic adaptive immune systems that facilitate protection of bacteria and archaea against infection by foreign mobile genetic elements. The model pathogen *Francisella novicida* encodes a CRISPR-Cas12a (FnoCas12a) system and a CRISPR-Cas9 (FnoCas9) system, the latter of which has an additional and non-canonical function in bacterial virulence. Here, we investigated and compared the functional roles of the FnoCas12a and FnoCas9 systems in transformation inhibition and bacterial virulence. Unlike FnoCas9, FnoCas12a was not required for *F. novicida* virulence. However, both systems were highly effective at plasmid restriction and acted independently of each other. We further identified a critical protospacer adjacent motif (PAM) necessary for transformation inhibition by FnoCas12a, demonstrating a greater flexibility for target identification by FnoCas12a than previously appreciated, and specificity that is distinct from that of FnoCas9. The two systems exhibited different patterns of expression, suggesting they may confer distinct benefits to the bacterium in different conditions. Together, these data suggest that the differences between these systems provide *F. novicida* with a more comprehensive defense against foreign nucleic acids. We subsequently harnessed this information and demonstrated that the FnoCas12a and FnoCas9 machineries can be simultaneously reprogrammed to restrict the same non-native target, expanding the toolset for prokaryotic genome manipulation.

Keywords: CRISPR-Cas, DNA targeting, plasmid restriction, Cas12a, Cpf1, Cas9, *Francisella*

Introduction

CRISPR (clustered, regularly interspaced, short, palindromic repeats) - Cas (CRISPR-associated) are prokaryotic adaptive immune systems that protect against mobile genetic elements like bacteriophages and plasmids (14, 16, 89, 122, 123). These conserved genomic loci contain a CRISPR array and adjacent *cas* genes (10, 11, 36). The CRISPR array consists of alternating repetitive sequences, interspaced by unique spacer sequences that often correspond with sequences of foreign nucleic acid targets, and these targets are referred to as protospacers (1) (8, 9).

Upon introduction of a foreign nucleic acid, the invading DNA is selected, processed and integrated into the CRISPR array with flanking repeat sequences (38, 124). To protect against subsequent infection by a nucleic acid that shares the same sequence, the CRISPR array is transcribed and processed into individual CRISPR RNAs (crRNAs), each comprised of one spacer and part, if not all, of the repeat sequence. The crRNAs form complexes with Cas protein(s) and when the spacer sequence of the crRNA binds to the protospacer in the target, the associated Cas protein(s) cleave the invading foreign DNA to protect the cell (18-20).

CRISPR-Cas systems are mechanistically diverse and can be divided into multiple classes (36). Class I systems require a multi-protein complex for target recognition and cleavage, while Class II systems consist of a single effector protein for these processes. Both can be further subdivided into Types based on phylogeny, gene clustering, and associated RNAs (36). The Class II system that uses the Cas9 protein for DNA recognition and cleavage has been extensively characterized, and the relative simplicity of a single effector protein has enabled the development of revolutionary Cas9-derived tools for genetic engineering (125-127). Other systems with single effector proteins have been discovered that have critical differences

compared to Cas9. Of these systems, the CRISPR-Cas12a system (also known as Cas12a) has been the focus of a notable research due to unique characteristics that benefit genome engineering applications. Distinctions of Cas12a include processing of its own crRNAs by the Cas12a protein, use of a single crRNA and no accessory RNAs to cleave DNA targets, and production of staggered double strand breaks. Despite the extensive study of these Cas9 and Cas12a CRISPR systems in non-native cell types and for engineering, few of these functions have been studied in their native bacterial contexts, which is essential for understanding the natural biological functions of these wide-spread prokaryotic defense systems.

Interestingly, the intracellular pathogen *Francisella novicida* encodes both a CRISPR-Cas9 (FnoCas9) and a CRISPR-Cas12a (FnoCas12a) system (119). *F. novicida* is a model for the native functions of Cas9 beyond DNA cleavage because in this bacterium Cas9 is also essential for virulence of the pathogen. Although the role of FnoCas9 in virulence has been established, the role of FnoCas9 and FnoCas12a in DNA defense, as well as the role of FnoCas12a in virulence have yet to be explored. Both systems encode CRISPR repeats targeting various sites throughout the same putative *F. novicida* prophage genome. This prophage has only been found in strains lacking complementary spacers, suggesting functional *in vivo* activity of these systems in bacteriophage defense (119). *cas12a*-encoding CRISPR-Cas loci are conserved in diverse bacterial species (e.g. *Prevotella*, *Flavobacterium)*, and are found almost exclusively in mammalian host-associated (commensal and pathogenic) bacteria, similar to the ecological distribution of CRISPR-Cas9 systems (35, 36, 119). Interestingly, unlike CRISPR-Cas9 systems that are unique to bacteria, Cas12a-associated systems are also found in archaea and have both DNase and RNase activity, allowing DNA targeting and crRNA processing (36, 128).

Herein, we demonstrate that the CRISPR-Cas12a system defends against foreign DNA invasion within its native bacterial host using Cas12a and a crRNA. Since *F. novicida* contains both CRISPR-Cas9 and -Cas12a systems, we compared their physiological activities, revealing that their functions in DNA defense were independent and that each effector protein exhibited a different pattern of expression, providing the first description of how two Class II systems function relative to one another in the same bacterium. Conversely, we found that only FnoCas9 was important for *F. novicida* virulence, while a loss of FnoCas12a did not alter the fitness of the bacterium in a mammalian host. To better elucidate the functional differences between these two systems in DNA cleavage, we demonstrated that FnoCas12a has underappreciated flexibility in the sequence used to identify self and non-self sequences, known as the protospacer adjacent motif (PAM). Finally, we compared the baseline efficiency of the *F. novicida* Cas12a and Cas9 systems in DNA defense by reprogramming CRISPR arrays for the same, non-native target, and found that they protected the bacterium with remarkably similar efficacy. This similarity suggests that functional differences between FnoCas12a and FnoCas9 are not in their baseline ability to restrict infection or the molecular differences in their DNA cleavage mechanism, but rather in other aspects of their function, including regulation and PAM requirements. These findings indicate that multiple Class II CRISPR-Cas systems can exist in a bacterium to provide both robust and comprehensive protection from mobile genetic elements and that they can be further utilized to modulate bacterial physiology through the programmed targeting of DNA.

Results

FnoCas9 contributes to *F. novicida* virulence independently of FnoCas12a

Given the described role of FnoCas9 in *F. novicida* virulence, we first characterized the role of FnoCas9 and FnoCas12a during infection, together and independently. Mice were infected subcutaneously and the bacterial burden in the skin and spleen was measured 48 hr post infection (Figure 1 A-B). WT *F. novicida* established robust infection in both organs, while a Δ*cas9* mutant was highly attenuated. Mutation of *cas12a* from WT *F. novicida* had no effect on the ability of the pathogen to establish infection, nor was a Δ*cas9*Δ*cas12a* double mutant further attenuated from the levels of the Δ*cas9* strain *in vivo*. These results are the first to indicate a divergence in the functional roles of these two systems in their native contexts, with FnoCas9 exhibiting an alternative function in virulence, while FnoCas12a did not alter bacterial fitness *in vivo*. This suggests that FnoCas12a, and possibly FnoCas9, may be providing the more canonical function of CRISPR-Cas systems, protecting *F. novicida* from potentially harmful foreign genetic elements.

Endogenous *F. novicida* CRISPR systems function independently in DNA defense

The first indication that FnoCas9 and FnoCas12a defend *F. novicida* against foreign DNA was in the spacer sequences in the CRISPR arrays. The *F. novicida* U112 *cas12a*-associated CRISPR array contains 9 spacers, two of which target a putative prophage that was identified in *F. novicida* strain 3523, also targeted by the *F. novicida* U112 CRISPR-Cas9 system (119). To determine if *F. novicida* U112 is resistant to infection by protospacer-containing sequences, and in the absence of viable bacteriophage to use in such experiments, we performed a plasmid inhibition transformation assay to mimic infection by a mobile genetic element. Wild-type *F. novicida* was transformed with a plasmid containing a 50 base pair (bp) sequence from the putative prophage encompassing the 30 bp protospacer and 10 bases on either side (referred to as the "Cas12a_target_5" plasmid, as it corresponds to the 5th spacer within the Cas12a-CRISPR

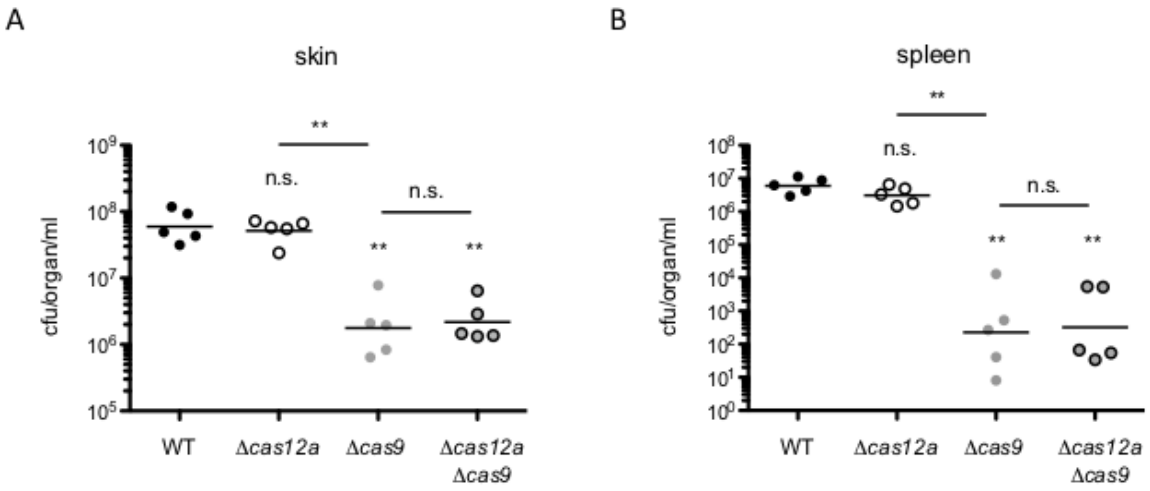


Figure 1. Cas9 contributes to F. novicida virulence independently of Cas12a. A-B) CFU burden in mouse (A) skin and (B) spleens 48 hr post-infection with U112 (WT), **Δ***cas12a*, **Δ***cas9***, Δ***cas9***Δ***cas12a* (*n*=5, bar represents geometric mean; **p≤0.005).

array) inserted into a non-transcribed portion of the plasmid (Figure 2A). Wild-type bacteria resisted transformation with the Cas12a_target_5 plasmid while remaining permissive to transformation by the plasmid control (the target plasmid backbone without the inserted prophage region; from here on referred to as the control plasmid).

We next tested whether there was interdependency between the endogenous *F. novicida* Cas9 and Cas12a systems in DNA cleavage by measuring the ability of Cas12a to inhibit transformation with a Cas12a_target_5 plasmid in the presence and absence of Cas9. *F. novicida* restricted transformation with the Cas12a_target_5 plasmids while remaining permissive to transformation with the control, while the Δ*Cas12a* strain was permissive to transformation with both plasmids. However, when a Δ*cas9* strain was transformed with the Cas12a_target_5 plasmid, it was able to restrict transformation with this plasmid, suggesting that Cas12a-dependent targeting occurs independently of Cas9 activity (Fig. 2A). We next tested the independence of the Cas9 system of *F. novicida* in DNA cleavage. Transformation inhibition assays were performed with a Cas9_target_13 plasmid. The Cas9_target_13 plasmid contained the protospacer for spacer #13 of the CRISPR-Cas9 array flanked by a PAM sequence that had previously been determined for *F. novicida* Cas9 by *in vitro* and *in vivo* DNA targeting assays (88, 129, 130). We observed that *F. novicida* could inhibit transformation by the Cas9_target_13 plasmid, and that this inhibition was dependent on an intact Cas9 (Fig. 2B). Interestingly, *F. novicida* inhibited transformation by both Cas12a- and Cas9-targeted plasmids by approximately 3.5 logs, suggesting that in these transformation conditions, both proteins exhibited a similar ability to restrict foreign DNA (Fig. 2A-B). This is the first examination of plasmid restriction by

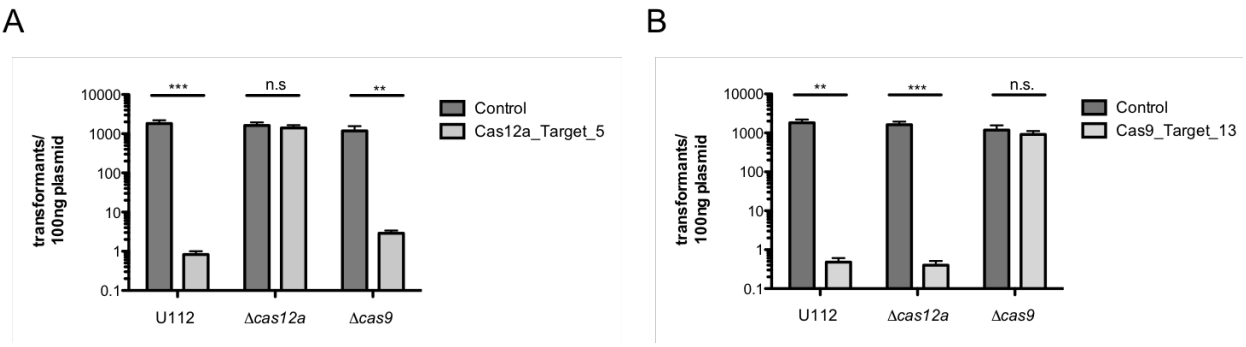


Figure 2. Cas12a and Cas9 have distinct targets that they inhibit with similar efficiencies. A) Wild-type *F. novicida*, Δ*cas9*, and Δ*cas12a* were transformed with the control and the Cas12a_target_5 plasmid **B)** Wild-type *F. novicida*, Δ*cas9*, and Δ*cas12a* were transformed with the control and a Cas9_target_13 plasmid that contains a protospacer that is complementary to spacer 13 from the *cas9*-associated CRISPR array. Both Cas9 and Cas12a inhibit transformation by spacers in their respective loci with similar efficiencies, and inhibition occurs in the absence of the effector from the other locus. **, $P<0.005$; ***, $P<0.0001$.

two distinct Class II CRISPR could occur efficiently and independently of the other. Thus, the presence of these two systems gives *F. novicida* dual defenses against invading nucleic acid threats.

Cas12a exhibits PAM promiscuity in native host

The ability of *F. novicida* to protect against transformation by sequences targeted by Cas12a-associated spacers led us to investigate the flexibility of the PAM sequence the CRISPR-Cas12a system (22, 25, 131, 132). *In vitro* structural and heterologous expression studies have identified that the FnoCas12a spacer sequence required for plasmid interference is 5' TN (24, 78, 133, 134). To identify the preferred PAM of FnCas12a within *Francisella*, we aligned the 5' and 3' flanking nucleotides of putative prophage regions that are complementary to 9 unique spacers found in different strains of *Francisella* (119). This analysis predicted a 5' TTTN PAM motif.

Considering this PAM in the context of the high AT content genome of *F. novicida*, we questioned the stringency of FnoCas12a PAM requirements *in vivo*. To determine whether each of the three T residues within the putative PAM are required for target recognition, we conducted a transformation assay with point mutants in the predicted PAM sequence of the Cas12a_target_5 plasmid, and compared transformation efficiency as the ratio of colony forming units (CFUs) recovered from the Δ*Cas12a* mutant to that of the CFUs recovered from wild-type. We mutated the -2 position T (5' to the protospacer) to a C in the Cas12a_target_5 plasmid, resulting in a 5' TTCA mutant PAM (construct A, Fig. 3C). Targeting of the 5' TTCA plasmid was largely abrogated, indicating that the -2 T is important for target recognition. However, there was also a 24-hour delay in growth of the transformants with this mutant PAM, further

suggesting that additional bases are involved in plasmid recognition. Conversely, mutation of the -3 position T to C (construct B) did not alter plasmid targeting efficiency, indicating that this residue is non-essential for targeting (Fig. 3C). We then tested a target plasmid harboring a -4 T to C mutation (construct C). Targeting of this plasmid was slightly less efficient compared to the Cas12a_target_5 plasmid encoding the 5' TTTN PAM, suggesting that the -4 PAM residue plays a role in DNA recognition (Fig. 3C). The combined mutation of the -3 and -4 position TT residues to CC yielded a 5' CCTN mutant PAM, which more severely inhibited plasmid targeting compared to the -4 T to C mutation alone (Fig. 3C). This suggests the -3 nucleotide plays a role in target recognition, but does so in tandem with the flanking bases in the PAM (Fig. 3B-C, construct D). Because each of the -2 to -4 position bases in the predicted 5' TTTN PAM had a partial contribution to plasmid targeting, we generated a protospacer construct with the 5' TTTN PAM nucleotides mutated to 5' CCCN (construct E). The 5' CCCN plasmid was not targeted effectively, and transformation efficiency was recovered to almost that of the control plasmid (Fig. 3C). From these data, we conclude that *F. novicida* Cas12a requires a 5' TTTN PAM, and that the three T residues play an additive role in target DNA inhibition. This is distinct from the PAM required for the other Class II CRISPR system in *F. novicida*, Cas9, which requires a 3' NGG, highlighting the mechanistic differences between these systems that together provide more comprehensive protection from mobile genetic elements (88, 129, 130).

Cas12a and Cas9 follow different patterns of expression during transformation

Due to the functional redundancy in DNA cleavage and distinct differences in PAM requirements of the *F. novicida* Cas9 and Cas12a systems, we considered the possibility that the

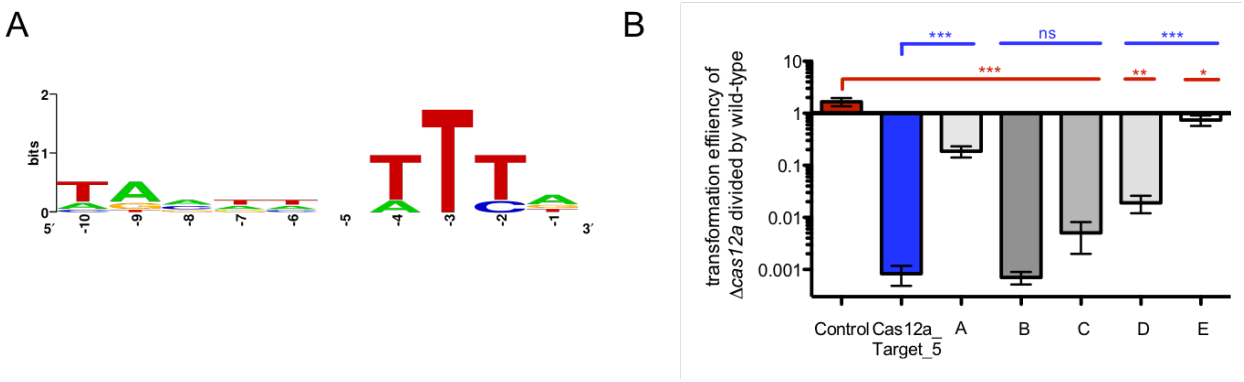


Figure 3. Cas12a requires a 5' protospacer-adjacent motif (PAM). A) sequence logo plot for the 5' flanking sequence of 9 protospacers complementary to 9 unique spacers from different strains of *Francisella*. Letter height represents the nucleotide frequency at each position relative to the protospacer for these samples. B) Wild-type *F. novicida* U112 and a Δ*cas12a* mutant were transformed with Control and Cas12a_Target_5 plasmid derivatives in (C), with a 5' CCCN PAM mutation was most able to restore transformation efficiency after 48 hours post transformation. (red indicates significance to control by two-tailed t-test, blue indicates significance relative to target by two-tailed t-test) C) 5' sequence alignment for PAM mutations made in the Cas12a_Target_5 plasmid. Bases that have been mutated from the predicted PAM are underlined and bolded. *, $P<0.05$; **, $P<0.005$; ***, $P<0.0001$.

two CRISPR systems may be specialized to protect *F. novicida* from mobile genetic elements that arise in different environmental niches. *F. novicida* likely needs adaptive immunity in the soil, brackish water, and mammalian hosts, among other environments. To test whether the two systems can exhibit different patterns of expression, we quantified expression of *Cas9* and *Cas12a* during the transformation conditions used in plasmid interference experiments (competent cells prior to transformation, during transformation, and after recovery from transformation in growth media) relative to the starting culture of log-phase *F. novicida* U112 (OD_{600} 1.0). Interestingly, we observed that *Cas12a* expression increased between the starting culture and the competent cell conditions, appearing to continue to increase slightly during transformation. *cas9* expression is highest in the competent cells, spiking when the cells were moved into a cold ionic buffer and then dropping over the remainder of transformation, to below the expression level of the starting culture in the recovery stage (Fig. 4). These results suggest that the *F. novicida* CRISPR-Cas systems may be regulated to provide protection in different environments, and that the CRISPR-Cas systems' mechanistic differences may specialize them for threats specific to those different conditions.

Cas9 and Cas12a have the same baseline ability to restrict foreign DNA in *F. novicida*

We next sought to directly compare the efficiency of plasmid restriction of the *F. novicida* CRISPR systems, without the competing spacers in the CRISPR arrays and normalized by using both systems to target the same non-native target sequence. The PAM sequences for Cas12a and Cas9 are located on opposing sides of the protospacers, allowing for the same

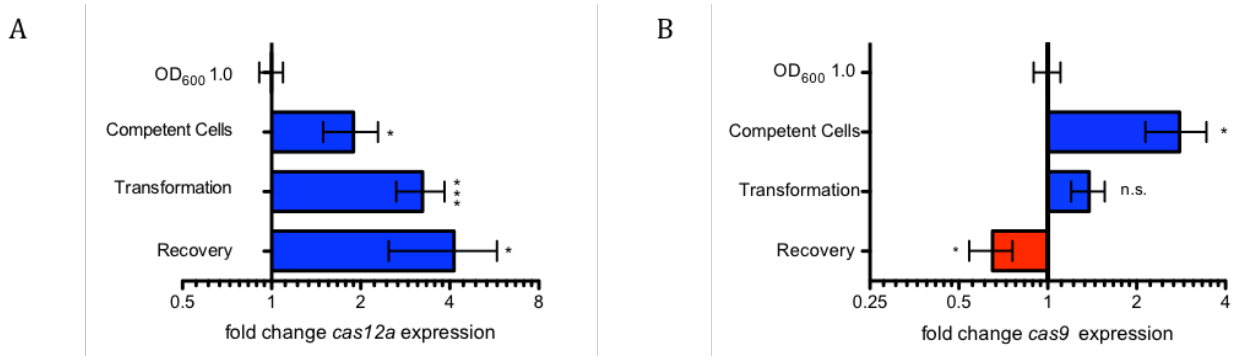


Figure 4. *cas12a* and *cas9* have different expression patterns during transformation. Relative expression of *cas12a* and *cas9* was examined by RT-PCR during plasmid transformation conditions at OD_{600} 1.0, competent cells, transformation, and after 45 minutes and 120 minutes of recovery. Each experiment was normalized to a housekeeping gene and the expression level in the starting culture (OD_{600} 1.0), resulting in the fold change in expression between conditions. Blue bars indicate an increase in expression from the starting culture, red bars indicate a decrease in expression. **A)** fold-change in *cas12a* expression over transformation conditions normalized to OD_{600} 1.0. Cpf1 expression increases between the starting culture and the transformation conditions. **B)** fold-change in *cas9* expression over transformation conditions normalized to OD_{600} 1.0. Cas9 expression is the highest in the competent cells, and drops below the starting expression level in the transformation recovery. *, $P<0.05$.

sequence to be targeted by both systems. Each CRISPR array was replaced with a single non-native spacer flanked on both sides by the complete crRNA repeat sequence of the respective system. The two CRISPR systems were otherwise unmodified, with the repeat-spacer-repeat sequences placed downstream the native Cas9 or Cas12a CRISPR array promoters (Fig. 5A). A strain with the Cas12a CRISPR array reprogrammed for the new target and an intact WT Cas9 system, and a strain with Cas9 reprogrammed for the new target but with an intact WT Cas12a system, were transformed with a plasmid containing a protospacer target for the reprogrammed systems (Fig. 5B). In this vector, called the "Cas12a&Cas9_target" plasmid, the protospacer was flanked by the PAM for both CRISPR systems to provide a universal target plasmid.

Reprogramming of both the Cas12a and Cas9 systems successfully restricted transformation with the target plasmid to just above the limit of detection of the assay (Fig. 5B). We then reprogrammed both CRISPR loci for the new target in the same strain, and evaluated the ability to restrict transformation with the target plasmid, finding that this strain also restricted transformation with high efficiency (Fig. 5B). This first controlled comparison of two Class II CRISPR system activities in their native bacterial host suggests that in spite of mechanistic differences, in the absence of competing crRNAs, the Cas9 and Cas12a CRISPR systems of *F. novicida* have remarkably similar abilities to restrict infection with a spacer-encoded target, with an almost undetectable level of plasmid escape.

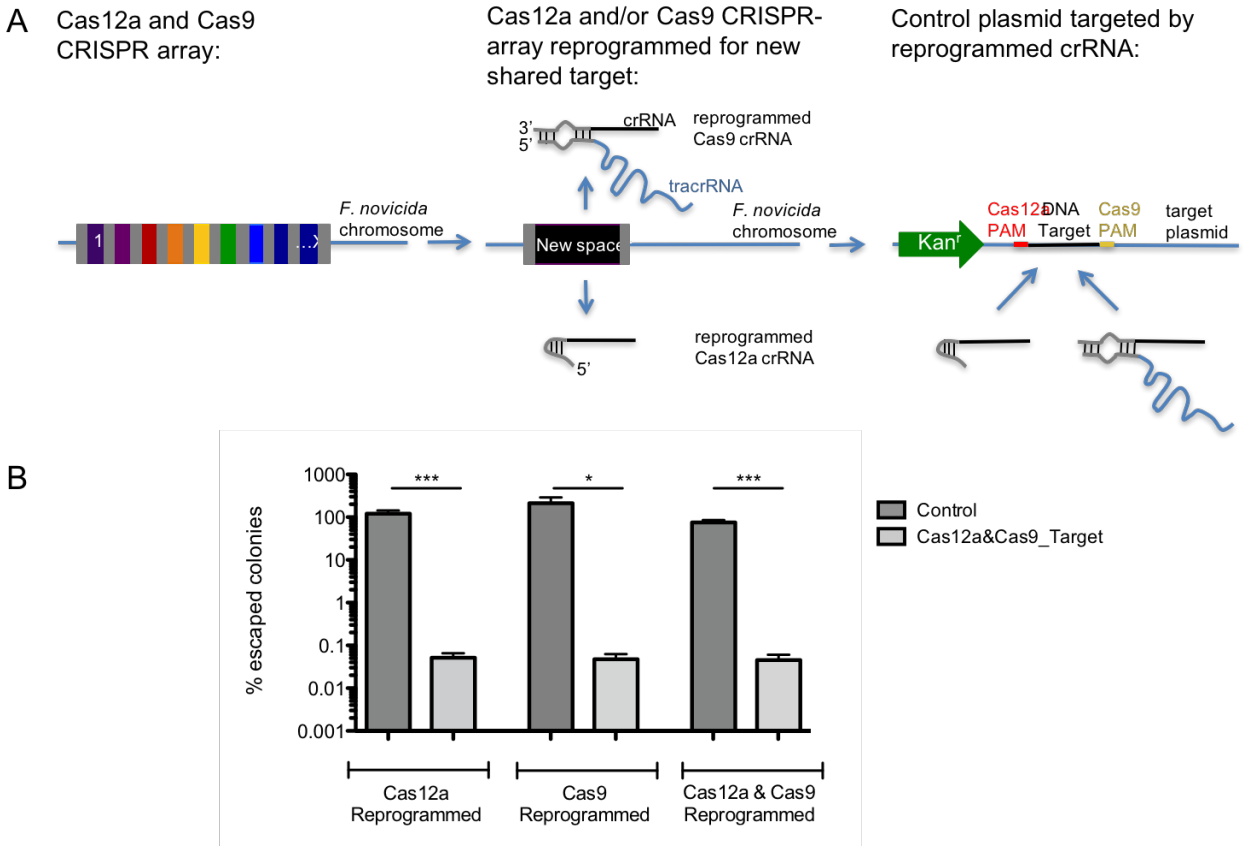


Figure 5. Cas9 and Cas12a exhibit similar endogenous DNA targeting efficiencies when reprogrammed for the same artificial target. A) Schematic of a CRISPR array and the reprogrammed Cas9 and Cas12a CRISPR loci. In the reprogrammed loci, the ΔcrRNA complemented with a repeat-(non-native spacer)-repeat, using repeats specific for each CRISPR system. The reprogrammed Cas12a spacer is transcribed and processed into a mature crRNA that interacts with Cas12a independently of accessory RNAs. The reprogrammed Cas9 spacer is transcribed and processed into a mature crRNA that interacts with the tracrRNA, and the resulting RNA duplex binds to Cas9 and guides the effector to its target. B) Wild-type *F. novicida* U112, a strain with Cas12a reprogrammed, a strain with Cas9 reprogrammed, and a strain with both Cas12a and Cas9 loci reprogrammed for the same non-native spacer were transformed with Control and Cas12a&Cas9_Target plasmids. The % of colonies escaping targeting was calculated relative to the untargeted control plasmid *, $P<0.05$; **, $P<0.005$; ***, $P<0.0001$.

Discussion

F. novicida harbors two distinct endogenous Class II CRISPR systems, CRISPR-Cas12a and CRISPR-Cas9, providing a unique model for studying natural Class II CRISPR system functions relative to one another (119). Not only does the FnoCas9 system enable virulence in a mammalian host, but the native presence of two CRISPR-Cas systems in this bacterium represents a two-pronged CRISPR-defense; both systems contain spacers that target the same putative prophage, which suggests that there is a fitness benefit to retaining the two DNA targeting Class II systems (119). We demonstrated that these two systems inhibit transformation with their respective targets with similar efficiencies, and that they do so independently (Fig. 1).

Furthermore, we show that both Cas12a and Cas9 CRISPR systems can be reprogrammed for a new target. Upon reprogramming, in the absence of competing variables, both system restrict transformation with a target plasmid at the same efficiency, with an extremely low frequency of plasmid escape (Fig. 5). When both systems are reprogrammed for the same target in a single strain, the efficiency of plasmid restriction is not increased above that of each individual system (Fig 5). This suggests that there may not be significant importance to having the second system available to compensate for possible PAM mutations in the target that would prevent recognition by one of the systems. Therefore, the repurposing of both systems for different targets within a bacterium could provide maximally efficient protection from two sequences. These results indicate that with simple genome manipulations, bacteria and archaea encoding CRISPR-Cas12a and/or Cas9 systems can be re-programmed to target invading and emerging nucleic acid threats that are not already targeted by complementary sequences in the native spacer array.

In spite of their baseline similarities in DNA restriction upon reprogramming, we identify key differences between the two systems that allow them to provide *F. novicida* with more comprehensive protection from nucleic acid threats. We show that the non-overlapping PAM requirements of Cas9 and Cas12a, and the flexibility of the Cas12a PAM, provide a level of target diversity that far exceeds that of either system individually (Fig 3). Likewise, we show that the two systems follow distinct expression patterns over the course of transformation (Fig. 4). We hypothesize that these differences allow the bacterium to take maximal advantage of the mechanistic differences between the two systems for adaptive immunity, without expending energy on producing large effector proteins unnecessarily. Additional investigation into the differential regulation of these two CRISPR-Cas systems in the presence of different types of nucleic acid predation may provide insight into the unique specializations of these two systems.

The differences in the molecular DNA cleavage mechanisms of Cas9 and Cas12a, which include inducing different double stranded breaks in their targets, crRNA processing, and presence of a tracrRNA, likely result in functional differences in exogenous DNA protection. We confirmed and interrogated aspects of the FnoCas12a mechanism in the native bacterial host, a context they had not been examined in previously. However, due to the lack of a viable phage for *Francisella*, we are unable to test how these differences effect adaptive viral immunity in *F. novicida*. Never-the-less, we can make informed predictions based on the composition of the CRISPR arrays about the Cas9 and Cas12a system functions. Given the presence of spacers for the same putative prophage in both CRISPR arrays, both systems can likely protect against overlapping invading elements, suggesting that the conditions they are expressed in and differences in their target selection (PAM) and adaptation requirements likely play as, if not

more, of a significant role in their respective differences in DNA defense than their molecular mechanisms of crRNA processing and cleavage.

Interestingly, unlike Cas9, the FnoCas12a PAM is located to the 5' of the protospacer. Because of this, the same highly virulent sequence could theoretically be targeted by both the Cas9 and Cas12a systems to increase protection of the bacterium. We tested this theory by reprogramming both systems in the same strain for the same non-native target, and transforming with a plasmid containing the protospacer flanked by the PAMs for both systems. Interestingly, this did not increase the efficiency of protection from the strains with individually reprogrammed loci, suggesting that the difference in PAM orientation is the product of the cleavage mechanism of each system which evolved independently of the sustained presence of the other system.

However, one benefit of two similarly functional Class II systems is the recognition diversity that is enabled by two Cas effectors with distinct PAMs. *In vitro* and heterologous expression FnoCas12a studies have suggested that a 5' TN is the necessary PAM sequence (128, 135). We bioinformatically predicted that the preferred PAM for FnoCas12a is 5' TTTN PAM, and further interrogated the PAM requirements of FnoCas12a by evaluating its ability to restrict plasmids with systematic mutation in the TTTN PAM sequence while in its natural bacterial host. We found that FnoCas12a actually exhibits a high level of flexibility in the PAM sequences that it recognizes. We observed that of the three Ts in the PAM sequence, the -2 position T is the most important and that the -4 T was dispensable for target DNA recognition and targeting *in vivo* (Fig. 4B). While the -2 position T is important, it is not sufficient to fully restore targeting efficiency. We observe that a 5' TTTN is the preferred PAM sequence *in vivo*, with the 5' CCCN mutant PAM having the largest effect on abrogating target plasmid inhibition (Fig. 4B). This differs from the *in vitro* cleavage data for FnoCas12a that suggests a 5' TN PAM, and more

closely resembles the 5' TTTN PAM of AsCas12a (136). These results highlight the importance of studying CRISPR-Cas systems within their native bacterial hosts in addition to through heterologous expression and *in vitro* models. Not only have we identified differences in the requirements of the Cas12a system for target recognition, but we have demonstrated the importance of studying their interactions and functional characteristics for understanding their relative contributions to the immune system of a bacterial pathogen. These differences and comparisons are critical for both understanding basic bacterial physiology and expanding the existing CRISPR toolset.

The reprogramming of native CRISPR-Cas systems in bacteria, primarily Cas9 dependent systems, is an evolving tool to modify the bacterial genome, combat antibiotic resistance, virulence, and manipulate the population balance in complex environments (131, 137-145). The diverse applications of reprogramming bacterial systems represent one of the new frontiers of CRISPR-Cas technologies, and Cas12a may provide an ideal alternative for these tools. Studies have shown that the reprogramming of native and heterologously expressed CRISPR-Cas systems can be used to make targeted mutations in the bacterial chromosome, conduct genetic screens, or alter transcription of a gene (131, 137, 143-151). Cas12a and Cas9 are found primarily in mammalian commensals and pathogens in both livestock and humans, and providing tools for making highly relevant microbes tractable for reverse genetic studies. Likewise, in both native and non-native bacterial hosts, Cas12a and Cas9 systems have additional applications in human health and agriculture as an antimicrobial and anti-virulence tool (152-154). The ability to transfer functional CRISPR-Cas systems between diverse bacteria and lethality of CRISPR-Cas targeting of the bacterial genome has led to the highly sensitive removal of specific bacteria from mixed cultures (142, 153). Similarly, CRISPR systems can be

used to selectively alter the fitness of specific organisms, for example through the conjugated targeting of pathogenic strains as well as emerging virulence factors and antibiotic resistance cassettes (25, 148). The reprogramming of multiple systems for new targets could be combined, native or supplemented into a prokaryote of interest, to effectively avoid resistance mutations or anti-CRISPR proteins that exist for one CRISPR-Cas system (155-157). Therefore, the use of multiple single-effector CRISPR systems within a host provides a more comprehensive toolset, much like how naturally in *F. novicida*, as demonstrated herein, the presence of both Cas12a and Cas9 provides a more comprehensive DNA defense program.

Materials and methods

Bacterial strains and growth conditions

All strains and plasmids can be found in Table S1. All strains of *Francisella novicida* U112 were grown at 37 °C, shaking, in Tryptic Soy Broth (VWR International Inc) supplemented with 0.2% Cysteine (BD Biosciences, Sparks, MD). All strains were plated on Tryptic Soy Agar (VWR International Inc) plates supplemented with 0.1 % Cysteine. *Escherichia coli* DH5-α were grown in Tryptic Soy Broth and plated on Tryptic Soy Agar. Kanamycin sulfate (30 μg/ml, Fisher Scientific Company) was used for selection for both liquid cultures and solid agar.

Mouse infections

Mice were infected subcutaneously with ~2 x 10^5 cfu bacteria. Spleens were harvested at 48 hours post infection, homogenized in PBS, and the bacterial burden per organ was determined. *in vivo* protocols are described in detail in the “STAR methods” section of Chapter 3.

PAM prediction

Francisella spp CRISPR-Cas12a spacers that target a putative prophage in *F. novicida U112 3523* were confirmed using BLASTn (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch). The protospacer was identified by the location of the hairpin secondary structure in the CRISPR repeat relative to the spacer for each strain, and the sequence flanking both sides of the protospacer was identified by BLASTn. The flanking sequences for all *Francisella spp* 3523 protospacers were compiled and the relative nucleotide contribution of each base at each position in these 5' and 3' sequences were determined using a sequence logo plot (link to weblogo), indicated by relative letter height.

Francisella novicida strain construction

Null deletion and point mutants listed in Table S1 were constructed by allelic exchange using the primers in Table S2. Homologous sequences (500-1000bp) up and downstream of the region of interest were amplified by PCR from genomic DNA of *F. novicida* U112, using Phusion high-fidelity DNA polymerase (New England Biolabs). These fragments were used to construct the allelic exchange substrate, a kanamycin selectable marker containing Flp recombinase target sites (FRT) was inserted between the flanking sequences using overlapping PCR, as previously described (158). Point mutants were constructed using the same technique and primers encoding the single amino acid substitution (Table S2). Linear fragments were transformed into chemically competent *F. novicida* and mutants were selected on kanamycin plates. DNA was isolated using Qiaquick Tissue Extraction Kit (QIAGEN, INC), and mutants were confirmed by PCR amplification from outside of the recombined region followed by sequencing (Genewiz). Strains were unmarked using a temperature sensitive suicide vector, pFFlp, carrying the Flp recombinase in *trans*, as described previously (159). Strains were complemented in *cis*.

Plasmid manipulations

Plasmids were transformed and isolated from competent E. coli (NEB 5-alpha) using Zyppy Mini, midi and Maxi prep kits (Zymo Research). The broad host range vector pBAV1K-T5-GFP (pBAV) was used as the control plasmid and backbone for the target plasmids in all assays. Overlapping PCR was used to construct Cas12a_target and Cas9_target plasmids, and the panel of PAM mutants in the Cas12a_target_5 plasmid (Table S1-2). In the target plasmids the prophage region and/or protospacer sequence inserted immediately downstream of the EcoR1 site of the pBAV backbone with the primers indicated in Table S1.

Plasmid Inhibition Transformation Assays

Overnight cultures of *F. novicida* were used to make competent cells by diluting 1:100 and growing to an OD_{600} of 1.0-1.2. To make competent cells, cultures were concentrated ten-fold in 4°C chemical transformation buffer (CTB) and incubated at 4°C for 15 minutes. Competent cells (300 µl) were incubated with plasmid DNA (500ng or 1000ng) for 25 minutes, shaking at 37°C before recovering in 1 ml of TSB+0.2% cysteine for 2.5 hours (shaking, 37°C). Transformations were plated on TSA plates with kanamycin selection and incubated at 37°C overnight. A no-DNA transformation was used as a negative control.

Quantitative real time-PCR

RNA was isolated over the course of the Plasmid Inhibition Transformation Assay (at OD_{600} 1.0, Competent Cells, Transformation, Recovery (45 minutes), Recovery (120 minutes)) using TRI-reagent and a Direct-zol RNA MiniPrep Kit (Zymo Research, Irvine, CA). Wild-type *F. novicida* was grown to late log phase (OD_{600} 1.0). The OD_{600} 1.0 *F. novicida* cultures were concentrated 10x in 4°C chemical transformation buffer (CTB) and stored at 4°C for 20 minutes (Competent

cells). For each transformation, 300µl of competent cells and shaken at 37°C for 25 minutes (Transformation). After 25 minutes, 1 ml of *F. novicida* media (TSB+0.2%cys) was added to each transformation and shaken at 37°C for 120 minutes (Recovery). DNA was removed from the samples using Turbo DNaseI (Ambion Biosciences). Power Sybr Green RNA-to-CT one-step kit (Applied Biosystems) was used to generate cDNA and conduct the quantitative real-time PCR using the primers indicated in Table S2 (qRT-PCR). Relative transcript levels were calculated by normalizing C_T values to the *Francisella novicida* housekeeping gene DNA helicase II (*uvrD* and *FTN_1594*) to determine $2^{-\Delta\Delta Ct}$ for each condition. Each experiment was normalized to the expression level in the starting culture (OD_{600} 1.0), resulting in the fold change in expression between conditions. Results were plotted as fold change in expression relative to the starting culture (OD_{600} 1.0 culture).

Statistical analysis

Unpaired, two-tailed, Student's t-tests were used to determine significance for all of the figures except for the in vivo infection in Figure 1, for which a Mann-Whitney test was used.

Acknowledgements.

We thank Dr. Siddharth Jaggavarapu for thoughtful comments on the manuscript and Drs. Tim Sampson and Chui Yoke Chin for her insight and assistance troubleshooting assays on this topic.

Funding information.

CRISPR work in the lab is supported by National Institutes of Health (NIH) grants U54-AI057157 from the Southeastern Regional Center of Excellence for Emerging Infections and Biodefense, as well as R01-AI110701 to David S. Weiss, who is also supported by a Burroughs Wellcome Fund Investigator in the Pathogenesis of Infectious Disease award.

Chapter 5. Extent of crRNA complementarity shifts CRISPR-Cas systems between transcriptional repression and DNA defense

Hannah K. Ratner[1,2,3] and David S. Weiss,[2,3,4,*]

[1]Microbiology and Molecular Genetics Program, Emory University, Atlanta, GA 30329.

[2]Emory Vaccine Center, Emory University, Atlanta, GA 30329.

[3]Yerkes National Primate Research Center, Emory University, Atlanta, GA 30329.

[4]Division of Infectious Diseases, Department of Medicine, Emory University School of Medicine, Atlanta, Georgia 30329.

*Corresponding author:

David S. Weiss

Emory Vaccine Center

954 Gatewood Rd, Room 2028

Atlanta, GA 30329

Tel: (404) 727-8214

Keywords: CRISPR-Cas, DNA targeting, transcriptional regulation, Cas9, Cas12a, *Francisella*

Abstract

CRISPR-Cas systems are widespread in prokaryotes and function as adaptive immune systems that use small RNAs (crRNAs) to guide Cas protein(s) to recognize and cleave harmful foreign nucleic acids. There is increasing evidence for broader endogenous roles of these systems, and a large percentage of the crRNA targets remain unidentified. The CRISPR-Cas9 system of pathogenic *Francisella novicida* (FnoCas9) regulates endogenous gene expression using Cas9, tracrRNA, and a small CRISPR-associated RNA, scaRNA, which is distinct from crRNAs. scaRNA leads to repression of transcription without cleavage through reduced complementarity to the DNA target and binding near the transcriptional start site. It was unclear if FnoCas9 crRNAs have a similar ability to repress transcription in the absence of DNA cleavage. We observed that, indeed, FnoCas9 can repress transcription using partial crRNA complementarity. Furthermore, this ability extended to the Cas12a system of *F. novicida* (FnoCas12a), indicating that it is not restricted to Cas9. We further found that when reprogrammed, FnoCas12a can regulate the transcription of physiologically important genes in a native bacterial host. These results highlight the bifunctional capacity of Cas9, Cas12a, and possibly other CRISPR-Cas systems for DNA cleavage or transcriptional repression as directed by the extent of complementarity of associating small RNAs and their targets. This paradigm of CRISPR-Cas functionality suggests that a broad re-analysis of crRNAs for endogenous targets with reduced complementarity could reveal new and diverse roles for CRISPR-Cas systems in prokaryotic biology.

Author Summary

The CRISPR-Cas9 system is best known as an adaptive immune system in prokaryotes that recognizes and cleaves foreign DNA. However, the breadth of CRISPR-Cas functions have not been completely elucidated. In the bacterial pathogen *Francisella*, Cas9 has an additional role in binding and repressing transcriptional read-through of endogenous genes, using a non-canonical small RNA (scaRNA). We have now tested whether crRNAs, which unlike scaRNA are ubiquitous in every CRISPR-Cas system, could similarly guide Cas9 to repress endogenous transcription. Indeed, we found that both native CRISPR-Cas systems of *F. novicida*, Cas9 and Cas12a, can transcriptionally repress genes without cleaving DNA. Furthermore, the difference in functionality from DNA cleavage to repression is directed by the reduced extent of crRNA complementarity to its target. Since myriad crRNAs have not previously been explored for limited complementarity to endogenous targets, many instances of endogenous transcriptional regulation may not yet have been uncovered. This work therefore suggests that reanalysis of crRNAs for their potential to regulate endogenous targets could lead to a paradigm shift in our understanding of the roles of these systems in prokaryotic biology.

Introduction

CRISPR (clustered, regularly interspaced, short, palindromic repeats) - Cas (CRISPR-associated) systems are widespread in prokaryotes, where their best known role is as adaptive immune systems that protect against foreign genetic elements (13, 22). While these systems are diverse in composition and mechanism, the CRISPR-Cas9 and CRISPR-Cas12a (Cpf1) systems in particular have been widely repurposed for molecular biological and therapeutic genome engineering applications (37, 160-162). These two systems use single

effector proteins (Cas9 or Cas12a) to cleave foreign nucleic acid targets by forming a complex with a small RNA, CRISPR RNA (crRNA) (15, 135, 163). Each crRNA contains a repeat which is bound by the effector protein and a spacer that binds to complementary sequences on foreign nucleic acid targets, guiding Cas9 or Cas12a to cleave the target (called the protospacer) and protect the cell (15, 135, 163). While the Cas12a crRNA independently directs the protein to its target, Cas9 interacts with the crRNA as an RNA duplex with tracrRNA, a second small RNA transcribed from the CRISPR-Cas9 locus (82).

The pathogen *Francisella novicida* encodes CRISPR-Cas9 (FnoCas9) and -Cas12a (FnoCas12a) systems, both of which are capable of directing DNA cleavage (88, 119, 135). In addition to crRNA-directed cleavage, FnoCas9 also represses the expression of endogenous genes, promoting virulence by facilitating evasion of the host immune system (164, 165). Cas9 has been shown in other pathogens to regulate traits that are important for virulence, such as attachment to host cells and intracellular survival (118, 166-168). We recently interrogated the mechanism for endogenous gene repression by FnoCas9 and developed an updated model. We found that FnoCas9 can act as a transcriptional repressor, using tracrRNA and an additional CRISPR-associated RNA, scaRNA, to guide FnoCas9 to bind but not cleave endogenous DNA targets located near transcriptional start sites (TSSs). Limited scaRNA:DNA target complementarity facilitated the cleavage-independent DNA binding of FnoCas9, leading to transcriptional repression (165). Interestingly, computational predictions of scaRNAs in organisms with type II CRISPR-Cas systems suggest a correlation between the presence of these small RNAs and increased strain virulence (168).

Given the ability of both scaRNA and crRNA to direct Cas9 to DNA targets, and the ability of scaRNA to direct transcriptional repression, we hypothesized that crRNAs of the *F.*

novicida CRISPR-Cas systems may also be capable of regulating transcription through a similar mechanism. We observed that small modifications to crRNA spacers that reduced target complementarity were sufficient for transitioning the activity of FnoCas9 from cleavage to transcriptional repression. We subsequently found that FnoCas12a displayed similar bi-functionality and repressed transcription upon reduced crRNA complementarity. Based on this finding, we engineered the FnoCas12a CRISPR array, successfully harnessing the transcriptional regulatory activity of this protein to repress endogenous gene expression. The ability of partially complementary crRNAs to regulate endogenous targets may represent a widespread function of the CRISPR-Cas9 and CRISPR-Cas12a machineries, and may be a general activity of other CRISPR-Cas systems as well. Thus, in addition to demonstrating that the bi-functionality of CRISPR-Cas systems can be rationally engineered by modifications to crRNAs, these data suggest that a broad re-analysis of known crRNAs for partial complementarity to endogenous targets could reveal new regulatory functions of CRISPR-Cas systems in prokaryotic biology.

Results

FnoCas9 crRNAs with reduced target complementarity can repress transcription

We previously showed that scaRNA mediates FnoCas9 binding to the *F. novicida* chromosome and repression of transcription, using 15 bases of complementarity between its tail and the target DNA (165). When scaRNA complementarity to the chromosome was artificially extended to 20 bases, lethal cleavage of the chromosome or target plasmid occurred instead (165). It was unclear if crRNAs could act similarly to scaRNA and direct these distinct functions. To determine whether Cas9 crRNAs with reduced complementarity to a DNA target could repress transcription without inducing cleavage, we developed a panel of FnoCas9 crRNA target

plasmids. Each plasmid contained a construct in which a synthetic constitutive promoter (p146) drives the expression of *gfp*, with 0, 8, 11, or 15 bases of complementarity to an FnoCas9 crRNA (the first crRNA spacer in the CRISPR array was used) (169). In addition, an FnoCas9 protospacer adjacent motif (PAM), a short nGG sequence adjacent to the DNA target, was included between the transcriptional start site and *gfp* as this is required for FnoCas9 binding (Figure 1A-B) (165). Each plasmid was transformed into WT and Δ*cas9* strains, and *gfp* expression was quantified. The full spacer plasmid was not included as it has been shown previously to be targeted by Cas9 for cleavage (129). For plasmids containing 11 and 15 bases of complementarity to crRNA, *gfp* expression was lower in WT as compared to *Δcas9*, indicating that FnoCas9 significantly repressed expression (Figure 1C). In contrast, FnoCas9 did not similarly repress *gfp* from plasmids with 0 or 8 bases of complementarity (Figure 1C). These data are consistent with the level of complementarity required for scaRNA-mediated *gfp* repression from a plasmid (8 bases is insufficient while 11 bases is sufficient)(165). Together, the data indicate that an *F. novicida* crRNA can direct FnoCas9 to repress transcription from a target plasmid.

To determine whether the observed transcriptional repression occurred in the absence of DNA cleavage, and to confirm that full complementarity between the crRNA and DNA target leads to cleavage, we measured the ability of FnoCas9 to inhibit transformation with each of the plasmids as well as a plasmid with the full protospacer target which includes the 20 bp of complementary to the FnoCas9 crRNA that is required for cleavage (Figure 1A-B). Plasmids with 0, 8, 11, or 15 bases of complementarity to the crRNA were transformed at the same efficiency into WT *F. novicida* and Δ*cas9*, indicating the lack of plasmid restriction (Figure 1D). Conversely, transformation with the plasmid harboring the full (34 bp) sequence

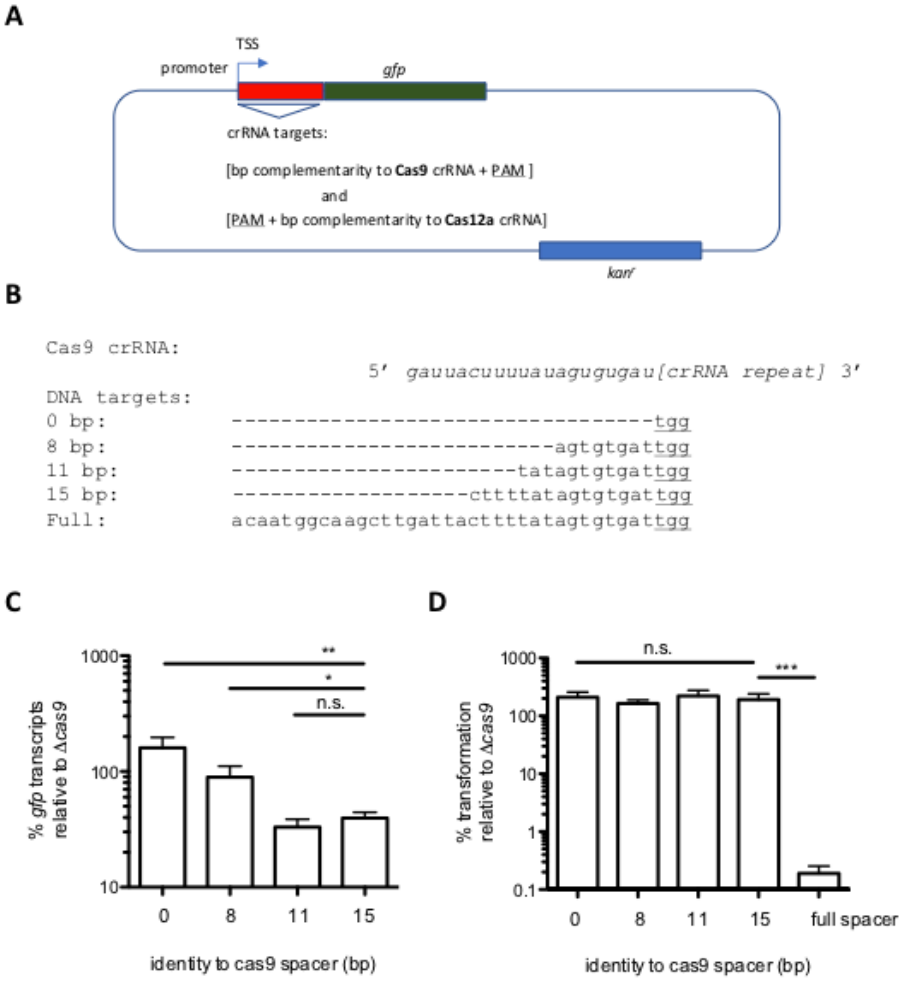


Figure 1. Reduced crRNA:target complementarity shifts Cas9 function from DNA cleavage to transcriptional repression. A) Diagram of the pBav-derived vectors used in B-D. B) Sequence alignment of the Cas9 protospacer target and PAM (NGG; underlined) inserted into the vector (A) with varying complementarity to the Cas9 crRNA, used in C-D. C) Relative *gfp* transcript levels were measured by qRT-PCR from constructs containing targets with varying lengths of complementarity to the Cas9 crRNA (0, 8, 11, 15 bp). Results are presented as % *gfp* expression from each plasmid in WT relative to that in Δ*cas9* (n=6, error bars represent s.e., *p≤0.05; **p≤0.005). D) Transformation of WT *F. novicida* and Δ*cas9* with the plasmids from (C) with 0 bp, 8 bp, 11 bp, 15 bp, and full (34 bp) complementarity to the Cas9 crRNA. Results are presented as percent transformation into WT relative to Δ*cas9*. (n=6, error bars represent s.e., ***p≤0.001).

complementary to the FnoCas9 crRNA spacer resulted in significantly reduced levels of transformation in WT compared to Δ*cas9* (Figure 1D). These results indicate that similar to scaRNA, the ability of FnoCas9 to repress transcription or cleave DNA is determined by the extent of complementarity between the target and the guiding crRNA.

Cas12a crRNAs can direct different functions based on target complementarity

We next sought to investigate whether this RNA-determined bifunctionality of FnoCas9 could be extrapolated to the other native CRISPR-Cas system of *F. novicida*, CRISPR-Cas12a. To test the ability of crRNAs to direct Cas12a to repress transcription, we again used a panel of plasmids containing a construct with the p146 promoter driving the expression of *gfp*, with an FnoCas12a PAM (TTTn) and 0, 8, 11, or 15 bases of complementarity to an FnoCas12a crRNA located between the transcriptional start site and *gfp* (Figure 1A and 2A). Following transformation of each plasmid into WT and Δ*cas12a* strains, the relative *gfp* expression from the plasmids was measured. *gfp* expression from the plasmid with 15 bases of complementarity to the crRNA was significantly lower in WT compared to Δ*cas12a*, indicating Cas12a-dependent repression (Figure 2B). In contrast, no repression of *gfp* was observed from plasmids with 0, 8, or 11 bases of complementarity to the FnoCas12a crRNA in either WT or Δ*cas12a* (Figure 2B). This highlights a difference with FnoCas9, where 11 bases of complementarity was sufficient to repress *gfp* expression from the plasmid construct (Figure 1B). These results indicate that FnoCas12a is capable of directing transcriptional repression using crRNAs.

We measured the ability of FnoCas12a to restrict transformation of plasmids with 0, 8, 11, 15, or 29 bp (full) complementarity to the fifth spacer in the FnoCas12a CRISPR array (Figure 1A and 2A). FnoCas12a did not inhibit transformation of plasmids containing 0, 8, 11, or

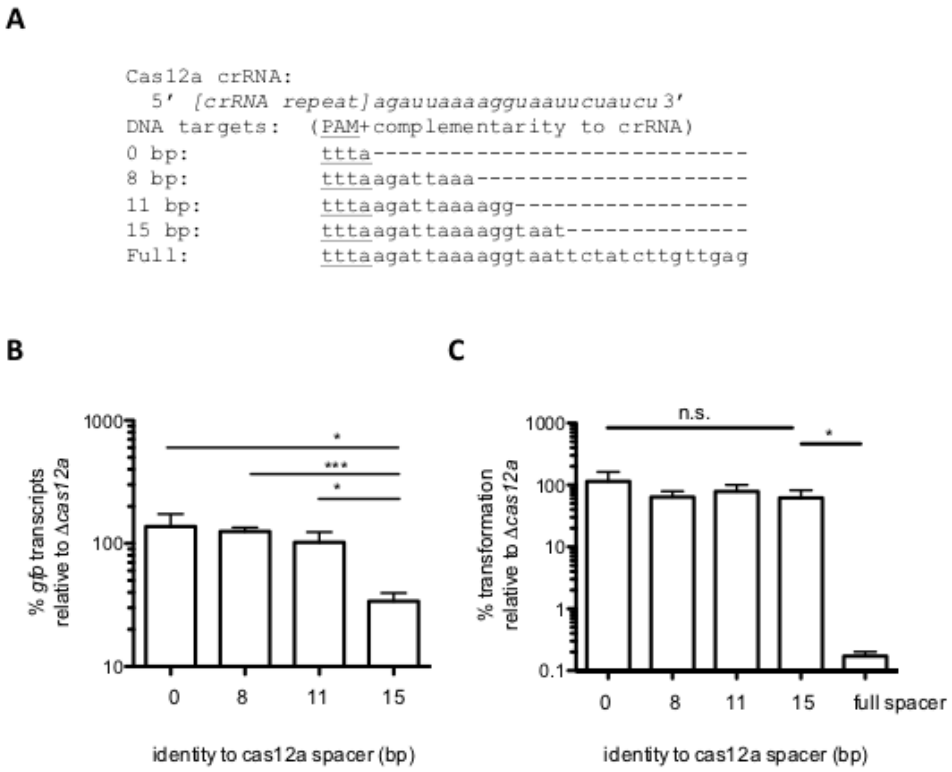


Figure 2. crRNA complementarity to target DNA shifts Cas12a from transcriptional repression to DNA cleavage. A) sequence alignment of the PAM + Cas12a protospacer sequences with varying complementarity to the Cas12a CRISPR array spacer, used in B-C. The PAM sequence (TTTA) is underlined. B) Relative *gfp* transcript levels were measured by qRT-PCR from constructs containing sequences with different lengths of complementarity to the Cas12a spacer (0, 8, 11, 15 bp) between a synthetic promoter and *gfp*. Results are presented as % *gfp* expression from each plasmid in WT relative to that in Δ*cas12a* (n=6, error bars represent s.e., *p≤0.05; ***p≤0.001). C) Transformation of WT *F. novicida* and Δ*cas12a* with plasmids with 0 bp, 8 bp, 11 bp, 15 bp, and full complementarity to the Cas12a spacer between the promoter and *gfp*. Results are presented as percent transformation into WT relative to Δ*cas12a*. (n=6, error bars represent s.e., *p≤0.05).

15 bases of complementarity to the crRNA (Figure 2C). However, transformation efficiency with the construct containing the full protospacer was strongly inhibited in WT relative to Δ*cas12a* (Figure 2C). These results indicate that catalytically active FnoCas12a can both repress transcription and direct target cleavage, and that this functional differentiation is determined by the extent of complementarity between the target and the crRNA.

Cas12a can be reprogrammed to repress endogenous Cas9 targets

FnoCas9 uses scaRNA to repress 4 genes encoded on 2 mRNA transcripts (*1104-1102* and *1101*)(Figure 3A) (165). We attempted to further validate the functionality of Cas12a-mediated transcriptional repression by reprogramming FnoCas12a to repress these genes. We reprogrammed the FnoCas12a CRISPR array to have partial complementarity to the 5' UTR of the *1104-1102* open reading frames (ORF), which are transcribed on a single mRNA, to determine whether the reprogrammed crRNA could repress endogenous targets (Figure 3B). To investigate the specificity of FnoCas12a-mediated targeting, we generated the construct such that the crRNA would not also bind the *1101* 5' UTR, unlike scaRNA which naturally targets both regions (Figure 3B). First, we measured the levels of each transcript in the native FnoCas9 regulon in WT, Δ*cas12a*, and Δ*cas9* strains, observing that the levels of *1104, 1103, 1102* and *1101* were not altered in the Δ*cas12a* mutant compared to WT, but were expressed at higher levels in the Δ*cas9* strain (Figure 3C-F). This confirmed the requirement of FnoCas9 in the repression of these genes and further demonstrated that FnoCas12a is not naturally involved (Figure 3C-F). We then measured the levels of the genes in the FnoCas9 regulon in strains with the reprogrammed FnoCas12a CRISPR array in the absence of FnoCas9 and in a Δ*cas9*/Δ*cas12a*

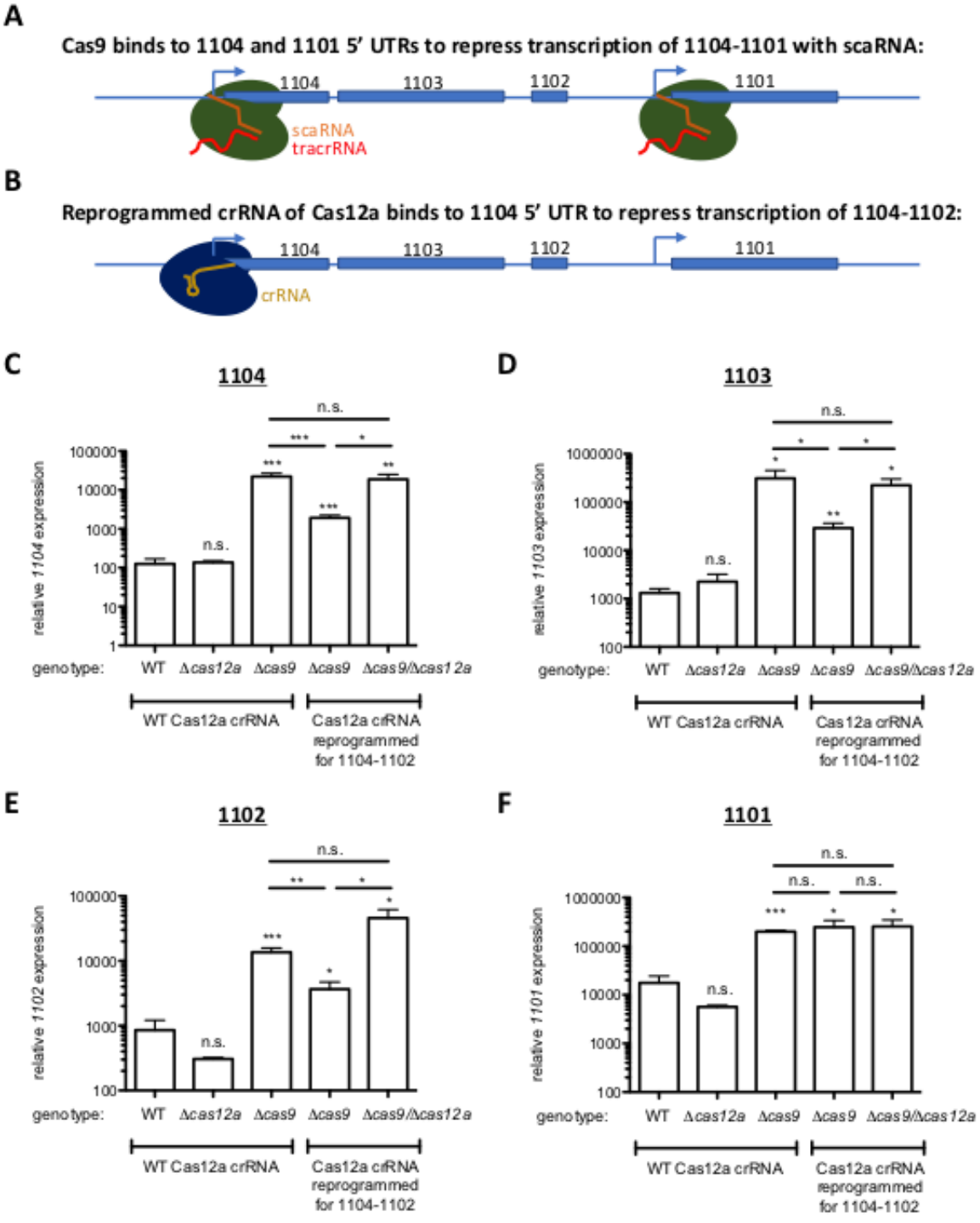

A
Cas9 binds to 1104 and 1101 5' UTRs to repress transcription of 1104-1101 with scaRNA:
1104
1103
1102
1101
scaRNA
tracrRNA
B
Reprogrammed crRNA of Cas12a binds to 1104 5' UTR to repress transcription of 1104-1102:
1104
1103
1102
1101
crRNA
C
1104
relative 1104 expression
genotype:
WT
Δcas12a
Δcas9
Δcas9
Δcas9/Δcas12a
WT Cas12a crRNA
Cas12a crRNA reprogrammed for 1104-1102
D
1103
relative 1103 expression
E
1102
relative 1102 expression
F
1101
relative 1101 expression
n.s.

Figure 3. Cas12a can be reprogrammed to repress endogenous targets. A) Model of Cas9 in complex with tracrRNA and scaRNA binding to the *1104* and *1101* 5' UTRs on the *F. novicida* chromosome with the scaRNA tail to repress *1104-1101* transcription. B) Model of Cas12a interacting with the 1104 5' UTR with the reprogrammed crRNA to repress *1104-1102* transcription. C-F) Relative levels of (C) *1104*, (D) *1103*, (E) *1102*, and (F) *1101* transcripts were measured by qRT-PCR in WT, Δ*cas9,* Δ*cas12a,* Δ*cas9* + Cas12a crRNA reprogrammed, and Δ*cas9*Δ*cas12a* + Cas12a crRNA reprogrammed strains. (n=6, error bars represent s.e., *$p \leq 0.05$; **$p \leq 0.005$; ***$p \leq 0.001$).

double mutant as a control for baseline expression. Compared to a Δ*cas9* strain that is unable to repress these genes naturally, the Δ*cas9* strain with the reprogrammed Cas12a crRNA had reduced levels of the *1104-1102* transcript (Figure 3C-E). This suggested that Cas12a was actively repressing these genes in the presence of the reprogrammed crRNA. To determine if this transcriptional repression in the reprogrammed strain was Cas12a-dependent, we compared the expression levels of *1104-1102* in the Cas12a reprogrammed strain (Δ*cas9*, intact Cas12a) and a reprogrammed strain without Cas9 and Cas12a (Δ*cas9*/Δ*cas12a*). We found that *1104-1102* expression was restored to the level of the Δ*cas9* strain in the absence of Cas12a, demonstrating that Cas12a can be reprogrammed to repress transcription through modification of the CRISPR array (Figure 3C-E). As an additional control, *1101* expression was unaltered between the Δ*cas9* and Δ*cas12a*/Δ*cas9* reprogrammed strains, consistent with the specific effect of FnoCas12a on *1104-1102* (Figure 3F). Taken together, these results show that Cas12a can repress endogenous transcription and highlight the functional flexibility of the *F. novicida* CRISPR-Cas systems.

Discussion

Based on our recent discovery that scaRNA uses limited target complementarity to guide FnoCas9 to repress transcription of endogenous genes (165), we explored whether crRNA:target complementarity could be modulated to facilitate transcriptional repression by the two native CRISPR-Cas systems of *F. novicida*. Both the FnoCas9 and FnoCas12a systems could efficiently repress transcription without cleaving the DNA when the target sequence contained 15 bases of complementarity to the crRNA (Figure 1C, 2B), but cleaved when there was full complementarity of the crRNA spacer to the target (Figure 1D, 2C). We harnessed this functionality to, for the first time, reprogram cleavage-capable Cas12a to repress transcription (Figure 3A-F).

The work presented here is supported by *in vitro*, structural, and heterologous expression studies that demonstrate that Cas9 crRNAs with partial target complementarity prevent cleavage of DNA (135, 141, 170). Coordinated double strand cleavage by Cas9 involves a conformational shift into a cleavage-activated state that depends on sufficient base pairing between the crRNA and DNA target (170). However, binding can still occur when Cas9 has insufficient crRNA:target complementarity to direct cleavage. We observed that when reduced complementarity between a crRNA and DNA target resulted in a loss of cleavage by Cas9, the effector gained the ability to repress transcription, and this crRNA-directed shift between DNA restriction and transcriptional repression could be extrapolated to a different CRISPR-Cas system, Cas12a. These data are supported by a recent study demonstrating the ability of an engineered Cas12a to exhibit crRNA-directed transcriptional activation in eukaryotic cells (171). Particularly interesting is that Cas12a, unlike Cas9, is able to process its own crRNAs without the assistance of cellular RNases, indicating that the system is both highly independent and functionally versatile, able to process crRNAs that could likely guide multiple distinct effector functions (128, 171, 172). Our data highlight the unexplored ability of cleavage-capable Cas12a systems to act as regulators of transcription, which may be an important natural role of these systems (Figure 2). Furthermore, the similar ability of both Cas9 and Cas12a to interfere with transcriptional repression and restrict foreign DNA suggests that crRNA-directed functional flexibility could be a feature that is conserved in other systems as well.

Other CRISPR-Cas systems, including those that use multi-protein effector complexes, may utilize crRNAs with partial complementarity for endogenous regulatory functions. A systematic analysis of Type I-E spacer sequences in an *E. coli* strain found that self-targeting spacers were more likely to target the host than foreign nucleic acids (173). Most interestingly,

the bioinformatic predictions of the crRNA binding sites suggest that they regulate endogenous transcription by binding in a PAM-dependent manner to transcriptionally active regions of the genome with partial complementarity(173), similar to what we describe here.

Furthermore, there exist examples of crRNAs influencing bacterial physiology. In the type I-E CRISPR-Cas system of *Pseudomonas aeruginosa*, a crRNA directs degradation of the mRNA for *lasR*, a transcriptional regulator which helps the pathogen evade the mammalian immune system (174). The type I-F CRISPR-Cas system of *P. aeruginosa* is involved in regulating motility through partial complementarity of a crRNA to a prophage (175). Additionally, some orthologs of Cas9 have been demonstrated to guide ssRNA cleavage through a ~20bp crRNA:target RNA interaction, although the biological importance of this mechanism is not fully understood (83-85).

Interestingly, self-targeting spacers were identified early in the study of CRISPR-Cas systems. While a recent analysis investigated predicted DNA targets of crRNAs with extensive complementarity (95% coverage and identity between spacer and predicted target), potential targets based on reduced complementarity were not studied (81). It remains unknown how many self-targeting spacers may exist with reduced complementarity to endogenous targets and could result in transcriptional regulation. Furthermore, the targets of the majority of spacers are not known (81). This could be due to the sequences of their exogenous targets being absent from current databases, but would also be consistent with them having endogenous regulatory targets of limited complementarity. Importantly, as we have shown, regulatory functions of CRISPR-Cas systems can co-exist with the canonical adaptive immune functions for multiple CRISPR-Cas systems. We hypothesize that as self-targeting spacers are investigated further, many more roles for CRISPR-Cas systems in diverse aspects of endogenous physiology will be revealed.

Materials and methods.

Strains and growth conditions

Francisella novicida strains were grown at 37 °C shaking. Tryptic Soy Broth (VWR international) with 0.2 % cysteine (BD biosciences) supplementation was used as the growth media for liquid cultures. For colony-based assays strains were grown on Typtic Soy Agar (VWR international) plates with 0.1% cysteine at 37 °C. For selection of mutants and maintanence of plasmids, media was supplemented with kanamycin sulfate (30 μg/ml, Fisher Scientific). For removal of the kanamycin resistance cassette, the system utilized tetracycline selection, in which media was supplemented with 15 μg/ml tetracycline (Alfa Aesar) and the strains were grown at 30 °C during the unmarking.

***Francisella novicida* mutant generation**

All strains used in experiments were derived from *Francisella novicida* U112 ("WT") (Table S1). Δ*cas12a and* Δ*cas9/*Δ*cas12a* deletion mutants and the reprogrammed crRNA array were made using allelic exchange with the primers listed in Table S3. The reprogrammed crRNA was made by replacing the Cas12a CRISPR array with a repeat-spacer-repeat in the same location, with the spacer targeting the 1104 5' UTR. Linear allelic exchange products were used as the substrates for the mutations, the 500-1000bp regions of homology for recombination into specific sites of the genome flanking the new inserted sequence (where necessary) and a selection cassette. For mutant selection a kanamycin selectable marker was used, containing Flp recombinase target sites (FRT) between the flanking sequences to enable unmarking (158). Fragments for construction of the allelic exchange products were amplified from genomic DNA (isolated with Qiaquick Tissue Extraction Kit) and sewn together using overlapping PCR as

described previously (165). Phusion high-fidelity DNA polymerase (New England Biolabs) was used for PCR amplification and the Qiaquick gel extraction kit (Qiagen) was used to purify DNA from PCR reactions. Chemically competent *F. novicida* was transformed with the allelic exchange substrates (158) and recombinants were selected for on kanamycin and confirmed by amplification and sequencing (Genewiz) using "ampli" and "seq" primers (Table S3) (165). The selection cassette was removed a vector containing the Flp recombinase in *trans*, using temperature and tetracycline selection as described previously (Table S2) (159). The Δ*cas9* strain was described in previous work (Table S1) (114) (117, 118).

Plasmid construction

Plasmids (Table S2) were constructed in the broad host range vector pBAV1K-T5-GFP (pBAV) using the primers listed in Table S3(176). In the plasmids used in the transformation and Cas9-/Cas12a-dependent transcriptional repression assays, the promoter and RBS that drive *gfp* expression in WT pBAV were replaced with a synthetic constitutive promoter "p146" followed by varying amounts of homology to the Cas9 or Cas12a crRNAs and respective PAMs, in front of *gfp* (169) (165). As per the functioning of each system, the Cas9 PAM was placed directly after the target sequence (-NGG) while the Cas12a PAM was inserted before the target sequence directly after the TSS (-TTTN) (Figure 1A, 2A). Plasmids were constructed, amplified and isolated from competent *E. coli* (NEB 5-alpha) as described previously (165).

Transformation Assays

Chemically competent *F. novicida* was made by 10C concentration of log-phase cultures ($OD_{600\ nm}$ of 0.8-1.0) in 4°C chemical transformation buffer (CTB) as described previously (158) (165).

As described in Ratner et al, 2019, competent cells were mixed with DNA and incubated at 37°C shaking (20 min), followed by a 2 hour recovery in 1 ml of liquid growth media at 37°C shaking and plating on selective agar plates overnight (37°C) (165). For transformations with plasmid vectors, the same amount of plasmid was transformed into WT and either the Δ*cas9* and Δ*cas12a* strains (~500 ng per plasmid). Transformants were enumerated for each strain and compared between WT and either Δ*cas9* or Δ*cas12a* strains to determine the % transformation relative to the CRISPR-effector protein deficient strain (165). As described in (165), Cas9 and Cas12a repression of *gfp* expression was measured by isolating transformants from each strain and for each plasmid, for growth in liquid growth media with kanamycin supplementation. At $OD_{600\ nm}$ of 0.8-1.0, RNA was isolated from the transformants and used to measure *gfp* transcript level in each strain (see qRT-PCR methods).

Quantitative real-time PCR

RNA extraction and qRT-PCR was conducted from cultures at OD_{600} nm of 0.8-1.0 as described in (165). RNA was extracted using TRI-reagent and a Direct-zol RNA MiniPrep Kit (Zymo Research) followed by Turbo DNaseI (Ambion Biosciences) (165). Power Sybr Green RNA-to-CT one-step kit (Applied Biosystems) was used for measurement of transcript levels by qRT-PCR (Primers in Table S3, (165). To determine 2-ΔΔct, *CT* values for *1104-1101* were normalized to DNA helicase II (*uvrD*, *FTN_1594*) and plasmid *gfp* expression was normalized to the kanamycin resistance cassette (118) (165). For the plasmid assays, the results are presented as % transcript in WT relative to the uninterrupted expression in the respective CRISPR-effector deficient strain (Δ*cas9* or Δ*cas12a* strains).

Statistical analysis

Unpaired, two-tailed, Student's t-tests were used to determine significance.

Acknowledgements.

We thank the members of the Weiss lab for their feedback and support on this project. This work was supported by National Institutes of Health (NIH) grant R01-AI110701 to David S. Weiss, who is also supported by a Burroughs Wellcome Fund Investigator in the Pathogenesis of Infectious Disease award.

Supplemental Information. Tables S1 and S2 are in Appendix G.

Chapter 6. Discussion and Relevance

BLP repression in *F. novicida* virulence

Cas9 protects against combinatorial killing by host antimicrobials

We demonstrated that in the intracellular pathogen *F. novicida*, the repression of 4 predicted BLPs (1104, 1103, 1102 and 1101) by Cas9 is required for virulence. It was previously demonstrated that 1103 helps *F. novicida* evade the Tlr-2 and the AIM2/ASC inflammasome innate immune defenses (115). This work suggested that repression of 1103 both decreases activation of Tlr-2 by reducing the abundance of its ligand and fortifies the bacterial membrane by decreasing permeability (115). Studies of 1103 have shown that deletion of this BLP from a *cas9* mutant strain can partially, but not fully, recover virulence (115, 118). We found that the virulence of a *cas9* mutant is fully restored upon the deletion of all four regulated genes, and that repression of each BLP contributes to infectivity in a murine model (177). Further elucidation of the connection between the repression of *1104*, *1102*, and *1101*, immune evasion, and changes to the bacterial envelope will provide insight into how envelope architecture contributes to the fitness of intracellular pathogens.

We have begun to investigate the role of *1104-1101* in *F. novicida* virulence, hypothesizing that Cas9, by repressing proteins near the host-pathogen interface, may promote intracellular survival by conferring resistance to antimicrobials produced in the phagosome to enable infection. We investigated the role of Cas9 in resisting reactive oxygen species (ROS) and lysozyme, which are important host-antimicrobials involved in killing intracellular pathogens. Interestingly, bacteria lacking Cas9 were more susceptible to combinatorial killing by H_2O_2 and lysozyme *in vitro* (Appendix A). Virulence of a Cas9 mutant, which is attenuated in WT mice, was increased in mice that were unable to produce both antimicrobials (Appendix B). Based on

the observation that lysozyme alone was unable to kill *F. novicida*, but when combined with a sub-MIC concentration of H_2O_2 *in vitro* resulted in a combinatorial killing effect (Appendix A), we interrogated the specific role of ROS in the Cas9-mediated virulence of *F. novicida*.

We found that Cas9 confers resistance to H_2O_2 *in vitro* through BLP repression (Appendix C). Interestingly, ROS-production by the host does not alter BLP-mediated activation of Tlr-2, suggesting that the role of BLP repression in evading killing by ROS is independent of this pathway (Appendix D). In further support of this finding, we found that Cas9-conferred resistance to H_2O_2 killing is physiologically important; in WT mice a *cas9* mutant is fully attenuated, while in a mouse that is unable to produce ROS as an innate immune defense, the *cas9* mutant is virulent (Appendix E). These data reveal an unexpected and previously uncharacterized role for Cas9 in promoting resistance to oxidative stress-dependent killing, and facilitating virulence.

Connection between the bacterial envelope and sensitivity to ROS

Next, it will be important to determine the mechanism of increased ROS sensitivity when BLPs are expressed. In some Gram-negative bacteria, it has been shown that regulation of membrane permeability is important for reducing susceptibility to hydrogen peroxide, a reactive oxygen species (ROS) antimicrobial that is produced by the host (178). The previous implication of 1103 in membrane integrity points to a possible contribution of 1104, 1102, and 1101 as well. Furthermore, lysozyme alone has been shown to potentiate H_2O_2 generation at physiologically relevant levels in eukaryotic cells (179, 180). The link between these two antimicrobials during infection is likely important for pathogens that passage through the phagosome and has been minimally explored in this context. It will be important to examine the contribution of membrane permeability to H_2O_2 sensitivity, how the oxidation of membrane BLPs alters the fitness of *F.*

novicida and interaction with other host-produced antimicrobials, and the direct interactions between ROS and lysozyme that may contribute to their synergy.

Future work should aim to identify the native function of these four genes in *Francisella*, the physiological changes occur at the membrane in a *cas9* mutant when they are highly expressed, and the mechanism and extent to which they alter fitness during passage through eukaryotic cells. It will be important for this investigation to determine the precise localization, membrane orientation, and interaction partners of each of the gene products of Cas9-regulated genes. Collectively, this will provide insight into how features of the bacterial envelope, in particular bacterial lipoproteins, membrane integrity, and interactions between ROS and antimicrobial peptides, contribute to the fitness and lifecycle of intracellular pathogens.

Implications of the ecological distribution of CRISPR-Cas9 systems for human health

Continued study of how CRISPR-Cas systems alter interactions between bacteria and their surroundings, such as host-pathogen interactions, will provide new insight into the alternative functions of these widespread systems. Our interrogation of the non-DNA targeting functions of the *F. novicida* CRISPR-Cas9 system may help to explain the unique ecological distribution of Class II CRISPR-Cas systems, which all use single proteins to recognize, bind, and cleave their targets (36). The prevalence of Cas9 is enriched in host-associated organisms such as pathogens and commensals. The reason for this clustering is not known, but does not appear to be linked by habitat or phylogeny. Conversely, Class I CRISPR system are found throughout the domain of prokaryotes and use a distinct multiprotein mechanism recognize and cleave their targets (36). One hypothesis is that this distribution is the result of the ability of systems with single effector proteins to adapt to environmental changes and mediate additional

functions in bacterial physiology beyond foreign DNA defense. More broadly, these findings highlight the need for a systematic bioinformatic and, when possible, experimental exploration of the functions of these systems in prokaryotic physiology beyond adaptive immunity.

Case to Broaden the Search for New CRISPR functions.

Considerations for the identification of RNAs that direct non-cleavage CRISPR-Cas9 functions.

We have described how the *F. novicida* CRISPR-Cas system regulates endogenous genes to facilitate virulence. Cas9 has been shown to be regulate traits that are important for virulence in other pathogens as well, such as attachment to host cells and intracellular survival, although the mechanism of underlying these functions have yet to be determined (118, 166, 167). Interestingly, computational predictions of scaRNAs in organisms containing Type II CRISPR systems suggest that scaRNAs may be more common in strains associated with increased virulence (168). To determine whether Cas9 is a potential endogenous regulatory factor via a small RNA such as scaRNA in other Cas9-harboring organisms, specific parameters for these analyses, based on the mechanism of scaRNA activity, should be taken into account.

In our work, the homology between scaRNA and the endogenous DNA targets was undetectable by searches for sequence similarities due to scaRNA's low number of consecutive bases of homology to the genome (11-12 bp). Identifying the location of scaRNA binding first required identification of the targeted locus and detailed genetic interrogation of the coding, non-coding, and promoter regions of the targeted locus to narrow in on the 5' UTR targets. From there, the location and target of scaRNA (DNA vs RNA) was identified and validated efficiently as the result a single putative PAM in each targeted UTR.

The most significant limitation to correctly identifying Cas9-associated small RNA targets remains the low sequence complementarity required for transcriptional repression. However, now that the Cas9 scaRNA mechanism has been elucidated, it may be possible to design bioinformatics tools, such as those being developed by the Djordjevic group, to identify these RNAs (168, 173). The search for the genomic target of the "spacer" or tail sequence of small RNAs containing CRISPR-like repeats, identified by small RNAseq, can be narrowed down by excluding targets without a PAM and requiring proximity to a promoter or TSS and complete complementarity within the seed sequence. The inclusion of the PAM sequence in the searches extends the minimum target sequence in the genome, reducing false targets. Therefore, when possible, it is necessary to experimentally interrogate the PAMs and PAM flexibility of culturable and genetically tractable organisms of interest. This is evidenced by our results indicating that Cas12a in *F. novicida* has a high level of flexibility in tolerated PAM sequences, although they do vary in cleavage efficiency, and this flexibility could not have been picked up by bioinformatics PAM prediction. In the identification of scaRNAs, applying an understanding of the level of repeat degeneration that is tolerated whilst retaining DNA binding function is also important, when possible. Analyses should extend beyond the host bacterium to the mobilome and potential interaction partners, prokaryotic and otherwise, because this may shed light on another new class of CRISPR functions, regulating the transcription of predating nucleic acids or interacting species. More broadly, it remains important to continue interrogating other functions and mechanistic flexibility of these systems, beyond this work, as evidenced by the recent findings that certain Cas9 orthologs can also degrade RNA (83-85). CRISPR-Cas systems may execute non-cleavage functions through a wide variety of mechanisms, some of which have yet to be discovered.

crRNA-directed Cas9 and Cas12a functional versatility

We demonstrated that through a mechanism similar to scaRNA, *F. novicida* Cas9 and Cas12a crRNAs are capable of repressing transcription without cleaving DNA (Chapter 5). This functional shift is determined by crRNA complementarity. When reprogrammed, FnoCas12a crRNAs can repress the native targets of scaRNA in the absence of Cas9. Importantly, the extension of this RNA-directed functional flexibility to the Cas12a system of *F. novicida* (FnoCas12a) suggests that it is a broader capability CRISPR-Cas systems that may include additional systems as well. The ability of Cas9 and Cas12a crRNAs to regulate gene expression while maintaining protective CRISPR systems indicates that there is not a fitness cost of having cleavage-capable CRISPR systems with additional beneficial functions in gene regulation, further suggesting that DNA defense is just one of many functions of CRISPR-Cas systems in prokaryotes. It is now pertinent to identify all of the systems that are capable of this shift. One hypothesis for the clustering of some systems, such as Cas9 and Cas12a, in host-associated organisms, is that the ability of certain systems to bind DNA or cleave DNA using the same effector (or possibly effector complex) results in a functional repertoire that is particularly beneficial for organisms with specific lifestyles or niche ecology.

The ability of crRNA and other small RNAs to direct diverse CRISPR-Cas activities likely encompasses a paradigm of CRISPR functions. As demonstrated herein, multiple Cas effector proteins can act as efficient, sequence specific transcriptional regulators by binding to the DNA, which could fulfill any endogenous regulatory role. They could also be altering transcription from mobile genetic elements to broaden the nature of their protective functions. DNA binding could have other functions within the bacterium as well, including but not restricted to occlusion of restriction modification sites and sequestration of the protein to the

DNA. Based on associations between CRISPR-Cas non-cleavage phenotypes and changes that occur at the membrane or in response to environmental changes, it is easy to theorize that these systems could more sensitively fine-tune interspecies and inter-domain interactions, proactively or defensively, by repressing transcription rather than cleaving DNA (33).

A need to re-analyze spacer targets

Fascinatingly, non-cleavage functions of CRISPR-Cas systems that result from partial RNA:nucleic acid interactions have been largely overlooked by the current analyses of crRNA spacer targets. These analyses typically look for protospacers with >90-98% homology to a spacer sequence, which would exclude spacers with partial homology to a target (81). In these cases that identify homology with perfect or near perfect homology to the bacterial chromosome, self-targeting spacers have been described, but exist alongside with mutations to the CRISPR system or PAM that would prevent lethal cleavage of the bacterial chromosome. An analysis of crRNA spacers in all available prokaryotic genomes suggested that the majority of spacers with identifiable targets (those with near-perfect homology and therefore likely to induce cleavage) were associated with prokaryotic mobile genetic elements (81). Importantly, these identifiable spacer targets made up just 7% of the spacers in the analysis, which was unable to identify targets (protospacers) for the vast majority (~93%) of spacers (81).

The work presented herein indicates that less than 60% homology of the crRNA that is involved in cleavage is required to direct transcriptional repression, and it is likely that many of the unidentified crRNA targets have partial complementarity that can direct non-cleavage functions. We propose that the targets of CRISPR array spacers, orphan crRNAs, and small RNAs containing repeat-like sequences such as scaRNA, need to be re-analyzed using the parameters we have outlined for non-cleavage Cas9 and Cas12a functions. These have also been

discussed in the previous section about identifying putative scaRNAs. The considerations include searching only for protospacers with undisrupted complementarity to the spacer seed sequence directly adjacent with a PAM and overall ~11-18 bp homology to the processed spacer. Promising binding sites are likely located at places that could contribute to a non-cleavage function, such as proximal to a promoter. This can be used to increase the stringency of the search. Alternatively, if the target sites with reduced complementarity to the spacer exhibit clear patterns of localization, it could independently signal particular functions of partial crRNAs. By changing the parameters of spacer sequence analysis, we will likely uncover many new CRISPR-Cas systems functions that extend far beyond nucleic acid defense and will likely provide insight into important mechanisms of prokaryotic physiology, such as adaptation to environmental changes, stress responses, and inter- and intra-species communication.

Type I-E CRISPR-Cas system of *E. coli* provides a fascinating example of how sequencing historical strains and innovating new approaches to analyzing spacer targets can be utilized to narrow in on unexpected targets and functions of CRISPR systems. The CRISPR arrays of the Type I-E system in *E. coli* have undergone minimal changes in the last 42000 years (181). Given the propensity of bacterial genomes to lose unnecessary genetic information, and previous work demonstrating that this system is activated in response to envelope stress, it has been hypothesized that the ancient spacers make some persistent contribution to *E. coli* physiology (181) (173).

A subsequent systematic analysis of Type 1-E spacers sequences in *E. coli* found that a high frequency of spacers have complementarity to putative targets on the host bacterial genome (173). Interestingly in the study of Type I-E spacers, there was an enrichment for spacer targets in genomic areas where transcription could be regulated, suggesting similarities to the

mechanism of scaRNA activity in *F. novicida (173)*. Although not yet validated experimentally, this study provides preliminary bioinformatic evidence that self-targeting crRNAs may promote binding but not cleavage extends beyond Cas9 and Cas12a systems, possibly even Class I systems.

Experimental evidence of non-cleavage and endogenous crRNA functions

Lending support for a re-analysis of spacers for low-homology targets, there is increasing evidence that the functions of CRISPR-Cas systems may extend widely beyond foreign DNA defense, and that crRNAs may play an important role in diversifying these functions though targeting endogenous nucleic acids. Although there are no natural examples of a crRNA with partial complementarity repressing transcription from endogenous DNA, the best examples of crRNAs influencing bacterial physiology are in the Type I-E and I-F systems of *Pseudomonas aruginosa*, which unlike Cas9 and Cas12a use a multiprotein complex to degrade DNA targets. In the type I-E CRISPR system, a crRNA directs degradation of the mRNA of *lasR*, a transcriptional regulator, which helps the pathogen to evade the mammalian immune system(174). The Type I-F CRISPR system of *P. aruginosa* is involved in regulating motility through partial complementarity of a crRNA to a prophage. For some orthologs of Cas9, crRNAs have been shown to direct Cas9 to degrade ssRNA, although what role this is playing in bacterial physiology has yet to be determined. Furthermore, imperfect binding of crRNAs to target DNA has been shown to inhibit cleavage by Cas9 but not interactions with the DNA altogether. Yet, the role of these Cas9 activities in microbial biology have yet to be fully elucidated, and the broader ability of crRNAs to modulate non-cleavage activities in other native systems, such as the ability of Cas12a to repress endogenous transcription was previously unknown.

Evolution of non-cleavage functions

Our results demonstrating that Cas9 guided by scaRNA represses endogenous transcription and that Cas9 and Cas12a are capable of a similar phenomenon guided by the crRNA, raise the question: can regulatory CRISPR-Cas functions can arise from protective CRISPR systems?

We discuss two non-exhaustive possibilities: 1) Acquisition of a spacer with off-target binding that serendipitously improves fitness and 2) selection for strains that escape self-targeting in conditions where it is also lethal to lose adaptive immunity functions.

In the first model, spacers are acquired that have regulatory activity due to random complementarity to off-target sites on the DNA that is sufficient for binding. This is selected for and maintained either because it is beneficial to the organism or simply, because it isn't costly and the spacer provides a fitness advantage because of its protective role. Given the low complementarity (~11bp) required for Cas9-dependent transcriptional repression in *F. novicida*, it is possible that it could occur by random chance.

We addressed the second possibility experimentally. Due to the lack of a viable phage for *F. novicida*, we were not able to test this with natural adaptation during predation, or to test what escape mutants arise from being forced to take up DNA into the chromosome with a crRNA self-target during phage predation. Instead, we simplified the system, hypothesizing that one route for the selection of regulatory CRISPR-Cas9 functions could be through the avoidance of self-targeting. We utilized the efficient ability of FnoCas9 to cleave DNA targets in order to select for insertion of a sequence into the chromosome that was targeted for cleavage by an FnoCas9 crRNA. The frequency of escape from lethal self-targeting was extremely rare, resulting in just three escape mutants. All three strains contained the same single base deletion in the protospacer

sequence (Appendix F). Fascinatingly, the single base deletion successfully shifted crRNA function from DNA cleavage to transcriptional repression. This experimental model provides one in silico example of how, in the escape from self-targeting, new CRISPR functions can be selected for while maintaining protective CRISPR function.

Multifunctional CRISPR effectors in research and engineering.

The results in Chapter 4 demonstrate that complementary activity of Cas9 and Cas12a in DNA defense and Chapter 5 demonstrates the shared ability of these systems to be commandeered to repress transcription. In both Chapters 4 and 5 we reprogram the crRNA loci for new targets, to cleave DNA (Cas9) or cleave DNA and repress transcription (Cas12a and Cas9). We furthermore reprogrammed scaRNA, the small RNA naturally responsible for guiding Cas9 to repress transcription in *F. novicida* to repress polymyxin resistance genes, re-sensitizing the pathogen to treatment with this antimicrobial. The engineering applications of Cas9 and Cas12a in DNA cleavage are discussed in Chapter 4. The utility of scaRNA-guided Cas9 transcriptional repression and reprogramming enables the same version of Cas9 to be expressed or delivered to a cell while simultaneously targeting DNA and repressing transcription, a technique that can be applied to any system that requires complex fine-tuning of the both RNA levels and DNA sequences, without burdening the cell by expressing a second large protein. Then, by extending these findings to Cas12a systems, we broaden the versatility of the technology. Cas12a has mechanistic differences such as sequence of the PAM and location relative to the spacer, and inducing staggered breaks rather than a blunt cut. Most importantly, Cas12a uses a single crRNA with no accessory tracrRNA that it processes from a CRISPR array itself. Cas12a can therefore be supplied with a CRISPR array containing spacers and can multiplex cleavage- and binding-based activities simultaneously within the cell.

The possibilities for extending these findings are open-ended. Cas12a could be used as the basis for a new category of synthetic repressors that modulate the level of any gene of interest, to a high level of sensitivity, by simply providing crRNAs. In prokaryotes, they could be used to easily screen for interaction partners in gene networks, synthetic lethality and functional redundancy, by simultaneously utilizing the cleavage and repression functions. It could be used to target certain proteins to specific places on the DNA for occlusion, recruitment, or to interrogate DNA packaging, among other possibilities. As more endogenous functions of CRISPR-Cas proteins are discovered, these possibilities will extend even further.

Conclusion

It was observed early in the study of CRISPR-Cas systems that many spacers have homology to foreign mobile genetic elements and their function as prokaryotic adaptive immune systems was demonstrated. The mechanism each system uses to target nucleic acids has been a research priority, and alongside this work, a vast majority of research has focused on developing and applying CRISPR-based technologies to new systems. As a result, CRISPR-Cas systems have revolutionized our capacity to make targeted alterations to nucleic acids in diverse cell types.

CRISPR-Cas systems are ubiquitous in prokaryotes making it imperative to understand their role in biology. Diverse associations between CRISPR-Cas systems and endogenous functions have been identified. It becomes increasingly important to address: what are the functions of CRISPR-Cas systems, and how do they occur?

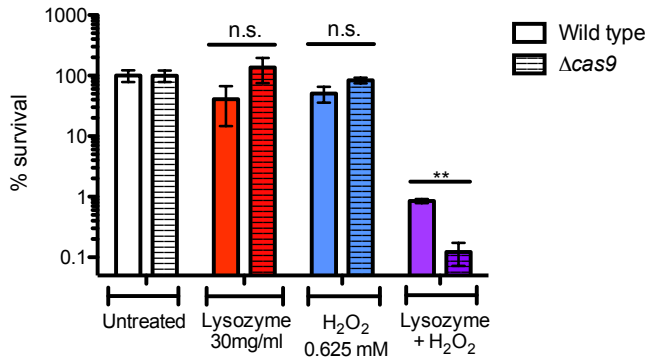
We demonstrated how Cas9 regulates virulence in *F. novicida* by repressing transcription, and extended these findings to show that through a similar mechanism, Cas9 and Cas12a can function similarly as directed by their associated crRNA. Importantly, we found that

these non-cleavage CRISPR functions can exist simultaneously with adaptive immune activity in *F. novicida*. This CRISPR-Cas based mechanism could regulate physiological changes in diverse organisms, and this work provides a framework for identifying new classes of CRISPR-cas system functions. Re-analysis of the crRNAs and small RNAs from the genomes of organisms in which the CRISPR-Cas systems have been associated with other physiological roles will likely uncover still more examples of CRISPR-Cas systems regulating transcription through this mechanism. This should be paired with exploration of the regulation of the CRISPR-Cas systems themselves. We propose that as more non-cleavage functions of CRISPR proteins and crRNA targets are explored, this work might lead to a broader shift in our understanding of CRISPR-Cas systems away from their initial description as primarily prokaryotic adaptive immune systems to a more comprehensive tool that prokaryotes utilize to broadly improve fitness in different environments. Furthermore, we have applied all of these findings to illustrate how both systems can be used as more versatile tools. As the biology of these adaptable and precise systems continues to be explored, mysteries surrounding their functions and crRNA targets will be solved, and new paradigms regarding the breadth of their activity in prokaryotes will continue to be established.

Appendix

Appendix A.

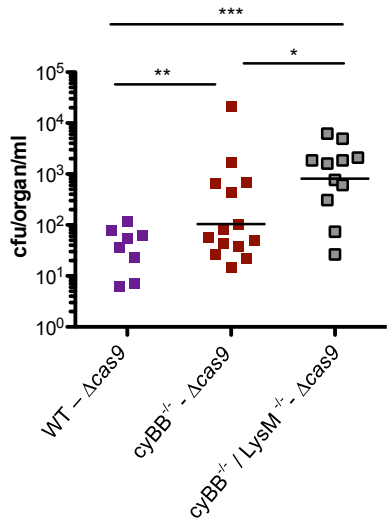
To investigate the role of Cas9 in resisting host antimicrobials, we first looked *in vitro* at the sensitivity of wild-type (WT) *F. novicida* compared to a Cas9 mutant (Δ*cas9*) during exposure to antimicrobials, measured in percent survival relative to untreated samples. We examined the effects of two host antimicrobials that are found in the host, lysozyme and hydrogen peroxide. Hydrogen peroxide (H_2O_2) damages cellular pathways through production of ROS via the Fenton reaction, and lysozyme inserts into and disrupts the membrane, much like polymyxin, while also having peptidoglycan degradation activity. *F. novicida* is highly resistant to lysozyme, and at the highest soluble concentration we were unable to see killing of WT or Δ*cas9* (Figure 1A). We then assessed the survival of *F. novicida* following combination treatment with a concentration of H_2O_2 that is just below the MIC for WT and the highest concentration of lysozyme, both of which were unable to kill WT or Δ*cas9 F. novicida* individually. When combined, the antimicrobials had a synergistic effect, decreasing survival of WT by approximately 2 log from the untreated and individually treated samples. Δ*cas9* was more susceptible to this combination than WT, with survival decreasing by approximately 3 log. This data suggests that Cas9 plays a role in resisting combinatorial killing by host antimicrobials.



Cas9 protects against combinatorial killing by antimicrobial peptides and ROS. Bacteria were incubated at 37 °C shaking in growth media with sub-inhibitory concentrations of antimicrobials both individually and in combination. CFU were enumerated at 4 hrs and represented as % survival of treated relative to untreated. Lysozyme kills by disrupting the membrane and degrading peptidoglycan. *F. novicida* is resistant to lysozyme at the highest soluble concentration (30 mg/mL). When combined with a sub-inhibitory concentration of H_2O_2 (0.625 mM) there was a striking increase in killing of *F. novicida*, with Δ*cas9* more susceptible than WT, n.s.; not significant, **; P<0.005.

Appendix B.

We examined whether Cas9 is important for evasion of these host antimicrobials within the host by infecting a panel of WT and knock out mice with Δ*cas9* and asking if disruption of lysozyme and H_2O_2 production within the context of infection could recover virulence of a Cas9 mutant. We measured the bacterial burden in the liver 48 hours post infection with a low dose of Δ*cas9*. The Δ*cas9* is almost fully attenuated in WT mice, while Mice deficient in H_2O_2 production (cyBB$^{-/-}$) were slightly more susceptible to low-dose infection. LysM$^{-/-}$ cyBB$^{-/-}$ double knockout mice had a significant increase in sensitivity to Δ*cas9* infection compared to WT mice and cyBB$^{-/-}$ mice. This suggests that Cas9 is important for evading combinatorial killing from lysozyme and ROS during infection and dissemination through the mammalian host, suggesting that ROS stress may play an underlying role individually.

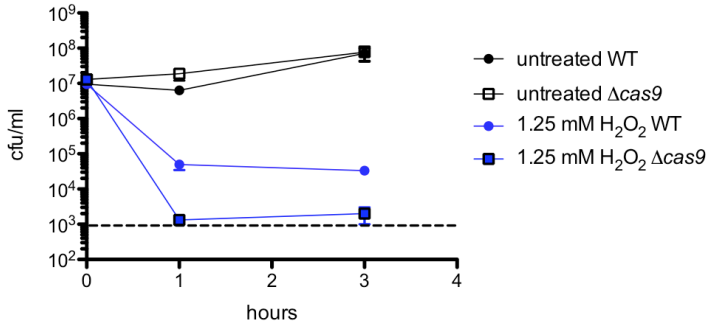


Virulence of Δcas9 in host antimicrobial knockout mice. Mice were infected subcutaneously with a low dose of a Δ*cas9* strain (10^{2-3} colony forming units (cfu)/mouse). Bacterial burden in the liver was measured by cfu 48 hours post infection.
*; P< 0.05, **; P<0.005, ***; P<0.0001

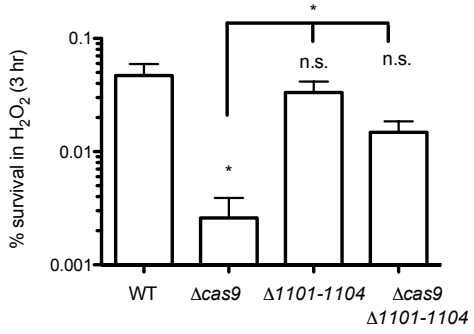
Appendix C.

We compared the sensitivity of WT and Δ*cas9 F. novicida* at higher concentrations of H_2O_2, 2x the MIC. At these concentrations, Δ*cas9* is more susceptible to killing by of H_2O_2 than WT *in vitro* (A). This data suggests that Cas9 may play an important role in resisting H_2O_2 produced by the host. To determine if the sensitivity of Δ*cas9* was due to *1104-1101* expression, we compared the sensitivity of WT, Δ*cas9,* Δ*1104-1101*, and Δ*cas9*Δ*1104-1101* strains to H_2O_2. Δ*cas9,* which expresses 1104-1101 is more susceptible to killing by H_2O_2 than WT *F. novicida* or a 1104-1101 mutant (Δ*1104-1101)* which has no 1104-1101 and thus resembles WT (B). In a Δ*cas9*Δ*1104-1101* strain resistance to H_2O_2 is restored (B). These results indicate that BLP repression is responsible for Cas9-based resistance to H_2O_2.

A



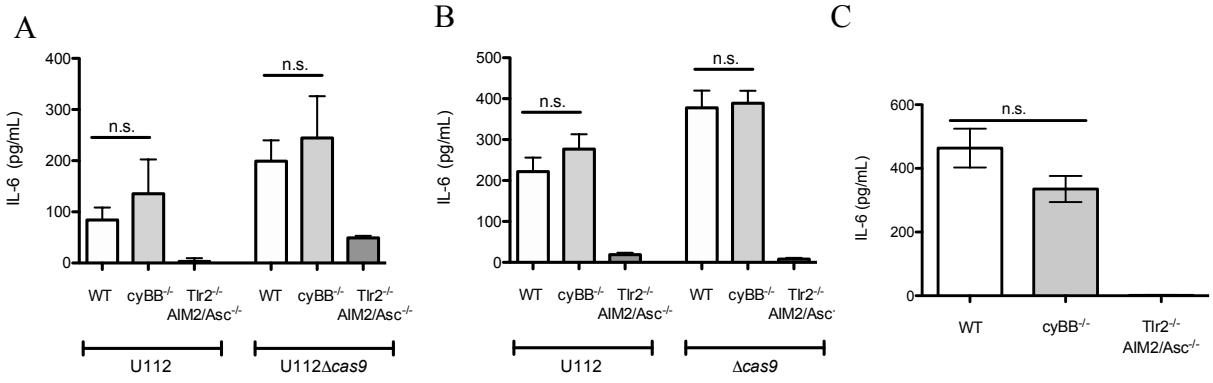
B



BLP repression underlies Cas9-mediated resistance to H2O2. A) WT *F. novicida* and Δ*cas9* were incubated at 37 °C shaking in growth media with 1.25 mM H_2O_2. Colony forming units (cfu) were enumerated at 1 and 3 hrs. B) Bacteria were incubated at 37 °C shaking in growth media with 1.25 mM H_2O_2. CFU were enumerated at 4 hrs, *P<0.05.

Appendix D.

Although Cas9 confers resistance to host antimicrobials, it remains unclear whether the *cas9* mutant is restored in virulence in the knockout mice due to an increased ability to survive in the absence of host antimicrobials, or reduced signaling through the innate immune pathways as a result of decreased membrane disruption in absence of ROS production in cybb$^{-/-}$ mice. To compare the inflammatory response of cybb$^{-/-}$ vs WT mice, we infected BMDM from cybb$^{-/-}$ and WT mice with *F. novicida* U112 (WT) and Δ*cas9*, and measured IL-6 production. Interestingly, we did not see a difference in IL-6 activation between the WT and *cybb*$^{-/-}$ macrophages for either strain (A). To remove external variables like bacterial replication and other physiological differences between Δ*cas9* and WT *F. novicida*, we treated macrophages with heat-killed U112 and Δ*cas9*, as well as the purified immunogenic component of BLP (B-C). In both cases, there was no difference in IL-6 production between the WT and *cybb*$^{-/-}$ macrophages, further confirming that ROS do not contribute to the attenuation of the Δ*cas9* strain through altered Tlr-2 activation.



The absence of ROS production from the host (cyBB-/-) does not alter signaling from WT or Δcas9 F. novicida to TLR2. A) WT, *cyBB*$^{-/-}$, *Tlr2*$^{-/-}$*+AIM2/Asc*$^{-/-}$, BMDM were infected with a 20:1 MOI of WT and Δ*cas9*, and IL-6 production was measured by ELISA 6 hr post infection (n=6-12). B) WT and Δ*cas9* were heat killed at 80°C for 1 hr and BMDM were infected with an MOI of 20:1. IL-6 production was quantified after 6 hr (n=6-12). C) BMDM were treated with 10 ng/mL of Pam3CSK4 (invitrogen) a triacylated lipopeptide that mimics BLPs, and IL-6 was measured after 8 hrs (n=12-13).

Appendix E.

We infected mice deficient in ROS production (*cybb*$^{-/-}$) with the standard infectious dose of *F. novicida* to assess the contribution of ROS alone through in the classic murine model of *F. novicida* pathogenesis. The *cas9* mutant is recovered for virulence in both organs, with a 3 log increase in bacterial burden relative to WT mice in the spleen and 2 log in the liver. This suggests that Cas9 is especially important for evading ROS stress during infection, paralleling the *in vitro* sensitivity of the cas9 mutant to H_2O_2.

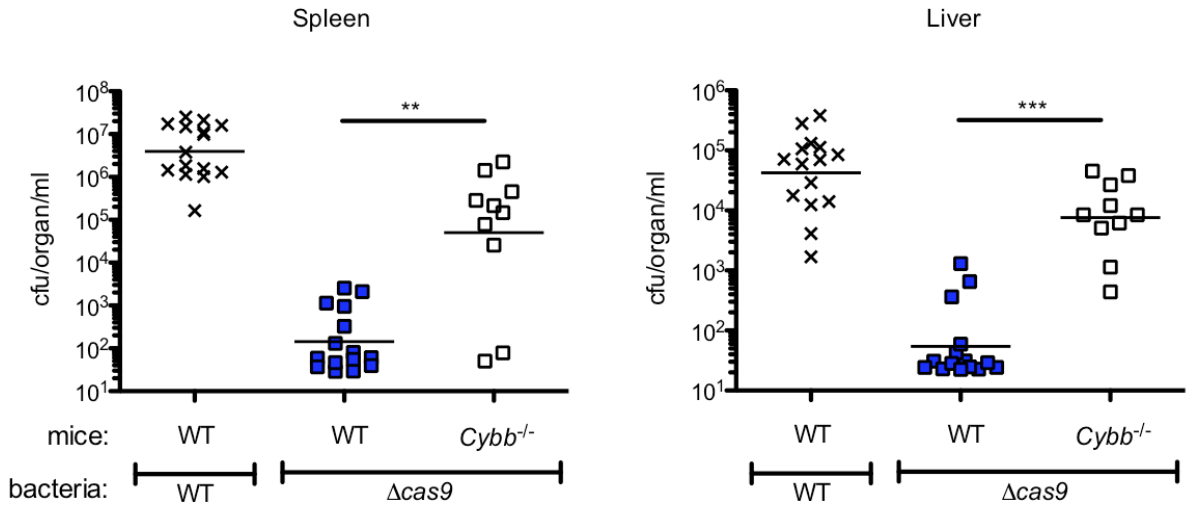
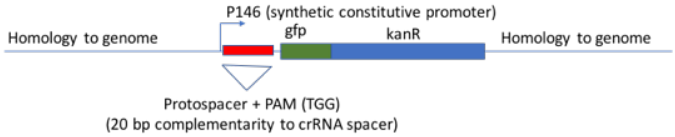


Figure. Cas9 helps evade ROS stress during infection to promote dissemination through a mammalian host. WT (C57BL/J6) and *Cybb*$^{-/-}$ mice were subcutaneously infected with WT or Δ*cas9* bacteria ($2*10^5$ cfu/mouse). Bacterial burden in the spleen and liver was measured by cfu 48 hours post infection. **; P<0.005, ***; P<0.0001.

Appendix F.

F. novicida was forced to take up protospacer into the chromosome with perfect complementarity to a Cas9 crRNA spacer, located between a promoter and *gfp* (A). Escape mutants were very infrequent, with just three isolated from the transformations. All three contained a deletion in a string of Ts in the target site (B). The mutation resulted in transcriptional repression of the downstream *gfp* (C). We confirmed that the transcriptional repression was due to the mutation in the protospacer by amplifying the linear fragment containing the mutated protospacer, and complementing that back into WT and *cas9* mutant strains. In the complemented strains, the same repression phenotype was observed (C). Furthermore, transformation of the allelic exchange fragment with the mutated protospacer was not restricted into WT relative to a *cas9* mutant strain (D). These results suggest that the mutation prevents cleavage but enables Cas9 to bind to the target and repress transcription.

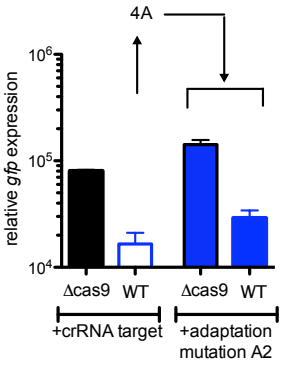
A



B

crRNA target in Δ*cas9*:
gattacttttatagtgtgattgg
Compensatory mutation to crRNA target:
A2. gattac-tttatagtgtgattgg

C



D

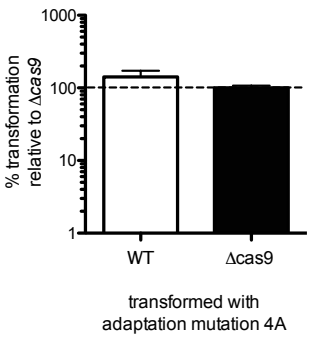


Figure. Self-targeting spacers result in the selection of a subtle target variation that alters Cas9 function from restriction to transcriptional repression. A) Schematic of the allelic exchange fragment (AEF) selected for uptake into the chromosome with kanamycin. A protospacer targeted for cleavage by *F. novicida* Cas9 via a native crRNA was placed between a synthetic promoter and *gfp*. B) The protospacer sequence of the transformants in a Cas9 mutant (same as the original transformed sequence in the AEF) and the WT strain. The WT (+Cas9) strains that survived transformation contained the protospacer mutation "A2". C) qRT-PCR of the relative *gfp* transcript levels in the "crRNA target" and "adaptation mutation" complement tranformants in WT and Δ*cas9* strains. The first two colums showing the "crRNA target" is the *gfp* transcript level in a Δ*cas9* transformed with the "crRNA target" compared to a WT transformant with the "crRNA target", which harbored the mutation A2. This mutation was complemented by amplification of the AEF from the WT transformant DNA, and re-transforming into Δ*cas9* and WT. The relative *gfp* levels in the "adaptation mutation A2" complement was determined. D) transformation efficiency of "adaptation mutation A2" into WT and Δ*cas9* during complementation.

Appendix G.

Table S1. Strains and plasmids used in this study (Chapter 5)

Name	Genotype (Δmutants::insertion constructs)	Reference
CRISPR-Cas9 system mutants and complement strains		
WT	*Francisella novicida* U112 (background strain for all mutants)	(117)
Δ*cas9*	Δ*cas9*	(117)
Δ*cas12a*	Δ*cas12a*	(118)
Δ*cas9* +Cas12a crRNA reprogrammed for 1104-1102	Δ*cas9* + Cas12a CRISPR array replaced with repeat-spacer-repeat to repress the 1104-1102 transcript	This study
Δ*cas9* Δ*cas12a* +Cas12a crRNA reprogrammed for 1104-1102	Δ*cas9* Δ*cas12a* + Cas12a CRISPR array replaced with repeat-spacer-repeat to repress the 1104-1102 transcript	This study
Plasmids with Cas9 crRNA spacer complementarity between promoter and gfp pBAV1K-T5-GFP derived plasmids.[a] T5/lac promoter region driving *gfp* was replaced with p146[b] + (0, 8, 11, 15, and 34 bp (full) crRNA complementarity +PAM)		
0 bp complementarity	pBAV1K-T5-GFP :: p146 + (PAM)	(Ratner et al., 2019)
8 bp complementarity	pBAV1K-T5-GFP :: p146 + (agtgtgattgg)	This study
11 bp complementarity	pBAV1K-T5-GFP :: p146 + (tatagtgtgattgg)	This study
15 bp complementarity	pBAV1K-T5-GFP :: p146 + (cttttatagtgtgattgg)	This study
34 bp complementarity	pBAV1K-T5-GFP :: p146 + (acaatggcaagcttgattacttttatagtgtgattgg)	Price et al., 2015)
Plasmids with Cas12a crRNA spacer complementarity between promoter and gfp pBAV1K-T5-GFP derived plasmids.[a] T5/lac promoter region driving *gfp* was replaced with p146[b] + (PAM + 0, 8, 11, 15, and 29 bp (full) crRNA complementarity)		
0 bp complementarity	pBAV1K-T5-GFP :: p146 + (PAM)	This study
8 bp complementarity	pBAV1K-T5-GFP :: p146 + (tttaagattaaa)	This study
11 bp complementarity	pBAV1K-T5-GFP :: p146 + (tttaagattaaaagg)	This study
15 bp complementarity	pBAV1K-T5-GFP :: p146 + (tttaagattaaaaggtaat)	This study
29 bp complementarity	pBAV1K-T5-GFP :: p146 + (tttaagattaaaaggtaattctatcttgttgag)	This study

[a] pBav vector (176)

[b] p146 (169)

Table S2. RT-PCR and Cloning Primer Sequences (Chapter 5)

#	Name	Sequence 5′-3′	Usage
1	1104_rt_fwd	CAGAGTCAAACTCCGCTGCG	qRT-PCR
2	1104_rt_rev	TGGCAGCCAAAATGCGGTAG	qRT-PCR
3	1103_rt_fwd	ATGGTGGGCAGTCTAGCGCA	qRT-PCR
4	1103_rt_rev	ACCCAACTCACCATCGCCACA	qRT-PCR
5	1102_rt_fwd	GAAGCTTTTGATCGCGATTATAGTA	qRT-PCR
6	1102_rt_rev	TGTAACTTGCCAATATAGCAACACT	qRT-PCR
7	1101_rt_fwd	TGAACAAGCAGATACGCCATGGGT	qRT-PCR

#	Name	Sequence 5′-3′	Usage
8	1101_rt_rev	CCAACAGCCTGCCTGCCACT	qRT-PCR
9	HelC_RT_Fwd	GGGATGTCGCCTTTTGATTTTC	qRT-PCR housekeeping gene
10	HelC _RT_Rev	CTCTTTTGTCCCTTGTGCTTGC	qRT-PCR housekeeping gene
11	gfp_Fwd	GTGCCATGCCCGAAGGTTAT	qRT-PCR
12	gfp_Rev	GCGTCTTGTAGTTCCCGTCA	qRT-PCR
13	Pbav_kan_Fwd	CTTCGGGCTTTTCCGTCTTT	qRT-PCR
14	Pbav_kan_Rev	GCCGATGTGGATTGCGAAAA	qRT-PCR
Cas12a crRNA reprogrammed for 1104-1102			
15	Cpf1_bLPTarget_LarmFwd	GCTGTTTCCAAGCATTGATAC	cloning
16	crRNA2_L_rev_ArsR	TATCGATCCTGCAGCTATGCTCTGTATATTATTGATTTCTAAATTAGAATTTTCTAATAT	coning
17	crRNA2_Kan_fwd	ATATTAGAAAATTCTAATTTAGAAATCAATAATATACAGAGCATAGCTGCAGGATCGATA	cloning
18	ReprgCpf#1_RKanRev_Int (1st round of amplification with crRNA2_KanFwd)	CTTTAAATAATTTCTACTGTTGTAGATGAGGTCGACGGTATCGATAAGCGAAGTTCC	cloning
19	Cpf1_reprgspcr#3Rev Out (2nd round of amplification with crRNA2_KanFwd)	CTATGATATGTATACGTACTGTCGCTAACTAGTCTAAGAACTTTAAATAATTTCTACTG	cloning
20	Cpf1_bLPTarget_LarmRev (kan rev, 3nd round of amplification primer)	CTACTGTTGTAGATTATAAATTAGCTGTATACGTACTGTCGCTAACTAGTC	cloning
21	ReprgCpf#1_RarmFwd_int (1st round of amplficiation with Cpf1_bLPTarget_RarmRev)	CAACAGTAGAAATTATTTAAAGTTCTTAGACCCGTTTTTGCCTAAATCAGCAAAAAAAG	cloning
22	Cpf1_bLPTarget_RarmFwd (2st round of amplficiation with Cpf1_bLPTarget_RarmRev)	CGTATACAGCTAATTTATAATCTACAACAGTAGAAATTATTTAAAGTTCTTAGACCCGTTTTTGCCTAAATCAGC	cloning
23	Cpf1_bLPTarget_RarmRev	GACTTGGTATAAGCGTAAATC	cloning
24	Cpf1_bLPTarg_AmpliFwd	CTGTGATACTGTAAATGAAAGC	mutant confirmation
25	Cpf1_bLPTarg_AmpliRev	CATATAATGTTGTTTTAGACTTG	mutant confirmation
26	Cpf1_bLPTarg_seqRev	CAAATTTCTAGAAAGTTATGGTG	mutant confirmation
***Δcas12a*:Kan^r**			
27	cas12a_L_fwd	CAGAAAGCTTAGTTTTCACCTCATC	cloning
28	cas12a_L_rev	CTTATCGATACCGTCGACCTCTTCATTCAAGAATATATTACCCTGTCAG	cloning
29	cas12a-Kan-Fwd	CTGACAGGGTAATATATTCTTGAATGAAGAGGTCGACGGTATCGATAAG	cloning
30	cas12a-Kan-Rev	GAAATGTAGAGAATTTTATAAGGAGTCTTTATCGCATAGCTGCAGGATCGATATC	cloning
31	cas12a_R_fwd	GATATCGATCCTGCAGCTATGCGATAAAGACTCCTTATAAAATTCTCTACATTTC	cloning
32	cas12a_R_rev	GTATTGTTTGATATATGGAGTAGCTGG	cloning
33	Cpf1_amplf4seq Rev	GGATAACAAGTGAAAATAGCCAGAT	Mutant confirmation

#	Name	Sequence 5′-3′	Usage
34	Cpf1_amplif4seq Fwd	CTCTACAGAAAATTGCCCGAG	Mutant confirmation
35	Fwdseq_LKanjxn_Cpf 1	CCTTGCTAAACTTATATCTTGCTCAAC	Mutant confirmation
36	Fwdseq_RKanjxn_Cpf 1	CTCCTTCATTACAGAAACGGC	Mutant confirmation
Plasmids with Cas9 crRNA complementarity between promoter and gfp			
37	33-39 Pbav _p146_Rev_Universal	CCCTTATCTAATTTATATCAGAATCATTTATATATGTCAATTGGTATTT ATAAATTTTTTATGATTTCCTCTAG	cloning
38	39_Pbav_p146_0sca_ Fwd	GATATAAATTAGATAAGGGTGGATGCGTAAAGGAGAAGAACTTTTC	cloning
39	Pbav_RC_seq primer	GCACTCTTGAAAAAGTCATAC	sequencing
40	49Cas9_34bpSpcr1_P bav146_R	CACACTATAAAAGTAATCAAGCTTGCCATTGTCCCTTATCTAATTTAT ATCAGAATC	cloning
41	49Cas9_34bpSpcr1_P bav146_F	CAATGGCAAGCTTGATTACTTTTATAGTGTGATTGGATGCGTAAAGG AGAAGAACTTTTC	cloning
42	50Cas9_15bpSpcr1_P bav146_R	CCAATCACACTATAAAAGCCCTTATCTAATTTATATCAGAATC	cloning
43	50Cas9_15bpSpcr1_P bav146_F	CTTTTATAGTGTGATTGGATGCGTAAAGGAGAAGAACTTTTC	cloning
44	51Cas9_11bpSpcr1_P bav146_F	GATAAGGGTATAGTGTGATTGGATGCGTAAAGGAGAAGAACTTTTC	cloning
45	51Cas9_11bpSpcr1_P bav146_R	CCAATCACACTATACCCTTATCTAATTTATATCAGAATC	cloning
46	52Cas9_8bpSpcr1_Pb av146_F	GATAAGGGAGTGTGATTGGATGCGTAAAGGAGAAGAACTTTTC	cloning
47	52Cas9_8bpSpcr1_Pb av146_R	CCAATCACACTCCCTTATCTAATTTATATCAGAATC	cloning
Plasmids with cas12a crRNA complementarity between promoter and gfp			
48	40Cpf1_0bp_Pbav146 _R	CCTTTACGCATTAAACCCTTATCTAATTTATATCAG	cloning
49	40Cpf1_0bp_Pbav146 _F	GATAAGGGTTTAATGCGTAAAGGAGAAGAACTTTTC	cloning
50	41Cpf1_8bp_Pbav146 _R	TTTAATCTTAAACCCTTATCTAATTTATATCAG	cloning
51	41Cpf1_8bp_Pbav146 _F	GATAAGGGTTTAAGATTAAAATGCGTAAAGGAGAAGAACTTTTC	cloning
52	42Cpf1_11bp_Pbav14 6_F	GATAAGGGTTTAAGATTAAAAGGATGCGTAAAGGAGAAGAACTTTTC	cloning
53	43Cpf1_15bp_Pbav14 6_F	GATAAGGGTTTAAGATTAAAAGGTAATATGCGTAAAGGAGAAGAACT TTTC	cloning
54	44Cpf1_20bp_Pbav14 6_F	GATAAGGGTTTAAGATTAAAAGGTAATTCTATATGCGTAAAGGAGAA GAACTTTTC	cloning
55	45Cpf1_29bp_Pbav14 6_Fa	GATAAGGGTTTAAGATTAAAAGGTAATTCTATCTTGTTGAGATGCGT AAAGGAGAAGAAC	cloning
56	45Cpf1_29bp_Pbav14 6_Rb	CTCAACAAGATAGAATTACCTTTTAATCTTAAACCCTTATCTAATTTA TATCAG	cloning
57	45Cpf1_29bp_Pbav14 6_Fb	TTTAAGATTAAAAGGTAATTCTATCTTGTTGAGATGCGTAAAGGAGA AGAACTTTTC	cloning

Appendix H.

Cas9-mediated targeting of viral RNA in eukaryotic cells

Aryn A. Price[1,2,3,#], Timothy R. Sampson[1,2,3#§], Hannah K. Ratner[1,2,3], Arash Grakoui[2,3,4*], and David S. Weiss[2,3,4*]

[1]Department of Microbiology and Immunology, Microbiology and Molecular Genetics Program, Emory University, Atlanta, GA 30329.

[2]Emory Vaccine Center, Emory University, Atlanta, GA 30329.

[3]Yerkes National Primate Research Center, Emory University, Atlanta, GA 30329.

[4]Division of Infectious Diseases, Department of Medicine, Emory University School of Medicine, Atlanta, Georgia 30329.

[§]Current address: Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, California, 91125.

[#]These authors contributed equally

Published in PNAS, May12, 2015: https://www.pnas.org/content/112/19/6164

*Corresponding authors: Arash Grakoui
Emory Vaccine Center
954 Gatewood Rd, Rm 3020
Atlanta, GA 30329
Tel: (404) 727-7217
Email: arash.grakoui@emory.edu

David S. Weiss
Emory Vaccine Center
954 Gatewood Rd, Rm 2028
Atlanta, GA 30329
Tel: (404) 727-8214
Email: david.weiss@emory.edu

Abstract:

CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) - Cas (CRISPR-associated) systems are prokaryotic RNA-directed endonuclease machineries that act as an adaptive immune system against foreign genetic elements (26, 182). Utilizing small CRISPR RNAs (crRNAs) that provide specificity, Cas proteins recognize and degrade nucleic acids (13). We previously demonstrated that the Cas9 endonuclease from *Francisella novicida* (FnCas9) is capable of targeting RNA (164). Here, we show that FnCas9 can be directed by an engineered RNA-targeting guide RNA (rgRNA) to target and inhibit a human +ssRNA virus, hepatitis C virus, within eukaryotic cells. This work reveals a versatile and portable RNA targeting system that can effectively function in eukaryotic cells and be programmed as an antiviral defense.

Significance Statement: The Cas9 endonuclease has quickly become a revolutionary tool in genome engineering. Utilizing small guiding RNAs, Cas9 can be targeted to specific DNA sequences of interest, where it catalyzes DNA cleavage. We now demonstrate that Cas9 from the Gram-negative bacterium *Francisella novicida* (FnCas9) can be reprogrammed to target a specific RNA substrate in eukaryotic cells, the genome of the +ssRNA virus, hepatitis C virus (HCV). Further, this targeting results in inhibition of viral protein production. Overall, programmable Cas9-mediated viral RNA targeting likely represents one of myriad potential applications of FnCas9 in RNA targeting in eukaryotic cells.

Manuscript.

Our recent work revealed a unique form of prokaryotic gene regulation, whereby Cas9 from *Francisella novicida* (FnCas9) targets a bacterial mRNA, leading to decreased mRNA stability and ultimately gene repression (164). Given the ability of specific Cas9 proteins to be reprogrammed to target and cleave DNA in numerous biological systems (87, 183, 184), we hypothesized that FnCas9 could be retargeted to a distinct RNA in eukaryotic cells and lead to its inhibition. To eliminate any confounding interactions of FnCas9 with DNA we targeted FnCas9 to the +ssRNA virus, hepatitis C virus (HCV), which has no DNA stage in its lifecycle. HCV is an important human pathogen associated with liver fibrosis, cirrhosis, hepatocellular carcinoma, and is the leading cause of liver transplantation (185, 186).

To target the RNA of HCV, we engineered a small RNA, which we term an RNA-targeting guide RNA (rgRNA). The rgRNA is similar in structure to that naturally created by the *F. novicida* tracrRNA and scaRNA, which are utilized for endogenous mRNA targeting (164). It consists of a dsRNA region thought to be important for interaction with Cas9, and a ssRNA targeting sequence complementary to a portion of the highly conserved HCV 5' untranslated region (UTR), involved in both translation of the viral polyprotein and replication of the viral RNA (Fig. 1A, and Fig. S1). Vectors encoding either this rgRNA or FnCas9 (Fig. S2 and Supplementary Sequence File) were transfected into human hepatocellular carcinoma cells (Huh-7.5) and subsequently infected with a previously described HCVcc genotype 2a, recombinant virus encoding *Renilla* luciferase (187). Expression of both the 5' UTR-targeting rgRNA and FnCas9 together reduced the levels of viral proteins, as measured by immunostaining for the E2 glycoprotein (Figs. 1B, C) or quantification of luciferase production (Fig. 1D). Conversely,

expression of either the rgRNA or FnCas9 alone had no effect (Figs. 1B-D), nor did expression of a non-specific rgRNA and FnCas9 (Figs. 1B-D), demonstrating the specificity of this system. Additionally, an rgRNA complementary to a portion of the 3' UTR, necessary for replication of viral RNA, decreased viral protein levels similarly (Figs. 1A-D), demonstrating that the effect was not specific to a single site in the HCV genome. Therefore, FnCas9 can be programmed by a single rgRNA to target the RNA of a human virus in eukaryotic cells, leading to viral inhibition.

In order to determine if FnCas9 was directly associated with HCV RNA, we performed co-immuoprecipitation experiments. We transfected cells with an HA epitope-tagged version of the protein (which maintained its ability to inhibit HCV [Figs. S3A, B]) as well as the 5' UTR-targeting rgRNA and subsequently infected the cells with HCV. FnCas9 was immunoprecipitated from cell lysates, associated RNA was purified, and quantitative real-time PCR was performed for the rgRNA and HCV RNA. The rgRNA was present in association with FnCas9 (Fig. 2A), as was HCV RNA (Fig. 2B), suggesting that HCV RNA was directly targeted by the FnCas9:rgRNA complex. A non-specific rgRNA did not facilitate this interaction (Fig. 2B), but it did associate with FnCas9 (Fig. 2C). Thus, FnCas9 can be targeted to directly associate with viral RNA within eukaryotic cells.

We next sought to determine how FnCas9 was inhibiting HCV—whether through degrading viral RNA, inhibiting translation, or blocking viral replication. We first found that addition of a nuclear localization signal (NLS) to FnCas9 abrogated its repression of viral protein production (Fig. 3A and Fig. S4), in line with its targeting of cytosolic HCV RNA. Since Cas9 proteins including FnCas9 are known to cleave DNA through two conserved structural domains, the

RuvC and HNH endonuclease domains (87), it is possible that these regions are important for inhibiting HCV. We therefore generated alanine point mutations in the conserved RuvC and HNH active sites of FnCas9 (D11A and H969A, respectively). Despite mutation in one or both of these domains, FnCas9 maintained its ability to inhibit HCV (Fig. 3B). However, mutation in the RNA-binding arginine-rich motif (ARM; R59A), which is necessary for interaction with nucleic acids both within *F. novicida* and in orthologous Cas9 (91, 92, 164), resulted in diminished HCV inhibition (Fig. 3B). It is therefore likely that FnCas9 inhibits HCV not through target cleavage, but instead by associating with the target RNA and preventing the function of the translational and/or replication machineries.

We subsequently tested whether FnCas9 could inhibit translation of HCV. We performed an *in vitro* translation reaction using immunoprecipitated FnCas9 from transfected Huh-7.5 cells, purified RNA from cells transfected with either the 5' UTR-targeting rgRNA or the non-targeting RNA, as well as HCV genomic RNA. Addition of both FnCas9 and the 5' UTR-targeting rgRNA resulted in decreased translation of the HCV genome, as measured by viral luciferase production (Fig. 3C). Furthermore, a catalytically inactive FnCas9 (D11A/H969A) maintained its ability to inhibit translation of HCV, while the ARM (R59A) mutant displayed less translational inhibition (Fig. 3C). Taken together, these data suggest that FnCas9 is capable of directly inhibiting translation of target RNA. In order to determine whether FnCas9 also inhibits replication of HCV RNA, we targeted the negative-sense strand (generated during replication and which is untranslated) of the 5' UTR. Such targeting resulted in inhibition of HCV (Fig. S5), suggesting that FnCas9 is capable of inhibiting viral replication as well. Overall,

these data strongly support a model whereby FnCas9 binds HCV RNA, does not cleave the target, but inhibits the function of both translational and replication machineries.

We subsequently tested the sequence requirements for RNA targeting. Cas9 proteins require a short sequence adjacent to the targeted region, called a proto-spacer adjacent motif (PAM), to cleave DNA (188). We sought to determine if a similar conserved adjacent region was necessary for HCV inhibition. A 5' UTR targeting rgRNA shifted to lack similar adjacent sequences still inhibited HCV (Fig. 4A). In fact, no common features are observed in the sequences adjacent to the targets of rgRNAs utilized in this study (Fig. S6A). In contrast, DNA targeting by FnCas9 endogenously within *F. novicida* absolutely requires a PAM (Figs. S6B, C). Together, these data demonstrate that FnCas9-mediated inhibition of HCV is PAM-independent.

DNA targeting RNAs utilized by Cas9 require a 3' seed sequence within the targeting region, and even a single mismatch in this region can abrogate function (132, 189). To test if there was a similar requirement for RNA targeting, we generated a panel of rgRNA mutants containing mismatches within the targeting sequence. We found that mutation of up to 6 bases within the 3' targeting region of the rgRNA were tolerated, without loss of HCV inhibition (Fig. 4B). Longer regions of mismatched bases at the 3' end resulted in a loss of activity (Fig. 4B). Specific mismatches in the 5' region of rgRNA were also non-functional (Fig. 4C). A single mismatch in either the first or second base was sufficient to abrogate viral inhibition (Fig. 4C). However, a mismatch in the third base alone did not lead to a loss of activity (Fig. 4C). These data strongly suggest that unlike DNA targeting by other Cas9 proteins, FnCas9 inhibition of HCV requires a

critical sequence in the 5' end rather than the 3' end of the targeting region of the guiding RNA (132, 189).

The previous experiments demonstrated that FnCas9 could target an RNA and facilitate resistance to HCV infection. We next tested whether FnCas9 could be introduced following an established viral infection and inhibit the virus (Fig. 5A). Transfection of HCV-infected Huh-7.5 cells with FnCas9 and 5' or 3' UTR targeting rgRNAs resulted in decreased viral protein production (Fig. 5B), while expression of either FnCas9 or rgRNAs alone was not sufficient to inhibit HCV infection (Fig. 5B). Thus, the FnCas9:rgRNA machinery can combat both new and established viral infections.

These data demonstrate the successful adaptation of the CRISPR-Cas prokaryotic immune system as an intracellular eukaryotic antiviral defense. While other CRISPR-Cas systems can target RNA in archaea (23, 141, 190) and bacteria (17), and recently SpCas9 has been shown to cleave RNAs *in vitro* (191), this work represents the reprogramming of a Cas protein (FnCas9) to target RNA within a eukaryotic cell. Intriguingly, we find that orthologous Cas9 proteins from diverse Type II CRISPR-Cas families, including *Streptococcus pyogenes, S. thermophilus,* and *Neisseria meningitidis*, are also capable of inhibiting HCV during cellular infection (Figs. S7,8). This suggests a broader capability of diverse Cas9 proteins to target and associate with RNA targets. Our results further demonstrate that FnCas9 inhibition of HCV is PAM-independent, unlike the *in vitro* RNA targeting ability of *Streptococcus pyogenes* Cas9 which requires exogenous PAM encoding oligomers (191). Thus, FnCas9-mediated RNA inhibition may be more flexible in its targeting. Importantly, this targeting is highly specific. Compared to cells

with a vector control, significant changes in host cell gene expression were not observed in cells expressing Cas9 in conjunction with either the HCV targeting or non-specific rgRNA, demonstrating the specificity of these complexes (Fig. S9 and Supplementary Table 1). Furthermore, the presence of FnCas9 in the cytosol is necessary to inhibit HCV, in contrast to DNA targeting by Cas9 that relies on nuclear localization (192). Cytosolic RNA targeting would potentially limit FnCas9 off-target effects on host DNA. Since the lifecycles of viruses with both RNA and DNA genomes require an RNA stage (generated during transcription, replication, or both), and FnCas9 can target both negative- and positive-sense strands of RNA and inhibit by blocking both viral translation and replication machineries (Fig. S10), it is likely that the FnCas9:rgRNA machinery can be utilized to target diverse viruses, including both +ssRNA viruses (such as flavivirus, poliovirus, and rhinovirus) and –ssRNA virus (such as filovirus, paramyxovirus and orthomyxovirus). Furthermore, some eukaryotic viruses have mechanisms to circumvent eukaryotic RNA-targeting antiviral defenses, such as classical RNAi systems (193-195); however, these viruses have not evolved in the presence of Cas9, so it is unlikely that they have Cas9 evasion strategies. Thus, the FnCas9:rgRNA machinery could facilitate the targeting of viruses as soon as their genome sequences are available, without knowledge of the virus lifecycle or host receptors.

Given the vast success of Cas9 as a mediator of genome engineering in a multitude of species (87, 90, 132, 183, 184, 196-201) our data suggest that FnCas9 could be used in a broad range of systems, representing a new paradigm in Cas9-mediated genetic engineering. This work demonstrates a portable, inter-domain machinery capable of viral inhibition, likely just one of myriad potential biotechnological and medical applications of Cas9-mediated RNA targeting

Materials and Methods

Plasmid Construction. FnCas9 was amplified using the primers found in Supplementary Table 2 and cloned into the *XbaI* and *PmeI* sites of pcDNA3.3 (Invitrogen, Grand Island, NY). FnCas9 point mutants were amplified from strains published previously(164). StCas9, SpCas9, and NmCas9 were amplified from Addgene vectors #48669, 41815, 47872, respectively. To create rgRNA vectors, *F. novicida* CRISPR repeat sequences and the indicated targeting sequence were cloned into the gRNA-encoding plasmid from the Church Lab (Addgene #41824)(183), within the pCR-Blunt-II (Invitrogen) backbone, using overlapping PCR and the primers indicated in Supplementary Table S1.

Cell lines and culture conditions. Huh-7.5 cells were maintained in Dulbecco's Modified Eagle Medium (DMEM, BioWhittaker, Walkersville, MD) containing 10% fetal bovine serum (FBS, Hyclone, Logan, UT) and 100 ug/ mL of penicillin/ streptomycin (Cellgro, Herndon, VA) at 37 degrees in 5% CO_2.

Plasmid Transfection. Huh-7.5 cells were seeded to 80% confluence in 24 well plates in DMEM without antibiotics the day prior to transfection. Eight hundred ng of total plasmid DNA were transfected using Lipofectamine™ 3000 transfection reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. Six hours following transfection, cell supernatants were aspirated and replaced with DMEM supplemented with FBS and antibiotics.

Virus transcription and transfection. Rluc virus utilized for luciferase assays encodes the *Renilla* luciferase gene between the p7 and NS2 coding sequences of the previously described J6/

JFH genotype 2a infectious clone Cp7(187). Rluc and Cp7 viral RNA were prepared as previously described(187). Purified plasmid encoding the cDNA copy of the full-length viral genome was linearized and 5' overhangs were digested with Mung Bean Nuclease (New England Biolabs, Ipswich, MA). Linearized DNA was purified with phenol-chloroform extraction and ethanol precipitation. Transcription of the linearized DNA template was performed using a MEGAscript^{TM} T7 kit and the linear template was treated with DNAse I (Ambion, Austin, TX). RNA was purified with phenol-chloroform extraction and isopropanol precipitation, and concentration and purity were determined using a Nanodrop spectrophotometer and standard agarose gel electrophoresis.

Huh-7.5 cells were grown to 70% confluence, trypsinized, and washed twice in PBS. Ten μg purified RNA were mixed with .4 mL Huh-7.5 cells suspended at a concentration of 2 x 10^7 cells/ mL. Samples were electroporated in BTX 2 mm gap cuvettes (Harvard Apparatus, Inc., Holliston, MA) using an ECM 830 apparatus (BTX Genetronics, San Diego, CA) with five pulses of 99 usec at 820 V over 1.1 sec. Cells were suspended in 25 mL complete DMEM and virus was harvested and stored at –80 degrees C three days following transfection.

Immunoprecipitation. Anti-HA IP was performed according to the manufacturer's instructions from lysates of Huh-7.5 cells infected with HCVcc and transfected with FnCas9 and rgRNA expression vectors as indicated (Sigma-Aldrich, St. Louis, MO). Total RNA was extracted from the precipitated fraction using an RNeasy Mini Kit (Qiagen, Valencia, CA) according to the manufacturer's instructions.

In vitro translation assay. Immunoprecipitated FnCas9 was incubated with 1µg HCV RNA, and 500ng RNA from Huh-7.5 cells transfected with either the 5' UTR – targeting rgRNA or the Control rgRNA, in conjunction with the rabbit reticulocyte lysate *in vitro* translation kit (Life Technologies, Carlsbad, CA) according to the manufacturer's instructions. Translated viral luciferase was measured as described below.

Quantitative RT-PCR. Quantitative reverse transcription reactions were performed using Taqman One Step RT-PCR Master Mix Reagent (Applied Biosystems, Hammonton, NJ) and primers specific for the 5' untranslated region of HCV (Supplementary Table 1). Sample analysis was performed on an Applied Biosystems 7500 apparatus. rgRNAs were quantified via Syber Green One Step RT-PCR, with primers found in Supplementary Table 1.

Luciferase Assays. Huh-7.5 cells in a 96 well plate format were lysed and analyzed for relative light activity using the Renilla Luciferase Assay System (Promega, Madison, WI) according to the manufacturer's instructions. Briefly, 50 ul of Renilla substrate in assay buffer was added to 50 ul of cell lysate and relative light units were quantitated on a Clarity 4.0 microplate luminometer (Biotek, Winooski, VT).

Immunohistochemistry. Six thousand Huh-7.5 cells per well were plated in collagen coated 96 well plates. The following day, cells were infected with the J6/JFH genotype 2a virus Cp7. Following three days of incubation, cells were fixed with methanol, washed twice with PBS, and permeabilized with PBS containing .1% Tween-20 (PBS-T). Fixed cells were incubated in blocking buffer containing 1% bovine serum albumin and .2% skim milk for thirty minutes.

Endogenous peroxidase was blocked using 3% H_2O_2, then cells were washed twice with PBS and once with PBS-T. Cells were then incubated with the 2C1 monoclonal antibody to HCV E2 glycoprotein for one hour at room temperature. After two washes with PBS and one with PBS-T, cells were incubated with ImmPress anti-mouse HRP (Vector Laboratories, Burlingame, CA), washed, and developed using DAB substrate (Vector Laboratories).

Acknowledgements: We thank Rafi Ahmed for helpful discussions, Sarkis Mazmanian for critical support, and Steven Bosinger of the Yerkes Genomics Core for performance of RNAseq and its analysis. This project was supported by NIH grants R01-AI11070 and U54-AI057157 (Southeastern Regional Center of Excellence for Emerging Infections and Biodefense), and the Burroughs Wellcome Fund to D.S.W, and Yerkes Research Center Base Grant P51RR-000165 and NIH grants R01-AI070101 and R01-DK083356 to A.G. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. T.R.S. was supported by the NSF GRFP and the ARCS Foundation. The authors have filed a related patent. A.A.P., T.R.S., and H.K.R. performed the experiments; all authors conceived and designed the experiments, interpreted data and wrote the manuscript.

References

1. Sorek R, Lawrence CM, & Wiedenheft B (2013) CRISPR-mediated adaptive immune systems in bacteria and archaea. *Annu Rev Biochem* 82:237-266.
2. Makarova KS*, et al.* (2011) Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* 9(6):467-477.

3. Barrangou R*, et al.* (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315(5819):1709-1712.

4. Sampson TR, Saroj SD, Llewellyn AC, Tzeng YL, & Weiss DS (2013) A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. *Nature* 497(7448):254-257.

5. Mali P*, et al.* (2013) RNA-guided human genome engineering via Cas9. *Science* 339(6121):823-826.

6. Cong L*, et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339(6121):819-823.

7. Jinek M*, et al.* (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337(6096):816-821.

8. Shepard CW, Finelli L, & Alter MJ (2005) Global epidemiology of hepatitis C virus infection. *Lancet Infect Dis* 5(9):558-567.

9. Scheel TK & Rice CM (2013) Understanding the hepatitis C virus life cycle paves the way for highly effective therapies. *Nat Med* 19(7):837-849.

10. Mateu G, Donis RO, Wakita T, Bukh J, & Grakoui A (2008) Intragenotypic JFH1 based recombinant hepatitis C virus produces high levels of infectious particles but causes increased cell death. *Virology* 376(2):397-407.

11. Jinek M*, et al.* (2014) Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* 343(6176):1247997.

12. Nishimasu H*, et al.* (2014) Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* 156(5):935-949.

13. Sternberg SH, Redding S, Jinek M, Greene EC, & Doudna JA (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*:Epub ahead of print:Jan 31, 2014.

14. Jiang W, Bikard D, Cox D, Zhang F, & Marraffini LA (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol* 31(3):233-239.

15. Semenova E*, et al.* (2011) Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proceedings of the National Academy of Sciences of the United States of America* 108(25):10098-10103.

16. Hale CR*, et al.* (2009) RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* 139(5):945-956.

17. Zhang J & White MF (2013) Hot and crispy: CRISPR-Cas systems in the hyperthermophile Sulfolobus solfataricus. *Biochem Soc Trans* 41(6):1422-1426.

18. Zebec Z, Manica A, Zhang J, White MF, & Schleper C (2014) CRISPR-mediated targeted mRNA degradation in the archaeon Sulfolobus solfataricus. *Nucleic acids research* 42(8):5280-5288.

19. Staals RH*, et al.* (2013) Structure and activity of the RNA-targeting Type III-B CRISPR-Cas complex of Thermus thermophilus. *Mol Cell* 52(1):135-145.

20. O'Connell MR*, et al.* (2014) Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature*.

21. Hsu PD, Lander ES, & Zhang F (2014) Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell* 157(6):1262-1278.

22. Sullivan CS & Ganem D (2005) A virus-encoded inhibitor that blocks RNA interference in mammalian cells. *J Virol* 79(12):7371-7379.

23. Singh G, *et al.* (2009) Suppression of RNA silencing by Flock house virus B2 protein is mediated through its interaction with the PAZ domain of Dicer. *FASEB J* 23(6):1845-1857.

24. Chao JA, *et al.* (2005) Dual modes of RNA-silencing suppression by Flock House virus protein B2. *Nat Struct Mol Biol* 12(11):952-957.

25. Mali P, *et al.* (2013) CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol* 31:833-838.

26. Qi LS, *et al.* (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152(5):1173-1183.

27. Bikard D, *et al.* (2013) Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic acids research* 41(15):7429-7437.

28. Hou Z, *et al.* (2013) Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. *Proceedings of the National Academy of Sciences of the United States of America*.

29. Chang N, *et al.* (2013) Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. *Cell Res* 23(4):465-472.

30. Gratz SJ, *et al.* (2013) Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease. *Genetics*.

31. Perez-Pinera P, *et al.* (2013) RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat Methods* 10:973-976.

Figures.

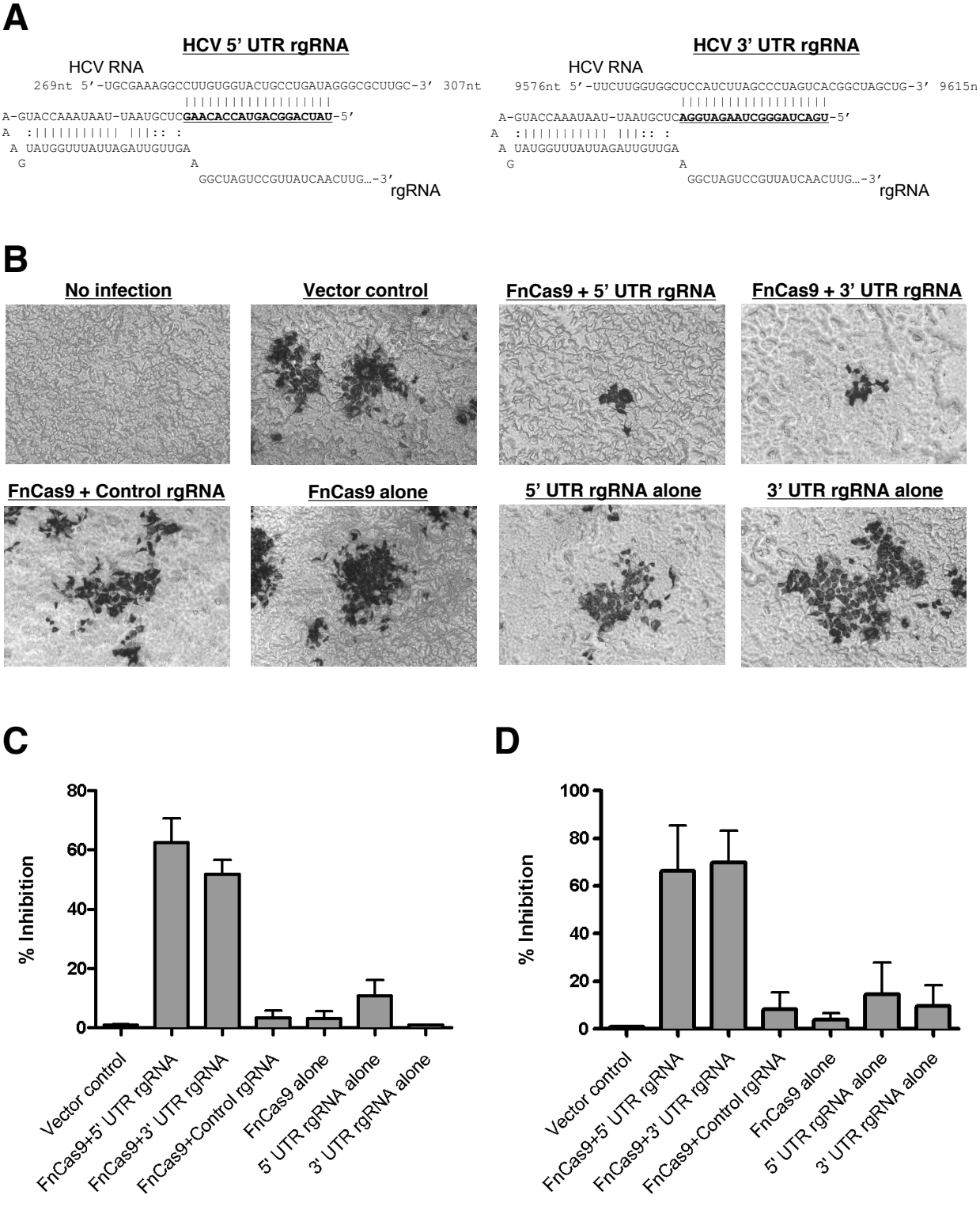

A
HCV 5' UTR rgRNA
HCV RNA
269nt 5'-UGCGAAAGGCCUUGUGGUACUGCCUGAUAGGGCGCUUGC-3' 307nt
A-GUACCAAAUAAU-UAAUGCUCGAACACCAUGACGGACUAU-5'
A UAUGGUUUAUUAGAUUGUUGA
GGCUAGUCCGUUAUCAACUUG…-3'
rgRNA
HCV 3' UTR rgRNA
HCV RNA
9576nt 5'-UUCUUGGUGGCUCCAUCUUAGCCCUAGUCACGGCUAGCUG-3' 9615n
A-GUACCAAAUAAU-UAAUGCUCAGGUAGAAUCGGGAUCAGU-5'
A UAUGGUUUAUUAGAUUGUUGA
GGCUAGUCCGUUAUCAACUUG…-3'
rgRNA
B
No infection
Vector control
FnCas9 + 5' UTR rgRNA
FnCas9 + 3' UTR rgRNA
FnCas9 + Control rgRNA
FnCas9 alone
5' UTR rgRNA alone
3' UTR rgRNA alone
C
% Inhibition
80
60
40
20
0
Vector control
FnCas9+5' UTR rgRNA
FnCas9+3' UTR rgRNA
FnCas9+Control rgRNA
FnCas9 alone
5' UTR rgRNA alone
3' UTR rgRNA alone
D
% Inhibition
100
80
60
40
20
0
Vector control
FnCas9+5' UTR rgRNA
FnCas9+3' UTR rgRNA
FnCas9+Control rgRNA
FnCas9 alone
5' UTR rgRNA alone
3' UTR rgRNA alone

Figure 1. FnCas9 can be reprogrammed to inhibit viral protein production in eukaryotic cells. (A) RNA-targeting guide RNA (rgRNA) schematic with targeting sequences (grey highlight) against the 5' or 3'UTR of HCV genomic RNA. **(B)** Huh-7.5 cells were transfected with the indicated combinations of FnCas9 and rgRNA and infected 48 hours later with HCV encoding *Renilla* luciferase. At 72 hours, cells were fixed and stained with anti-E2 antibody and imaged. **(C)** E2 positive foci from **(B)** were quantified and plotted as percent inhibition compared to the vector control. **(D)** Quantification of viral luciferase production displayed as percent inhibition compared to the vector control. (n=3, bars the SEM, data is representative of at least six experiments)

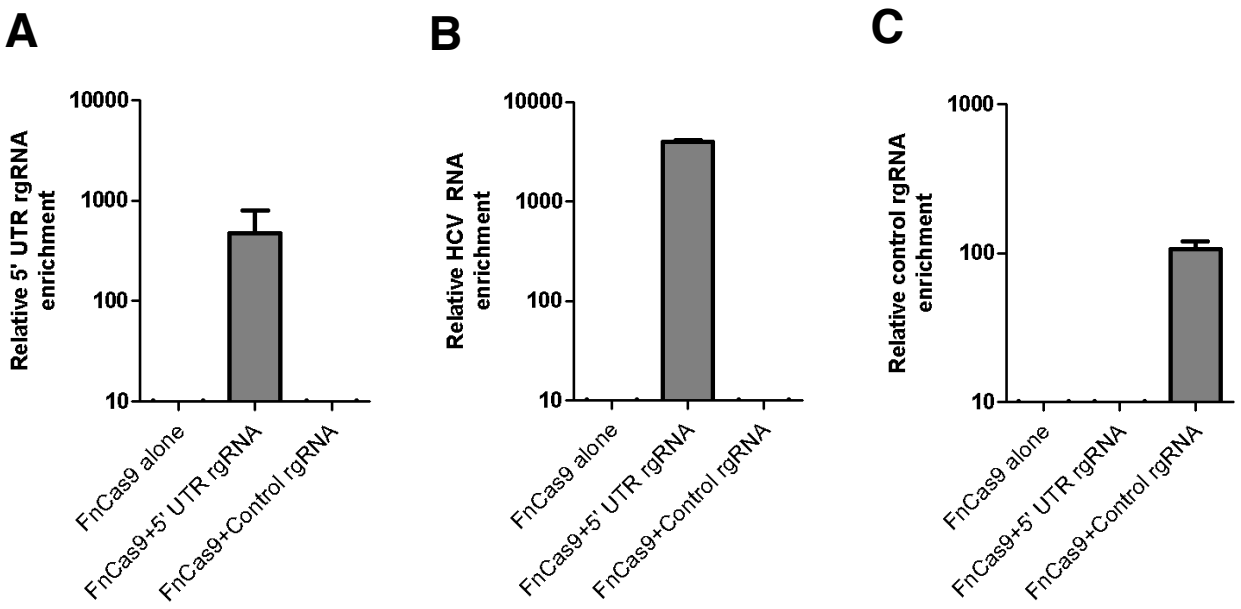


Figure 2. FnCas9 targets and associates with HCV RNA. Huh-7.5 cells producing an HA epitope-tagged FnCas9 alone, or with either the 5'UTR targeting rgRNA or the control rgRNA, were infected with HCV. At 72 hours post infection, lysates were immunoprecipitated with anti-HA. Co-precipitating RNA was purified and analyzed by quantitative real-time PCR to detect the relative enrichment of the **(A)** 5'UTR rgRNA, **(B)** HCV RNA, or **(C)** control rgRNA, normalizing to *gapdh* mRNA levels. (n=4, bars the SEM, data is representative of three experiments)

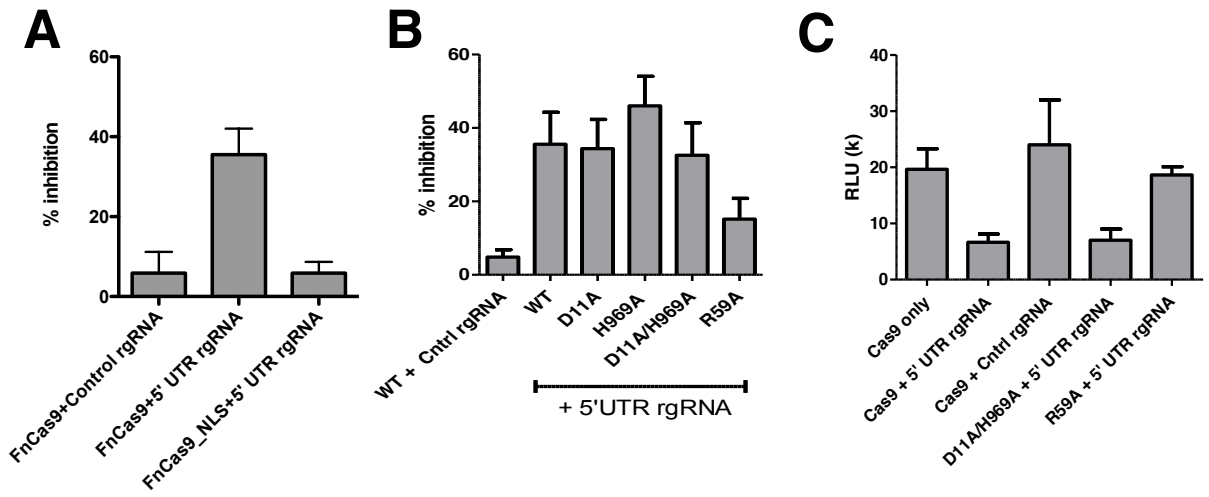


Figure 3. Molecular requirements for FnCas9 mediated HCV inhibition. (A) Huh-7.5 cells were transfected with FnCas9 ± nuclear localization signal (NLS), the 5' UTR targeting rgRNA, and HCV. At 72 hours, viral luciferase was quantified and the percent inhibition compared to the non-targeting rgRNA is displayed (n=8, data compiled from three independent experiments). **(B)** Experiments were performed as above, utilizing alanine point mutants in the RuvC domain (D11A), HNH domain (H969A), or the arginine rich motif (R59A) (n=8, data compiled from three independent experiments). **(C)** Rabbit reticulocyte lysate *in vitro* translation assays of HCV luciferase were performed utilizing the indicated Cas9 and RNAs and viral luciferase measured (n=4, data are representative of four experiments).

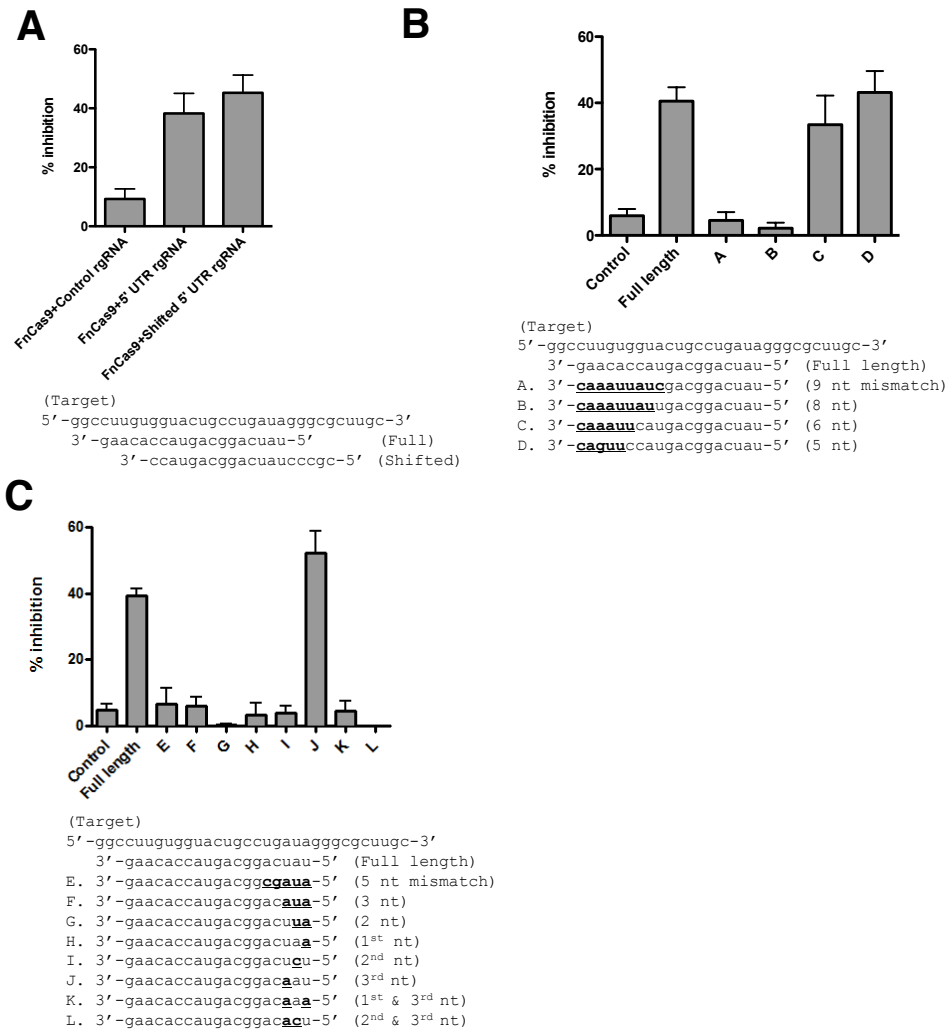


Figure 4. RNA sequence requirements for FnCas9 inhibition of HCV. (A) Huh-7.5 cells were transfected with FnCas9 utilizing the mutants in the indicated shifted alignments below. At 72 hours, viral luciferase was quantified and the percent inhibition compared to the non-targeting rgRNA is displayed (n=12, data is compiled from 3 independent experiments). **(B)** Experiments were performed as above, utilizing the mutants in 3' region indicated in the alignment below. (n = 12, bars the SEM, and data is compiled from 3 independent experiments). **(C)** Experiments were performed as above, utilizing the mutants in the 5' region indicated in the alignment below (n = 12, bars the SEM, and data is compiled from 3 independent experiments).

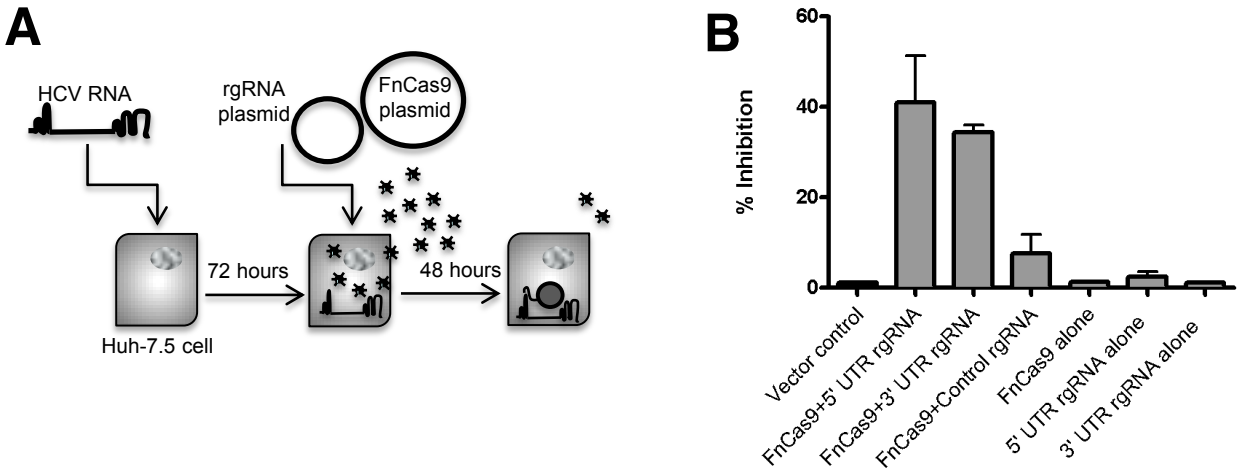


Figure 5. FnCas9 can inhibit an established viral infection. (A) Experimental outline. HCV transfected Huh-7.5 cells were transfected with FnCas9 and the indicated targeting RNAs, after 72 hours post infection **(B)** Quantification of viral luciferase production, displayed as percent inhibition compared to the vector control. (n=3, bars the SEM, data is representative of at least twelve experiments).

Supplementary Information. See publication online:

https://www.pnas.org/content/112/19/6164/tab-figures-data

References

For Chapters 1, 4-6

Due to space constraints and the rapidly expanding field, we have undoubtedly omitted notable work in the field; we sincerely apologize to those authors whose research we did not cite.

1. Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. J Bacteriol. 1987;169(12):5429-33.

2. Ishino Y, Krupovic M, Forterre P. History of CRISPR-Cas from Encounter with a Mysterious Repeated Sequence to Genome Editing Technology. Journal of Bacteriology. 2018;200(7):e00580-17.

3. Nakata A, Amemura M, Makino K. Unusual nucleotide arrangement with repeated sequences in the Escherichia coli K-12 chromosome. Journal of Bacteriology. 1989;171(6):3553-6.

4. Hermans PW, van Soolingen D, Bik EM, de Haas PE, Dale JW, van Embden JD. Insertion element IS987 from Mycobacterium bovis BCG is located in a hot-spot integration region for insertion elements in Mycobacterium tuberculosis complex strains. Infection and immunity. 1991;59(8):2695-705.

5. Groenen PMA, Bunschoten AE, Soolingen Dv, Errtbden JDAv. Nature of DNA polymorphism in the direct repeat cluster of Mycobacterium tuberculosis; application for strain differentiation by a novel typing method. Molecular Microbiology. 1993;10(5):1057-65.

6. Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, et al. Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for diagnosis and epidemiology. Journal of clinical microbiology. 1997;35(4):907-14.

7. Shariat N, Dudley EG. CRISPRs: Molecular Signatures Used for Pathogen Subtyping. Applied and Environmental Microbiology. 2014;80(2):430-9.

8. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology (Reading, England). 2005;151(Pt 8):2551-61.

9. Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. Journal of molecular evolution. 2005;60(2):174-82.

10. Jansen R, Embden JD, Gaastra W, Schouls LM. Identification of genes that are associated with DNA repeats in prokaryotes. Mol Microbiol. 2002;43(6):1565-75.

11. Haft DH, Selengut J, Mongodin EF, Nelson KE. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. PLoS computational biology. 2005;1(6):e60.

12. Makarova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV. 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. Biology direct. 2006;1:7.

13. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR provides acquired resistance against viruses in prokaryotes. Science. 2007;315(5819):1709-12.

14. Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, Snijders AP, et al. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science. 2008;321(5891):960-4.

15. Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(39):E2579-86.

16. Westra ER, van Erp PB, Kunne T, Wong SP, Staals RH, Seegers CL, et al. CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. Mol Cell. 2012;46(5):595-605.

17. Staals RH, Agari Y, Maki-Yonekura S, Zhu Y, Taylor DW, van Duijn E, et al. Structure and activity of the RNA-targeting Type III-B CRISPR-Cas complex of Thermus thermophilus. Mol Cell. 2013;52(1):135-45.

18. Barrangou R, Marraffini LA. CRISPR-Cas systems: Prokaryotes upgrade to adaptive immunity. Mol Cell. 2014;54(2):234-44.

19. Plagens A, Richter H, Charpentier E, Randau L. DNA and RNA interference mechanisms by CRISPR-Cas surveillance complexes. FEMS microbiology reviews. 2015.

20. Rath D, Amlinger L, Rath A, Lundgren M. The CRISPR-Cas immune system: Biology, mechanisms and applications. Biochimie. 2015.

21. Heler R, Marraffini LA, Bikard D. Adapting to new threats: the generation of memory by CRISPR-Cas immune systems. Molecular Microbiology. 2014;93(1):1-9.

22. Marraffini LA, Sontheimer EJ. CRISPR Interference Limits Horizontal Gene Transfer in Staphylococci by Targeting DNA. Science. 2008;322(5909):1843-5.

23. Hale CR, Zhao P, Olson S, Duff MO, Graveley BR, Wells L, et al. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. Cell. 2009;139(5):945-56.

24. Garneau JE, Dupuis ME, Villion M, Romero DA, Barrangou R, Boyaval P, et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature. 2010;468(7320):67-71.

25. Bikard D, Hatoum-Aslan A, Mucida D, Marraffini LA. CRISPR interference can prevent natural transformation and virulence acquisition during in vivo bacterial infection. Cell host & microbe. 2012;12(2):177-86.

26. Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P, et al. Evolution and classification of the CRISPR-Cas systems. Nat Rev Microbiol. 2011;9(6):467-77.

27. Chylinski K, Makarova KS, Charpentier E, Koonin EV. Classification and evolution of type II CRISPR-Cas systems. Nucleic Acids Res. 2014;42(10):6091-105.

28. Krupovic M, Makarova KS, Forterre P, Prangishvili D, Koonin EV. Casposons: a new superfamily of self-synthesizing DNA transposons at the origin of prokaryotic CRISPR-Cas immunity. BMC biology. 2014;12:36.

29. Koonin EV, Krupovic M. Evolution of adaptive immunity from transposable elements combined with innate immune systems. Nature reviews Genetics. 2015;16(3):184-92.

30. Bikard D, Marraffini LA. Control of gene expression by CRISPR-Cas systems. F1000prime reports. 2013;5:47.

31. Barrangou R. The roles of CRISPR-Cas systems in adaptive immunity and beyond. Current opinion in immunology. 2015;32:36-41.

32. Westra ER, Buckling A, Fineran PC. CRISPR-Cas systems: beyond adaptive immunity. Nat Rev Micro. 2014;12(5):317-26.

33. Ratner HK, Sampson TR, Weiss DS. I can see CRISPR now, even when phage are gone: a view on alternative CRISPR-Cas functions from the prokaryotic envelope. Current opinion in infectious diseases. 2015;28(3):267-74.

34. Makarova KS, Wolf YI, Koonin EV. Comparative genomics of defense systems in archaea and bacteria. Nucleic Acids Res. 2013;41(8):4360-77.

35. Vestergaard G, Garrett RA, Shah SA. CRISPR adaptive immune systems of Archaea. RNA Biology. 2014;11(2):156-67.

36. Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, Saunders SJ, et al. An updated evolutionary classification of CRISPR-Cas systems. Nat Rev Micro. 2015;13(11):722-36.

37. Koonin EV, Makarova KS, Zhang F. Diversity, classification and evolution of CRISPR-Cas systems. Current opinion in microbiology. 2017;37:67-78.

38. Amitai G, Sorek R. CRISPR-Cas adaptation: insights into the mechanism of action. Nat Rev Microbiol. 2016;14(2):67-76.

39. Datsenko KA, Pougach K, Tikhonov A, Wanner BL, Severinov K, Semenova E. Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. Nat Commun. 2012;3:945.

40. Yosef I, Goren MG, Qimron U. Proteins and DNA elements essential for the CRISPR adaptation process in Escherichia coli. Nucleic Acids Res. 2012;40(12):5569-76.

41. Nunez JK, Kranzusch PJ, Noeske J, Wright AV, Davies CW, Doudna JA. Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. Nature structural & molecular biology. 2014;21(6):528-34.

42. Nunez JK, Lee AS, Engelman A, Doudna JA. Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. Nature. 2015;519(7542):193-8.

43. Plagens A, Tjaden B, Hagemann A, Randau L, Hensel R. Characterization of the CRISPR/Cas subtype I-A system of the hyperthermophilic crenarchaeon Thermoproteus tenax. J Bacteriol. 2012;194(10):2491-500.

44. Swarts DC, Mosterd C, van Passel MWJ, Brouns SJJ. CRISPR Interference Directs Strand Specific Spacer Acquisition. PloS one. 2012;7(4):e35888.

45. van der Oost J, Westra ER, Jackson RN, Wiedenheft B. Unravelling the structural and mechanistic basis of CRISPR-Cas systems. Nat Rev Micro. 2014;12(7):479-92.

46. Heler R, Samai P, Modell JW, Weiner C, Goldberg GW, Bikard D, et al. Cas9 specifies functional viral targets during CRISPR-Cas adaptation. Nature. 2015;519(7542):199-202.

47. Burstein D, Harrington LB, Strutt SC, Probst AJ, Anantharaman K, Thomas BC, et al. New CRISPR–Cas systems from uncultivated microbes. Nature. 2016;542:237.

48. Ozcan A, Pausch P, Linden A, Wulf A, Schuhle K, Heider J, et al. Type IV CRISPR RNA processing and effector complex formation in Aromatoleum aromaticum. Nat Microbiol. 2019;4(1):89-96.

49. Jore MM, Lundgren M, van Duijn E, Bultema JB, Westra ER, Waghmare SP, et al. Structural basis for CRISPR RNA-guided DNA recognition by Cascade. Nature structural & molecular biology. 2011;18(5):529-36.

50. Gesner EM, Schellenberg MJ, Garside EL, George MM, MacMillan AM. Recognition and maturation of effector RNAs in a CRISPR interference pathway. Nature structural & molecular biology. 2011;18(6):688-92.

51. Garside EL, Schellenberg MJ, Gesner EM, Bonanno JB, Sauder JM, Burley SK, et al. Cas5d processes pre-crRNA and is a member of a larger family of CRISPR RNA endonucleases. RNA. 2012;18(11):2020-8.

52. Sashital DG, Jinek M, Doudna JA. An RNA-induced conformational change required for CRISPR RNA cleavage by the endoribonuclease Cse3. Nature structural & molecular biology. 2011;18(6):680-7.

53. Nam KH, Haitjema C, Liu X, Ding F, Wang H, DeLisa MP, et al. Cas5d protein processes pre-crRNA and assembles into a Cascade-like interference complex in Subtype I-C/Dvulg CRISPR-Cas system. Structure (London, England : 1993). 2012;20(9):1574-84.

54. Koo Y, Ka D, Kim E-J, Suh N, Bae E. Conservation and Variability in the Structure and Function of the Cas5d Endoribonuclease in the CRISPR-Mediated Microbial Immune System. Journal of molecular biology. 2013;425(20):3799-810.

55. Reeks J, Naismith James H, White Malcolm F. CRISPR interference: a structural perspective. Biochemical Journal. 2013;453(Pt 2):155-66.

56. Carte J, Pfister NT, Compton MM, Terns RM, Terns MP. Binding and cleavage of CRISPR RNA by Cas6. Rna. 2010;16(11):2181-8.

57. Carte J, Wang R, Li H, Terns RM, Terns MP. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. Genes & development. 2008;22(24):3489-96.

58. Haurwitz RE, Jinek M, Wiedenheft B, Zhou K, Doudna JA. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. Science (New York, NY). 2010;329(5997):1355-8.

59. Hatoum-Aslan A, Maniv I, Marraffini LA. Mature clustered, regularly interspaced, short palindromic repeats RNA (crRNA) length is measured by a ruler mechanism anchored at the precursor processing site. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(52):21218-22.

60. Haurwitz RE, Sternberg SH, Doudna JA. Csy4 relies on an unusual catalytic dyad to position and cleave CRISPR RNA. The EMBO journal. 2012;31(12):2824-32.

61. Wang R, Preamplume G, Terns MP, Terns RM, Li H. Interaction of the Cas6 riboendonuclease with CRISPR RNAs: recognition and cleavage. Structure (London, England : 1993). 2011;19(2):257-64.

62. Sternberg SH, Haurwitz RE, Doudna JA. Mechanism of substrate selection by a highly specific CRISPR endoribonuclease. RNA. 2012;18(4):661-72.

63. Niewoehner O, Jinek M, Doudna JA. Evolution of CRISPR RNA recognition and processing by Cas6 endonucleases. Nucleic Acids Research. 2014;42(2):1341-53.

64. Wiedenheft B, van Duijn E, Bultema JB, Waghmare SP, Zhou K, Barendregt A, et al. RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(25):10092-7.

65. Wiedenheft B, Lander GC, Zhou K, Jore MM, Brouns SJJ, van der Oost J, et al. Structures of the RNA-guided surveillance complex from a bacterial immune system. Nature. 2011;477(7365):486-9.

66. Lintner NG, Kerou M, Brumfield SK, Graham S, Liu H, Naismith JH, et al. Structural and Functional Characterization of an Archaeal Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR)-associated Complex for Antiviral Defense (CASCADE). The Journal of biological chemistry. 2011;286(24):21643-56.

67. Spilman M, Cocozaki A, Hale C, Shao Y, Ramia N, Terns R, et al. Structure of an RNA Silencing Complex of the CRISPR-Cas immune system. Molecular cell. 2013;52(1):10.1016/j.molcel.2013.09.008.

68. Hrle A, Su AAH, Ebert J, Benda C, Randau L, Conti E. Structure and RNA-binding properties of the Type III-A CRISPR-associated protein Csm3. RNA Biol. 2013;10(11):1670-8.

69. Zhang J, Rouillon C, Kerou M, Reeks J, Brugger K, Graham S, et al. Structure and mechanism of the CMR complex for CRISPR-mediated antiviral immunity. Mol Cell. 2012;45(3):303-13.

70. Rouillon C, Zhou M, Zhang J, Politis A, Beilsten-Edmands V, Cannone G, et al. Structure of the CRISPR interference complex CSM reveals key similarities with cascade. Mol Cell. 2013;52(1):124-34.

71. Staals RH, Zhu Y, Taylor DW, Kornfeld JE, Sharma K, Barendregt A, et al. RNA targeting by the type III-A CRISPR-Cas Csm complex of Thermus thermophilus. Mol Cell. 2014;56(4):518-30.

72. Peng W, Feng M, Feng X, Liang YX, She Q. An archaeal CRISPR type III-B system exhibiting distinctive RNA targeting features and mediating dual RNA and DNA interference. Nucleic Acids Res. 2015;43(1):406-17.

73. Samai P, Pyenson N, Jiang W, Goldberg GW, Hatoum-Aslan A, Marraffini LA. Co-transcriptional DNA and RNA Cleavage during Type III CRISPR-Cas Immunity. Cell. 2015.

74. Zhu Y, Klompe SE, Vlot M, van der Oost J, Staals RHJ. Shooting the messenger: RNA-targetting CRISPR-Cas systems. Biosci Rep. 2018;38(3):BSR20170788.

75. Kazlauskiene M, Kostiuk G, Venclovas Č, Tamulaitis G, Siksnys V. A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. Science. 2017;357(6351):605-9.

76. Niewoehner O, Garcia-Doval C, Rostøl JT, Berk C, Schwede F, Bigler L, et al. Type III CRISPR–Cas systems produce cyclic oligoadenylate second messengers. Nature. 2017;548:543.

77. Wang L, Mo CY, Wasserman MR, Rostøl JT, Marraffini LA, Liu S. Dynamics of Cas10 Govern Discrimination between Self and Non-self in Type III CRISPR-Cas Immunity. Molecular Cell. 2019;73(2):278-90.e4.

78. Deveau H, Barrangou R, Garneau JE, Labonte J, Fremaux C, Boyaval P, et al. Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. J Bacteriol. 2008;190(4):1390-400.

79. Marraffini LA, Sontheimer EJ. Self versus non-self discrimination during CRISPR RNA-directed immunity. Nature. 2010;463(7280):568-71.

80. Goldberg GW, Jiang W, Bikard D, Marraffini LA. Conditional tolerance of temperate phages via transcription-dependent CRISPR-Cas targeting. Nature. 2014;514(7524):633-7.

81. Shmakov SA, Sitnik V, Makarova KS, Wolf YI, Severinov KV, Koonin EV. The CRISPR Spacer Space Is Dominated by Sequences from Species-Specific Mobilomes. mBio. 2017;8(5):e01397-17.

82. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature. 2011;471(7340):602-7.

83. Dugar G, Leenay RT, Eisenbart SK, Bischler T, Aul BU, Beisel CL, et al. CRISPR RNA-Dependent Binding and Cleavage of Endogenous RNAs by the Campylobacter jejuni Cas9. Molecular Cell. 2018;69(5):893-905.e7.

84. Rousseau BA, Hou Z, Gramelspacher MJ, Zhang Y. Programmable RNA Cleavage and Recognition by a Natural CRISPR-Cas9 System from Neisseria meningitidis. Mol Cell. 2018;69(5):906-14.e4.

85. Strutt SC, Torrez RM, Kaya E, Negrete OA, Doudna JA. RNA-dependent RNA targeting by CRISPR-Cas9. eLife. 2018;7.

86. Chylinski K, Le Rhun A, Charpentier E. The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. RNA Biology. 2013;10(5):726-37.

87. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science. 2012;337(6096):816-21.

88. Fonfara I, Le Rhun A, Chylinski K, Makarova KS, Lecrivain AL, Bzdrenga J, et al. Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. Nucleic Acids Res. 2014;42(4):2577-90.

89. Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proceedings of the National Academy of Sciences. 2012;109(39):E2579–E86.

90. Hou Z, Zhang Y, Propson NE, Howden SE, Chu LF, Sontheimer EJ, et al. Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. Proceedings of the National Academy of Sciences of the United States of America. 2013.

91. Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, Ma E, et al. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. Science. 2014;343(6176):1247997.

92. Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell. 2014;156(5):935-49.

93. Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature. 2014;507(7490):62-7.

94. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31(2):166-9.

95. Wei Y, Terns RM, Terns MP. Cas9 function and host genome sampling in Type II-A CRISPR-Cas adaptation. Genes & development. 2015;29(4):356-61.

96. Stern A, Keren L, Wurtzel O, Amitai G, Sorek R. Self-targeting by CRISPR: gene regulation or autoimmunity? Trends in genetics : TIG. 2010;26(8):335-40.

97. Diez-Villasenor C, Guzman NM, Almendros C, Garcia-Martinez J, Mojica FJ. CRISPR-spacer integration reporter plasmids reveal distinct genuine acquisition specificities among CRISPR-Cas I-E variants of Escherichia coli. RNA Biol. 2013;10(5):792-802.

98. Celli J, Zahrt TC. Mechanisms of Francisella tularensis intracellular pathogenesis. Cold Spring Harb Perspect Med. 2013;3(4):a010314-a.

99. Kingry LC, Petersen JM. Comparative review of Francisella tularensis and Francisella novicida. Frontiers in cellular and infection microbiology. 2014;4:35.

100. Petersen JM, Mead PS, Schriefer ME. Francisella tularensis: an arthropod-borne pathogen. Vet Res. 2009;40(2):7-.

101. McCoy GW, Chapin CC. Studies of Plague, a plague-like disease and tuberculosis among roduents in California. Journal of Infectious Diseases. 1912;VI:170-80.

102. Francis E. Tularemia. J Am Med Assoc. 1925;84:1243-50.

103. Oyston PCF. Francisella tularensis: unravelling the secrets of an intracellular pathogen. Journal of Medical Microbiology. 2008;57(8):921-30.

104. Saslaw S, Eigelsbach HT, Prior JA, Wilson HE, Carhart S. Tularemia vaccine study. II. Respiratory challenge. Archives of internal medicine. 1961;107:702-14.

105. Pechous RD, McCarthy TR, Mohapatra NP, Soni S, Penoske RM, Salzman NH, et al. A Francisella tularensis Schu S4 Purine Auxotroph Is Highly Attenuated in Mice but Offers Limited Protection against Homologous Intranasal Challenge. PloS one. 2008;3(6):e2487.

106. Jones CL, Napier BA, Sampson TR, Llewellyn AC, Schroeder MR, Weiss DS. Subversion of Host Recognition and Defense Systems by Francisella spp. Microbiology and Molecular Biology Reviews. 2012;76(2):383-404.

107. Dennis DT, Inglesby TV, Henderson DA, Bartlett JG, Ascher MS, Eitzen E, et al. Tularemia as a biological weapon: medical and public health management. Jama. 2001;285(21):2763-73.

108. Sjostedt A. Tularemia: history, epidemiology, pathogen physiology, and clinical manifestations. Annals of the New York Academy of Sciences. 2007;1105:1-29.

109. Moreau GB, Mann BJ. Adherence and uptake of Francisella into host cells. Virulence. 2013;4(8):826-32.

110. Ozanic M, Marecic V, Abu Kwaik Y, Santic M. The Divergent Intracellular Lifestyle of Francisella tularensis in Evolutionarily Distinct Host Cells. PLoS pathogens. 2015;11(12):e1005208-e.

111. Jaggavarapu S, Weiss DS. Bacterial Prison Break: A Host Protein Mimic Paves the Way. Cell host & microbe. 2018;24(2):189-91.

112. Ledvina HE, Kelly KA, Eshraghi A, Plemel RL, Peterson SB, Lee B, et al. A Phosphatidylinositol 3-Kinase Effector Alters Phagosomal Maturation to Promote Intracellular Growth of Francisella. Cell host & microbe. 2018;24(2):285-95.e8.

113. Eshraghi A, Kim J, Walls AC, Ledvina HE, Miller CN, Ramsey KM, et al. Secreted Effectors Encoded within and outside of the Francisella Pathogenicity Island Promote Intramacrophage Growth. Cell host & microbe. 2016;20(5):573-83.

114. Jones CL, Sampson TR, Nakaya HI, Pulendran B, Weiss DS. Repression of bacterial lipoprotein production by Francisella novicida facilitates evasion of innate immune recognition. Cellular microbiology. 2012;14(10):1531-43.

115. Sampson TR, Napier BA, Schroeder MR, Louwen R, Zhao J, Chin CY, et al. A CRISPR-Cas system enhances envelope integrity mediating antibiotic resistance and inflammasome evasion. Proceedings of the National Academy of Sciences of the United States of America. 2014;111(30):11163-8.

116. Peng K, Broz P, Jones J, Joubert LM, Monack D. Elevated AIM2-mediated pyroptosis triggered by hypercytotoxic Francisella mutant strains is attributed to increased intracellular bacteriolysis. Cellular microbiology. 2011;13(10):1586-600.

117. Weiss DS, Brotcke A, Henry T, Margolis JJ, Chan K, Monack DM. In vivo negative selection screen identifies genes required for Francisella virulence. Proceedings of the National Academy of Sciences of the United States of America. 2007;104(14):6037-42.

118. Sampson TR, Saroj SD, Llewellyn AC, Tzeng Y-L, Weiss DS. A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. Nature. 2013;497(7448):254-7.

119. Schunder E, Rydzewski K, Grunow R, Heuner K. First indication for a functional CRISPR/Cas system in Francisella tularensis. International Journal of Medical Microbiology. 2013;303(2):51-60.

120. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21.

121. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology. 2014;15(12):550.

122. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. Science. 2007;315(5819):1709-12.

123. Staals RHJ, Agari Y, Maki-Yonekura S, Zhu Y, Taylor DW, van Duijn E, et al. Structure and activity of the RNA-targeting Type III-B CRISPR-Cas complex of Thermus thermophilus. Molecular cell. 2013;52(1):135-45.

124. Sternberg SH, Richter H, Charpentier E, Qimron U. Adaptation in CRISPR-Cas Systems. Mol Cell. 2016;61(6):797-808.

125. Barrangou R. Cas9 Targeting and the CRISPR Revolution. Science. 2014;344(6185):707-8.

126. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science. 2014;346(6213):1258096.

127. Ratner HK, Sampson TR, Weiss DS. Overview of CRISPR-Cas9 Biology. In: Doudna JA MP, editor. CRISPR-Cas: A laboratory manual. Cold Spring Harbor, New York: Cold Spring Harbor Press; 2016.

128. Fonfara I, Richter H, Bratovič M, Le Rhun A, Charpentier E. The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. Nature. 2016;532(7600):517-21.

129. Price AA, Sampson TR, Ratner HK, Grakoui A, Weiss DS. Cas9-mediated targeting of viral RNA in eukaryotic cells. Proceedings of the National Academy of Sciences of the United States of America. 2015.

130. Hirano H, Gootenberg JS, Horii T, Abudayyeh OO, Kimura M, Hsu PD, et al. Structure and Engineering of Francisella novicida Cas9. Cell. 2016;164(5):950-61.

131. Vercoe RB, Chang JT, Dy RL, Taylor C, Gristwood T, Clulow JS, et al. Cytotoxic chromosomal targeting by CRISPR/Cas systems can reshape bacterial genomes and expel or remodel pathogenicity islands. PLoS genetics. 2013;9(4):e1003454.

132. Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nature biotechnology. 2013;31(3):233-9.

133. Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology (Reading, England). 2009;155(Pt 3):733-40.

134. Shah SA, Erdmann S, Mojica FJM, Garrett RA. Protospacer recognition motifs: Mixed identities and functional diversity. RNA Biology. 2013;10(5):891-9.

135. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell. 2015;163(3):759-71.

136. Yamano T, Nishimasu H, Zetsche B, Hirano H, Slaymaker IM, Li Y, et al. Crystal Structure of Cpf1 in Complex with Guide RNA and Target DNA. Cell. 2016;165(4):949-62.

137. Selle K, Barrangou R. Harnessing CRISPR-Cas systems for bacterial genome editing. Trends in microbiology. 2015;23(4):225-32.

138. Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA. CRISPR-assisted editing of bacterial genomes. Nature biotechnology. 2013;31(3):233-9.

139. Jiang W, Marraffini LA. CRISPR-Cas: New Tools for Genetic Manipulations from Bacterial Immunity Systems. Annual review of microbiology. 2015;69:209-28.

140. Bikard D, Euler CW, Jiang W, Nussenzweig PM, Goldberg GW, Duportet X, et al. Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. Nature biotechnology. 2014;32(11):1146-50.

141. Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffini LA. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic acids research. 2013;41(15):7429-37.

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

143. Luo ML, Leenay RT, Beisel CL. Current and future prospects for CRISPR-based tools in bacteria. Biotechnology and bioengineering. 2016;113(5):930-43.

144. Luo ML, Mullis AS, Leenay RT, Beisel CL. Repurposing endogenous type I CRISPR-Cas systems for programmable gene repression. Nucleic Acids Research. 2015;43(1):674-81.

145. Li Y, Pan S, Zhang Y, Ren M, Feng M, Peng N, et al. Harnessing Type I and Type III CRISPR-Cas systems for genome editing. Nucleic Acids Research. 2016;44(4):e34-e.

146. Pyne ME, Bruder MR, Moo-Young M, Chung DA, Chou CP. Harnessing heterologous and endogenous CRISPR-Cas machineries for efficient markerless genome editing in Clostridium. Scientific Reports. 2016;6:25666.

147. Oh J-H, van Pijkeren J-P. CRISPR–Cas9-assisted recombineering in Lactobacillus reuteri. Nucleic Acids Research. 2014;42(17):e131-e.

148. Bikard D, Euler CW, Jiang W, Nussenzweig PM, Goldberg GW, Duportet X, et al. Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. Nat Biotech. 2014;32(11):1146-50.

149. Zhang K, Duan X, Wu J. Multigene disruption in undomesticated Bacillus subtilis ATCC 6051a using the CRISPR/Cas9 system. Scientific Reports. 2016;6:27943.

150. Wiktor J, Lesterlin C, Sherratt DJ, Dekker C. CRISPR-mediated control of the bacterial initiation of replication. Nucleic Acids Research. 2016;44(8):3801-10.

151. Selle K, Klaenhammer TR, Barrangou R. CRISPR-based screening of genomic island excision events in bacteria. Proceedings of the National Academy of Sciences of the United States of America. 2015;112(26):8076-81.

152. Yosef I, Manor M, Kiro R, Qimron U. Temperate and lytic bacteriophages programmed to sensitize and kill antibiotic-resistant bacteria. Proceedings of the National Academy of Sciences of the United States of America. 2015;112(23):7267-72.

153. Bikard D, Euler C, Jiang W, Nussenzweig PM, Goldberg GW, Duportet X, et al. Development of sequence-specific antimicrobials based on programmable CRISPR-Cas nucleases. Nature biotechnology. 2014;32(11):1146-50.

154. Citorik RJ, Mimee M, Lu TK. Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. Nature biotechnology. 2014;32(11):1141-5.

155. Bondy-Denomy J, Garcia B, Strum S, Du M, Rollins MF, Hidalgo-Reyes Y, et al. Multiple mechanisms for CRISPR-Cas inhibition by anti-CRISPR proteins. Nature. 2015;526(7571):136-9.

156. Pawluk A, Staals RHJ, Taylor C, Watson BNJ, Saha S, Fineran PC, et al. Inactivation of CRISPR-Cas systems by anti-CRISPR proteins in diverse bacterial species. Nature Microbiology. 2016;1:16085.

157. Bondy-Denomy J, Pawluk A, Maxwell KL, Davidson AR. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. Nature. 2013;493(7432):429-32.

158. Llewellyn AC, Jones CL, Napier BA, Bina JE, Weiss DS. Macrophage Replication Screen Identifies a Novel <italic>Francisella</italic> Hydroperoxide Resistance Protein Involved in Virulence. PloS one. 2011;6(9):e24201.

159. Gallagher LA, McKevitt M, Ramage ER, Manoil C. Genetic Dissection of the Francisella novicida Restriction Barrier. Journal of Bacteriology. 2008;190(23):7830-7.

160. Yao R, Liu D, Jia X, Zheng Y, Liu W, Xiao Y. CRISPR-Cas9/Cas12a biotechnology and application in bacteria. Synthetic and Systems Biotechnology. 2018;3(3):135-49.

161. Swarts DC, Jinek M. Cas9 versus Cas12a/Cpf1: Structure-function comparisons and implications for genome editing. Wiley interdisciplinary reviews RNA. 2018:e1481.

162. Wu WY, Lebbink JHG, Kanaar R, Geijsen N, van der Oost J. Genome editing by natural and engineered CRISPR-associated nucleases. Nature chemical biology. 2018;14(7):642-51.

163. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012;337(6096):816-21.

164. Sampson TR, Saroj SD, Llewellyn AC, Tzeng YL, Weiss DS. A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. Nature. 2013;497(7448):254-7.

165. Ratner HK, Escalera-Maurer A, Le Rhun A, Jaggavarapu S, Wozniak JE, Crispell EK, et al. Catalytically Active Cas9 Mediates Transcriptional Interference to Facilitate Bacterial Virulence 2019.

166. Ma K, Cao Q, Luo S, Wang Z, Liu G, Lu C, et al. cas9 Enhances Bacterial Virulence by Repressing the regR Transcriptional Regulator in Streptococcus agalactiae. Infection and immunity. 2018;86(3).

167. Louwen R, Horst-Kreft D, de Boer AG, van der Graaf L, de Knegt G, Hamersma M, et al. A novel link between Campylobacter jejuni bacteriophage defence, virulence and Guillain-Barre syndrome. European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology. 2013;32(2):207-26.

168. Guzina J, Chen WH, Stankovic T, Djordjevic M, Zdobnov E, Djordjevic M. In silico Analysis Suggests Common Appearance of scaRNAs in Type II Systems and Their Association With Bacterial Virulence. Frontiers in genetics. 2018;9:474.

169. McWhinnie RL, Nano FE. Synthetic promoters functional in Francisella novicida and Escherichia coli. Appl Environ Microbiol. 2014;80(1):226-34.

170. Sternberg SH, LaFrance B, Kaplan M, Doudna JA. Conformational control of DNA target cleavage by CRISPR-Cas9. Nature. 2015;527(7576):110-3.

171. Breinig M, Schweitzer AY, Herianto AM, Revia S, Schaefer L, Wendler L, et al. Multiplexed orthogonal genome editing and transcriptional activation by Cas12a. Nature Methods. 2019;16(1):51-4.

172. Zetsche B, Heidenreich M, Mohanraju P, Fedorova I, Kneppers J, DeGennaro EM, et al. Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. Nature biotechnology. 2017;35(1):31-4.

173. Bozic B, Repac J, Djordjevic M. Endogenous Gene Regulation as a Predicted Main Function of Type I-E CRISPR/Cas System in E. coli. Molecules (Basel, Switzerland). 2019;24(4).

174. Li R, Fang L, Tan S, Yu M, Li X, He S, et al. Type I CRISPR-Cas targets endogenous genes and regulates virulence to evade mammalian host immunity. Cell research. 2016;26(12):1273-87.

175. Heussler GE, Cady KC, Koeppen K, Bhuju S, Stanton BA, O'Toole GA. Clustered Regularly Interspaced Short Palindromic Repeat-Dependent, Biofilm-Specific Death of Pseudomonas aeruginosa Mediated by Increased Expression of Phage-Related Genes. mBio. 2015;6(3):e00129-15.

176. Bryksin AV, Matsumura I. Rational design of a plasmid origin that replicates efficiently in both gram-positive and gram-negative bacteria. PloS one. 2010;5(10):e13244.

177. Ratner HK, Escalera-Maurer A, Le Rhun A, Jaggavarapu S, Wozniak JE, Crispell EK, et al. Catalytically Active Cas9 Mediates Transcriptional Interference to Facilitate Bacterial Virulence. Molecular Cell. 2019;75(3):498-510.e5.

178. van der Heijden J, Reynolds LA, Deng W, Mills A, Scholz R, Imami K, et al. Salmonella Rapidly Regulates Membrane Permeability To Survive Oxidative Stress. mBio. 2016;7(4):e01238-16.

179. Mink SN, Jacobs H, Gotes J, Kasian K, Cheng ZQ. Ethyl gallate, a scavenger of hydrogen peroxide that inhibits lysozyme-induced hydrogen peroxide signaling in vitro, reverses hypotension in canine septic shock. Journal of applied physiology (Bethesda, Md : 1985). 2011;110(2):359-74.

180. Mink SN, Kasian K, Santos Martinez LE, Jacobs H, Bose R, Cheng ZQ, et al. Lysozyme, a mediator of sepsis that produces vasodilation by hydrogen peroxide signaling in an arterial preparation. American journal of physiology Heart and circulatory physiology. 2008;294(4):H1724-35.

181. Savitskaya E, Lopatina A, Medvedeva S, Kapustin M, Shmakov S, Tikhonov A, et al. Dynamics of Escherichia coli type I-E CRISPR spacers over 42 000 years. Molecular ecology. 2017;26(7):2019-26.

182. Sorek R, Lawrence CM, Wiedenheft B. CRISPR-mediated adaptive immune systems in bacteria and archaea. Annu Rev Biochem. 2013;82:237-66.

183. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. Science. 2013;339(6121):823-6.

184. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013;339(6121):819-23.

185. Shepard CW, Finelli L, Alter MJ. Global epidemiology of hepatitis C virus infection. Lancet Infect Dis. 2005;5(9):558-67.

186. Scheel TK, Rice CM. Understanding the hepatitis C virus life cycle paves the way for highly effective therapies. Nat Med. 2013;19(7):837-49.

187. Mateu G, Donis RO, Wakita T, Bukh J, Grakoui A. Intragenotypic JFH1 based recombinant hepatitis C virus produces high levels of infectious particles but causes increased cell death. Virology. 2008;376(2):397-407.

188. Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature. 2014:Epub ahead of print:Jan 31, 2014.

189. Semenova E, Jore MM, Datsenko KA, Semenova A, Westra ER, Wanner B, et al. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(25):10098-103.

190. Zebec Z, Manica A, Zhang J, White MF, Schleper C. CRISPR-mediated targeted mRNA degradation in the archaeon Sulfolobus solfataricus. Nucleic acids research. 2014;42(8):5280-8.

191. O'Connell MR, Oakes BL, Sternberg SH, East-Seletsky A, Kaplan M, Doudna JA. Programmable RNA recognition and cleavage by CRISPR/Cas9. Nature. 2014.

192. Hsu PD, Lander ES, Zhang F. Development and Applications of CRISPR-Cas9 for Genome Engineering. Cell. 2014;157(6):1262-78.

193. Sullivan CS, Ganem D. A virus-encoded inhibitor that blocks RNA interference in mammalian cells. J Virol. 2005;79(12):7371-9.

194. Singh G, Popli S, Hari Y, Malhotra P, Mukherjee S, Bhatnagar RK. Suppression of RNA silencing by Flock house virus B2 protein is mediated through its interaction with the PAZ domain of Dicer. FASEB J. 2009;23(6):1845-57.

195. Chao JA, Lee JH, Chapados BR, Debler EW, Schneemann A, Williamson JR. Dual modes of RNA-silencing suppression by Flock House virus protein B2. Nat Struct Mol Biol. 2005;12(11):952-7.

196. Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nature biotechnology. 2013;31:833-8.

197. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell. 2013;152(5):1173-83.

198. Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffini LA. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic acids research. 2013;41(15):7429-37.

199. Chang N, Sun C, Gao L, Zhu D, Xu X, Zhu X, et al. Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. Cell Res. 2013;23(4):465-72.

200. Gratz SJ, Cummings AM, Nguyen JN, Hamm DC, Donohue LK, Harrison MM, et al. Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease. Genetics. 2013.

201. Perez-Pinera P, Kocak DD, Vockley CM, Adler AF, Kabadi AM, Polstein LR, et al. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. Nat Methods. 2013;10:973-6.