ANNUAL REVIEWS

Annual Review of Genetics

The tracrRNA in CRISPR Biology and Technologies

Chunyu Liao[1] and Chase L. Beisel[1,2]

[1]Helmholtz Institute for RNA-based Infection Research (HIRI), Helmholtz Centre for Infection Research (HZI), 97080 Würzburg, Germany; email: Chase.Beisel@helmholtz-hiri.de

[2]Medical Faculty, University of Würzburg, 97080 Würzburg, Germany

Annu. Rev. Genet. 2021. 55:161–81

First published as a Review in Advance on August 20, 2021

The *Annual Review of Genetics* is online at genet.annualreviews.org

https://doi.org/10.1146/annurev-genet-071719-022559

Keywords

crystal structure, immune defense, sgRNA, RNA processing, RNase III

Abstract

CRISPR-Cas adaptive immune systems in bacteria and archaea utilize short CRISPR RNAs (crRNAs) to guide sequence-specific recognition and clearance of foreign genetic material. Multiple crRNAs are stored together in a compact format called a CRISPR array that is transcribed and processed into the individual crRNAs. While the exact processing mechanisms vary widely, some CRISPR-Cas systems, including those encoding the Cas9 nuclease, rely on a *trans*-activating crRNA (tracrRNA). The tracrRNA was discovered in 2011 and was quickly co-opted to create single-guide RNAs as core components of CRISPR-Cas9 technologies. Since then, further studies have uncovered processes extending beyond the traditional role of tracrRNA in crRNA biogenesis, revealed Cas nucleases besides Cas9 that are dependent on tracrRNAs, and established new applications based on tracrRNA engineering. In this review, we describe the biology of the tracrRNA and how its ongoing characterization has garnered new insights into prokaryotic immune defense and enabled key technological advances.

INTRODUCTION

Clustered regularly interspaced short palindromic repeats (CRISPR) and their CRISPR-associated (Cas) proteins protect bacteria and archaea against invading nucleic acids (3, 70). As the only known adaptive immune systems in prokaryotes, the systems store nucleic acid information in CRISPR arrays as short spacers between direct repeats. To enact immunity, the CRISPR array is transcribed into a long precursor CRISPR RNA (pre-crRNA) that is processed into individual mature CRISPR RNAs (crRNAs). The crRNAs then direct Cas effector nucleases to cleave DNA or RNA sequences that are complementary to the guide portion of the crRNA and flanked by a protospacer adjacent motif (PAM) (for DNA targets) or a protospacer flanking sequence (commonly abbreviated as PFS) lacking extensive complementarity to the crRNA repeat (for RNA targets) (61, 73). Because target selection is dictated by the acquired spacer, CRISPR-Cas systems can specifically clear future infections by the same invader. The spacer sequence can also be readily changed to direct Cas nucleases to other nucleic acid targets with appropriate flanking sequences, opening a range of applications (extensively reviewed in 83). In turn, the ongoing characterization of CRISPR-Cas immune systems has not only advanced our understanding of the ongoing arms race between prokaryotes and mobile genetic elements but also provided numerous benefits to society.

Arguably, the best-studied and mostly widely harnessed CRISPR-Cas systems are type II systems encoding the effector nuclease Cas9. As the first characterized single-effector nuclease, Cas9 was primed to be co-opted for genome editing and other applications (11, 14, 17, 34, 44, 47, 51, 71, 95, 107). However, completing the jump from biology to technology required a fundamental understanding of how the system's crRNAs are generated to guide DNA targeting. The discovery of the *trans*-activating crRNA (tracrRNA) as a core component of crRNA biogenesis became the missing piece that enabled the invention of the single-guide RNA (sgRNA) and the adoption of Cas9 as the centerpiece of CRISPR technologies (22). Accordingly, the discovery of the tracrRNA was central to the awarding of the 2020 Nobel Prize in Chemistry to Dr. Emmanuelle Charpentier and Dr. Jennifer Doudna for their development of CRISPR-Cas9 for genome editing (75). Since the discovery of the tracrRNA and its use in genome editing almost a decade ago, there have been numerous discoveries that have advanced our understanding of its biological functions and expanded the list of systems involving tracrRNAs. There have also been numerous technological advances that have improved and expanded the reach of tracrRNAs in CRISPR technologies. In this review, we describe the role of the tracrRNA in immune defense of CRISPR-Cas9 and beyond, nontraditional functions involving tracrRNAs, and how the tracrRNA has enabled an expanding set of CRISPR technologies.

THE tracrRNA IN *STREPTOCOCCUS PYOGENES*

Understanding the role of the tracrRNA in CRISPR biology first requires a perspective on the remarkable diversity of CRISPR-Cas systems and how each system confers adaptive immunity. As adaptive immune systems, all CRISPR-Cas systems store invader-derived spacers in CRISPR arrays, process individual crRNAs from a transcribed pre-crRNA, and use the crRNAs to guide the Cas nucleases to recognize and clear complementary nucleic acids. However, the commonalities largely end there. The Cas proteins associated with CRISPR-Cas systems vary widely and confer different mechanisms for each step of adaptive immunity. This diversity drove the field to create a hierarchical classification scheme that underwent multiple rounds of revision as new systems were discovered. The current scheme comprises classes 1 and 2, types I–VI, over 30 subtypes (e.g., I-E, II-A), and a scattering of subtype variants (e.g., I-F1, II-C2) (70). Because these proteins enact the different steps of adaptive immunity, many of the associated mechanisms partition with the CRISPR-Cas type and subtype. For example, in the early days of characterizing CRISPR biology,

type I and III systems were shown to rely on their encoded endoribonuclease Cas6 to cleave each repeat as part of crRNA biogenesis (7, 9, 10, 38, 70). This protein is absent in other CRISPR-Cas types, including type II systems, spurring efforts to elucidate how the different groups of systems process a pre-crRNA into individual crRNAs.

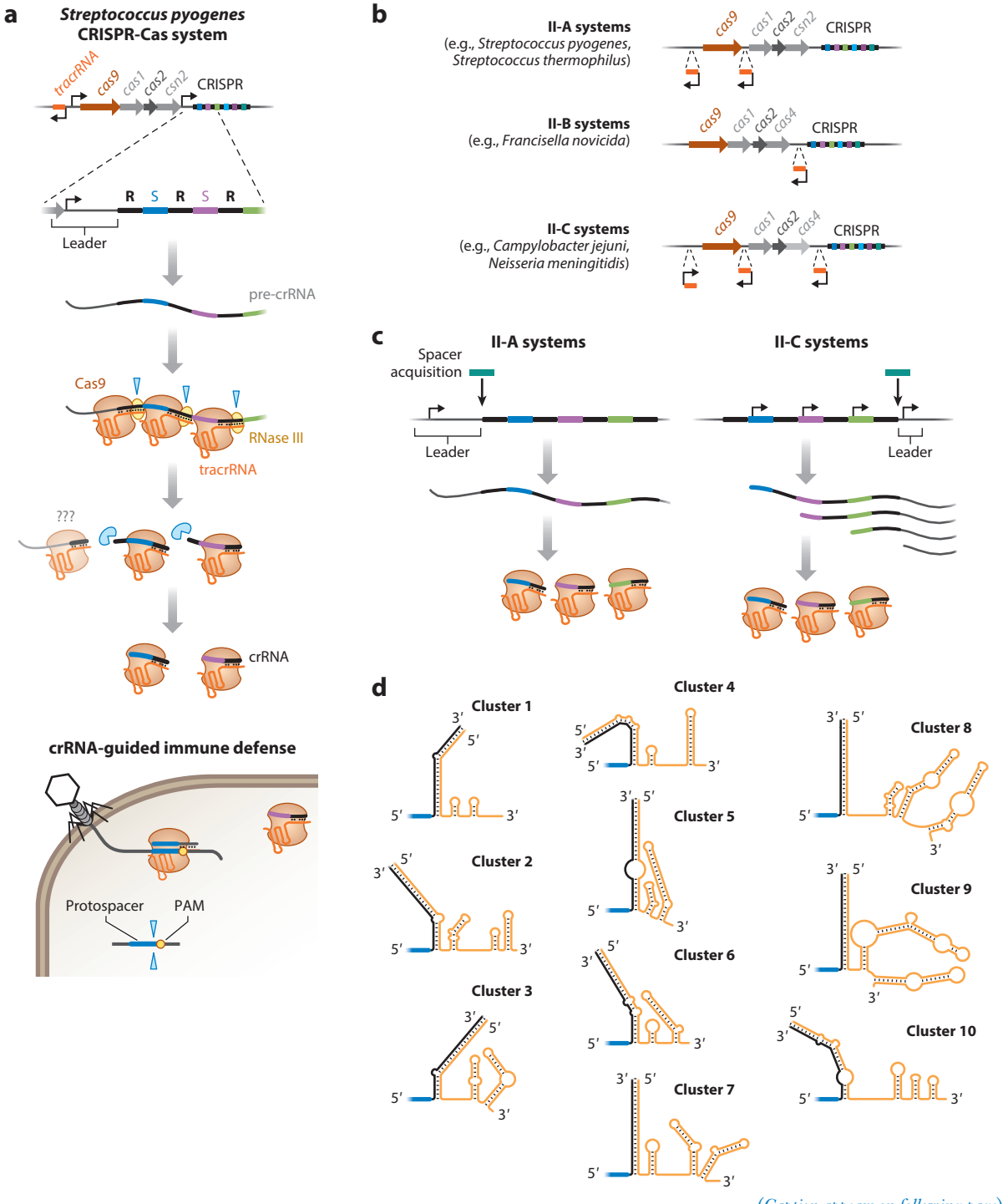
The tracrRNA was first reported in 2011 as part of a seminal study with the human pathogen *Streptococcus pyogenes* and its endogenous type II-A CRISPR-Cas system (22) (**Figure 1*a***). The tracrRNA was identified during a period when next-generation sequencing was first being applied to identify small RNAs missed by standard annotation algorithms (22, 98). One technique, differential RNA sequencing (dRNA-seq) (22, 94, 98), was applied to identify small RNAs and determine whether they possessed a dedicated promoter or were processed from a larger transcript. The sequencing analyses revealed the tracrRNA as one of the most abundant small RNAs in the cell. The RNA was also encoded adjacent and divergent to the *cas9* gene (then called *csn1*), suggesting that it had some role in the flanking CRISPR-Cas system. The tracrRNA further appeared as not one but three different transcripts with a shared 3′ end: ∼171-nucleotide (nt) and ∼89-nt products representing primary transcripts and an ∼75-nt species processed from both longer transcripts. The processing site was within a 24-nt stretch bearing extensive complementarity to the CRISPR repeats, suggesting that the tracrRNA is directly involved in crRNA biogenesis.

The ensuing experiments conducted in *S. pyogenes* informed our primary view of how the tracrRNA participates in crRNA biogenesis (22) (**Figure 1*a***). The CRISPR array is transcribed into a pre-crRNA containing multiple repeats. Each repeat then base-pairs with the 24-nt stretch of the tracrRNA termed the antirepeat domain. The AT-rich sequence within the repeat:antirepeat duplex presents a preferred substrate for the host endoribonuclease RNase III (78), where cleavage forms a 2-nt overhang on the 3′ end of the repeat. The cleavage event also divides the pre-crRNA into individual immature crRNAs consisting of a full-length spacer flanked by either half of the repeat. The RNA duplex is bound by Cas9, while the 5′ end of the immature crRNA is trimmed ∼10 nt into the spacer. The trimming is presumably conducted by host ribonucleases, although these enzymes remain to be identified. The exact order of events (e.g., duplex formation, Cas9 binding, RNase III cleavage) is also unclear, although in vitro experiments with the homologous II-A system from *Streptococcus thermophilus* reported that Cas9 promoted pre-crRNA:tracrRNA hybridization but did not exhibit measurable binding to the tracrRNA (55). Regardless of the exact mechanisms, the resulting ribonucleoprotein effector complex comprises Cas9 bound to an ∼40-nt mature crRNA and an ∼75-nt processed tracrRNA. The duplex of the mature crRNA and processed tracrRNA can be fused with a short tetraloop to create the sgRNA, which simplifies the characterization and implementation of Cas9 nucleases (50).

THE tracrRNA OUTSIDE OF *STREPTOCOCCUS PYOGENES*

Subsequent efforts to characterize tracrRNAs and crRNA biogenesis upheld these basic principles while also revealing variations. In line with the original work, the tracrRNA is considered a ubiquitous feature of CRISPR-Cas systems in all three subtypes of type II systems (II-A, II-B, II-C) and essential for crRNA biogenesis (15, 33, 40, 59). Various bioinformatics search tools have also been developed for identifying tracrRNAs (see the sidebar titled Bioinformatic Identification of tracrRNAs). The variations have been principally associated with the genomic location of the tracrRNA within the system, how crRNAs are generated, and the structures of the crRNA:tracrRNA complex.

The identification of tracrRNAs has revealed different encoding locations within each CRISPR-Cas system that depend on the subtype classification (**Figure 1*b***). For example, reported type II-A systems can encode the tracrRNA gene upstream or downstream of *cas9*, while



(*Caption appears on following page*)

Figure 1 (*Figure appears on preceding page*)

The tracrRNA in crRNA biogenesis and immune defense by type II CRISPR-Cas systems. (*a*) Genomic architecture and crRNA biogenesis pathway for the type II-A CRISPR-Cas system from *Streptococcus pyogenes*. The CRISPR array comprising conserved repeats (R, *black rectangles*) with intervening invader-derived spacers (S, *colored rectangles*) is transcribed into a long pre-crRNA. The tracrRNA then base-pairs with each repeat, driving cleavage by RNase III and binding by Cas9. The processed crRNA:tracrRNA duplex then directs bound Cas9 to DNA sequences matching the crRNA guide sequence (protospacer) and flanked by a PAM. Blue triangles represent processing/cleavage sites. A few steps in crRNA biogenesis remain unknown, including how the 5′ end of each crRNA is trimmed and the fate of the first repeat in the array, which lacks a corresponding spacer. (*b*) tracrRNA location within representative CRISPR-Cas systems from the three type II subtypes. Each system encodes one tracrRNA, although the tracrRNA location can vary between systems in each subtype. (*c*) Comparing crRNA biogenesis for II-A and II-C CRISPR-Cas systems. Type II-A systems initiate the transcription of the array through an upstream promoter encoded within the leader region. Type II-C systems initiate the transcription of the array within each spacer through promoter elements encoded in each repeat. The two systems also differ in the side of the array to which new spacers are added. The outside repeat is copied as part of acquisition to maintain the pattern of the array. (*d*) Predicted secondary structures capturing the vast majority of identified tracrRNAs. The structures were grouped into 10 clusters identified in Reference 25. Abbreviations: Cas, CRISPR-associated; CRISPR, clustered regularly interspaced short palindromic repeats; crRNA, CRISPR RNA; PAM, protospacer adjacent motif; pre-crRNA, precursor CRISPR RNA; R, repeat; S, spacer; tracrRNA, *trans*-activating crRNA.

the reported type II-C systems can also encode the tracrRNA gene upstream or downstream of *cas9* as well as adjacent to the CRISPR array. The one published example of a II-B system, which is found in *Francisella novicida* (1), has the tracrRNA gene located adjacent to the CRISPR array. Although tracrRNAs have been extensively identified, a systematic analysis remains to be performed to determine the diversity and frequency of tracrRNA gene locations and the extent to which the genomic context of the tracrRNA gene contributes to immune function.

While the tracrRNA consistently base-pairs with each repeat within the pre-crRNA, transcription of the crRNA and acquisition of new spacers also vary with the subtype classification (**Figure 1*c***). Type II-A systems initiate transcription from the leader region and proceed through the entire CRISPR array, generating one pre-crRNA transcript. In contrast, characterized type II-C systems contain −10 promoter elements near the 3′ end of each repeat (26, 118). Transcription therefore initiates within each downstream spacer, creating different lengths of pre-crRNA transcripts. RNase III processing was also shown to be dispensable for immune defense by the II-C system in *Neisseria meningitidis*—at least through the last spacer in the array (118). Spacer

BIOINFORMATIC IDENTIFICATION OF tracrRNAS

Predicting *trans*-activating CRISPR RNAs (tracrRNAs) is important for the characterization of new Cas9 proteins and their development as genome editing tools (33, 56). As the tracrRNA can be distinct among different CRISPR systems in terms of sequence, secondary structure, and placement within the CRISPR-Cas loci (70, 114), its computational identification is not trivial. Current prediction methods often consist of first identifying Cas9 through protein homology followed by searching for nearby CRISPR arrays. The CRISPR repeat sequence is then used to search for antirepeat sequences in the vicinity, while the end of the tracrRNA is determined by scanning for a downstream rho-independent terminator (6, 16, 25). These steps can be achieved using individual online or local tools (6, 15), although there are also bundled algorithm codes that allow for batch analyses (16, 25). It is recommended that the identified tracrRNA is confirmed through RNA-seq analyses either in the native microorganism or through heterologous expression in a lab strain. These methods can be modified for predicting tracrRNAs for the type V CRISPR-Cas systems possessing tracrRNAs, although more adjustments would be needed, particularly for predicting the V-C and V-D short-complementarity untranslated RNAs (scoutRNAs) given their lack of extensive complementarity to the CRISPR RNA (crRNA) repeat.

acquisition also appears to occur on opposite sides of the array (**Figure 1*c***), where Cas9 and the tracrRNA are involved in this poorly understood process (39, 41, 108, 118). The molecular details of spacer acquisition and crRNA biogenesis remain effectively unexplored within the remaining II-B subtype (1).

Beyond acquisition and transcription initiation, the structure of the crRNA:tracrRNA complex varies widely across type II systems. One comprehensive bioinformatics search for tracrRNAs predicted 10 main groups based on predicted secondary structures (25) (**Figure 1*d***). These assigned groups could be differentiated by the bulge between the RNA duplex as well as by the secondary structure downstream of the tracrRNA antirepeat.

THE tracrRNA IN CAS9 RECOGNITION

Crystal structures of *S. pyogenes* Cas9 (SpyCas9) have proven invaluable for understanding the molecular details of how Cas9 specifically recognizes the crRNA:tracrRNA duplex (2, 43, 45, 46, 52, 80, 120). The processed crRNA:tracrRNA complex bound by SpyCas9 can be divided into seven domains that can generally be found in other crRNA:tracrRNA complexes. The guide, upper stem, bulge, and lower stem domains form the repeat:antirepeat duplex, while three domains comprising a stem-loop termed the nexus (5), linker, and two additional stem-loops form the tracrRNA tail (**Figure 2*a***). SpyCas9 forms numerous contacts with all domains except the upper stem, freeing the upper stem for unhindered cleavage by RNase III (**Figure 2*b***). These interactions are principally based on RNA secondary structure through recognition of the sugar-phosphate backbone, although base-specific interactions occur through the bulge and the nexus as well as the bottom of the lower stem, the second stem-loop, and the loop region of the third stem-loop. Both lobes of SpyCas9 are also involved in binding the crRNA:tracrRNA complex, with the repeat:antirepeat duplex bound by the recognition (REC) lobe, the last two stem-loops bound by the nuclease (NUC) lobe, and the nexus and linker bound by both lobes. Both structural and sequence-specific recognition allow Cas9 to differentiate this complex from the vast array of other RNAs present in a cell. Upon sgRNA binding, the REC lobe of Cas9 undergoes substantial conformational rearrangement typified by the formation of a central channel for target DNA

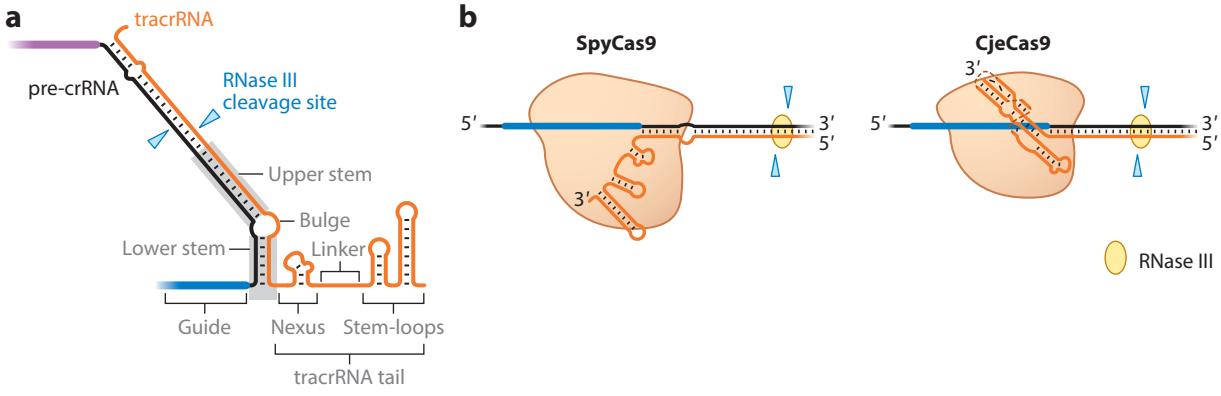


Figure 2

Recognition of the crRNA:tracrRNA duplex by Cas9. (*a*) Anatomy of the crRNA:tracrRNA complex. (*b*) Structure and processing of the crRNA:tracrRNA complex bound to Cas9 from the type II-A system in *Streptococcus pyogenes* (*left*) and the type II-C system in *Campylobacter jejuni* (*right*). The crRNA:tracrRNA duplex undergoes processing by RNase III (*yellow oval*). Abbreviations: Cas, CRISPR-associated; CjeCas9, *C. jejuni* Cas9; CRISPR, clustered regularly interspaced short palindromic repeats; crRNA, CRISPR RNA; pre-crRNA, precursor CRISPR RNA; SpyCas9, *S. pyogenes* Cas9; tracrRNA, *trans*-activating CRISPR RNA.

binding (52). This rearrangement represents the first of many for SpyCas9 to proceed from a holoenzyme to a ribonucleoprotein complex bound to cleaved DNA.

Besides SpyCas9, a number of other Cas9 nucleases have been crystallized with a bound sgRNA that highlights the shared and distinct features across these proteins. Within II-A systems, the small Cas9 from *Streptococcus aureus* (SauCas9) represents one of the few additional Cas9 nucleases that have been crystallized. Contrasting with SpyCas9, the REC lobe and a distinct wedge (WED) domain present in the NUC lobe interact with the repeat:antirepeat duplex nonspecifically through its sugar-phosphate backbone. The nexus is recognized by the REC lobe, similar to its recognition in SpyCas9, while the second and only other stem-loop in the tracrRNA tail is dispensable for the crystallization of the ribonucleoprotein complex. Structural studies of Cas9 proteins from II-B and II-C systems revealed notable structural differences in their sgRNA scaffolds as well as diverse features of domain interaction and arrangement (40, 102, 113). Most notably, the Cas9 from the II-C system in *Campylobacter jejuni* (CjeCas9) recognizes a complicated structure in the tracrRNA tail comprising a pseudoknot and triple helix (113) (**Figure 2*b***). Given the distinct modes of recognition, it is not surprising that phylogenetically distant Cas9 nucleases cannot utilize each other's sgRNAs for DNA targeting, even when the cognate PAM is present (5, 29). The tracrRNA tail also appears to be the primary determinant of specificity, as swapping these tails allowed recognition of the sgRNA by the noncognate Cas9 (5).

THE tracrRNA BEYOND crRNA BIOGENESIS

Type II CRISPR-Cas systems consistently use the crRNA:tracrRNA complex to recognize and cleave double-stranded DNA (dsDNA) as part of adaptive immunity. However, the number of examples of Cas9 nucleases performing other functions with the help of the tracrRNA is growing. One established example involves some Cas9 nucleases recognizing and cleaving RNA (**Figure 3*a***). Recognition occurs through pairing between the RNA and the guide sequence in the duplexed crRNA:tracrRNA, although the rules for recognition and cleavage vary. For instance, RNA binding and cleavage by SpyCas9 required a flanking PAM presented as dsDNA or a DNA:RNA duplex. While absent in natural RNAs, the duplex could be artificially created by annealing a PAM-presenting single-stranded DNA (ssDNA), known as a PAMmer (81). Later work showed that deactivating the HNH and RuvC endonuclease domains in SpyCas9 allowed RNA binding even in the absence of a PAM or PAMmer (4, 77). The Cas9 nucleases from *N. meningitidis* (NmeCas9) as well as SauCas9 and CjeCas9 were also reported to bind and cleave targeted RNA (27, 89, 101), but without a PAM or PAMmer. Evidence also exists that the Cas9 from *F. novicida* (FnoCas9) can also bind and cleave RNA (85). The study with CjeCas9 in particular provided evidence for RNA targeting through the endogenous type II-C CRISPR-Cas system (27). By identifying RNAs bound by the endogenous Cas9 in *C. jejuni*, the authors showed that the CRISPR-Cas system's crRNAs were directing the binding and cleavage of many cellular transcripts. It remains to be seen whether targeting of endogenous transcripts by Cas9 represents a biological function or a background activity tolerated by the cells.

Whether targeting RNA or DNA, Cas9 utilizes a crRNA derived from the CRISPR array. However, prior work showed that a separate RNA encoded in the vicinity of the CRISPR array can also direct Cas9. This small CRISPR-Cas-associated RNA (scaRNA) was identified in the intracellular pathogen *F. novicida* as an essential component of Cas9-mediated repression of genes related to immune avoidance (92) (**Figure 3*b***). The ~58-nt scaRNA is encoded immediately upstream of the CRISPR array and resembles a crRNA, with a guide sequence and repeat-like sequence complementary to the tracrRNA antirepeat. To exert its regulatory function, the scaRNA hybridizes with the tracrRNA, and the scaRNA:tracrRNA complex is bound by Cas9.

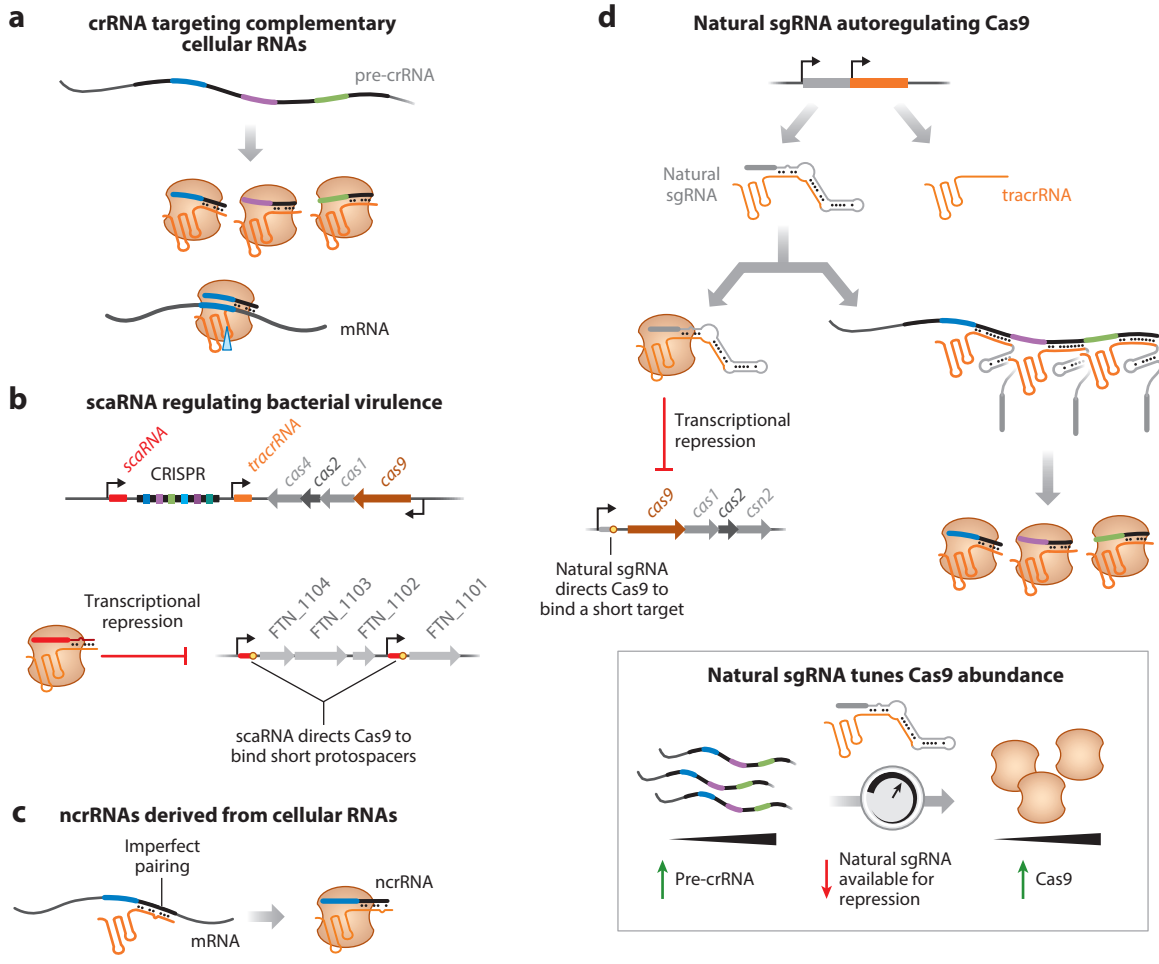


Figure 3

Alternative roles of the tracrRNA in Cas9 biology. (*a*) The crRNA processed with the tracrRNA can direct binding and cleavage of cellular RNAs. These RNAs encode sequences with extensive complementarity to the crRNA guide but do not require the presence of a PAM. At least one Cas9 (i.e., SpyCas9) binds complementary RNAs but does not cleave the RNA. (*b*) A scaRNA downregulates target genes to promote bacterial virulence. The scaRNA is encoded adjacent to the CRISPR array and hybridizes with the tracrRNA, similar to a mature crRNA. The RNA duplex then directs Cas9 to bind shorter PAM-flanked protospacers located downstream of two different promoters. Downregulating these genes promotes virulence in the human pathogen *Francisella novicida*. (*c*) Generation of ncrRNAs from cellular RNAs. The RNAs contain a sequence resembling the CRISPR repeat, promoting hybridization to the antirepeat portion of the tracrRNA and recognition of the RNA duplex by Cas9. The RNA duplex can then direct DNA targeting. (*d*) A long form of the tracrRNA directs autoregulation of Cas9 expression. The long form folds into a natural sgRNA. The natural sgRNA forms a complex with Cas9 and directs the nuclease to bind a PAM-flanked protospacer downstream of the promoter driving expression of the *cas* operon. The protospacer is sufficiently short to allow DNA binding but not cleavage, while binding blocks transcription of the operon. Alternatively, the natural sgRNA preferentially hybridizes to a repeat within the pre-crRNA. The long form of the tracrRNA then undergoes processing by RNase III to generate a mature crRNA. The natural sgRNA therefore serves as a feedback controller by tuning Cas9 levels to the abundance and length of the pre-crRNA. Abbreviations: Cas, CRISPR-associated; CRISPR, clustered regularly interspaced short palindromic repeats; crRNA, CRISPR RNA; mRNA, messenger RNA; ncrRNA, noncanonical CRISPR RNA; PAM, protospacer adjacent motif; pre-crRNA, precursor CRISPR RNA; scaRNA, small CRISPR-Cas-associated RNA; sgRNA, single-guide RNA; SpyCas9, *Streptococcus pyogenes* Cas 9; tracrRNA, *trans*-activating CRISPR RNA.

The scaRNA then directs Cas9 to two DNA targets with ~15 nt complementary to the scaRNA guide and flanked by a recognized PAM in the *F. novicida* genome (86). The limited base-pairing allows DNA binding but not cleavage, resulting in the transcriptional repression of the four downstream virulence genes. Given that a scaRNA would be hard to distinguish from the first repeat in a CRISPR array, more work will be needed to determine the prevalence of scaRNAs and the roles they play in microbial physiology.

Beyond CRISPR arrays and the scaRNA, recent work demonstrated that Cas9 can be guided by cellular RNAs encoded outside of CRISPR-Cas loci. These noncanonical crRNAs (ncrRNAs) were identified in *C. jejuni* as a subset of transcript fragments bound by Cas9, similar to the RNA targets of the crRNAs (27). Unlike the crRNA-targeted RNAs, though, the transcript fragments shared a motif complementary to the tracrRNA antirepeat. Further interrogation revealed that these ncrRNAs were derived through base-pairing with the tracrRNA and subsequent processing, presumably by host RNases (**Figure 3*c***). The extent of base-pairing was often imperfect, although it was sufficient to mediate tracrRNA binding and recognition of the complex by Cas9. In a few cases, combining the ncrRNA-encoding mRNA and the tracrRNA drove sequence-specific cleavage of the corresponding DNA target by Cas9. While no phenotype was associated with the assessed ncrRNAs in *C. jejuni*, the finding shows that Cas9 can be guided by RNAs unrelated to CRISPR-Cas systems. It also shows the potential impact of base-pairing between a cellular transcript and the tracrRNA antirepeat—for better or for worse—that could influence the horizontal transfer and evolution of CRISPR-Cas systems utilizing tracrRNAs.

In addition to interacting with scaRNAs and ncrRNAs to direct Cas9, the tracrRNA was recently shown to regulate Cas9 expression (110) (**Figure 3*d***). This discovery came from exploring the longer version of the tracrRNA in *S. pyogenes* (22). Both the shorter (89-nt) and longer (171-nt) versions encode the antirepeat that duplexes with a pre-crRNA repeat and undergo processing by RNase III to yield a mature crRNA. However, the extended 5′ end allows the longer version to fold into a structure mimicking an sgRNA, albeit with a 13-nt guide sequence, an imperfect stem that does not undergo processing by RNase III, and a large structure protruding from the stem. Despite these differences, this natural sgRNA binds Cas9 and directs it to recognize an 11-bp sequence flanked by a PAM in Cas9's own promoter. The limited pairing allows DNA binding but not cleavage, resulting in transcriptional repression, similar to the scaRNA. Because higher Cas9 levels would lead to stronger repression, the longer version of the tracrRNA provides feedback control of Cas9 levels. Furthermore, by either forming the natural sgRNA or interacting with a pre-crRNA repeat, the longer version of the tracrRNA adjusts Cas9 levels based on the abundance and length of the pre-crRNA. The feedback loop appeared to mitigate autoimmunity by preventing the acquisition of self-targeting spacers caused by overexpression of the *cas* genes. The natural sgRNA was predicted in almost half of assessed II-A systems.

THE tracrRNA BEYOND CAS9

While the tracrRNA was originally considered a unique feature of type II CRISPR-Cas systems with a Cas9 protein, efforts to expand the known repertoire of CRISPR-Cas systems revealed type V systems requiring tracrRNAs (70) (**Figure 4*a***). Type V systems encode Cas12 effector nucleases that, similar to Cas9, tend to target dsDNA. However, unlike Cas9, Cas12 nucleases create a 5′ overhang as part of target DNA cleavage and drive collateral degradation of ssDNA upon target recognition (13, 117). The mature crRNA also comprises a repeat preceding the guide sequence. While all type II systems rely on a tracrRNA, some Cas nucleases, like Cas12a or Cas12i, process the crRNA repeat without accessory factors (30, 114). Of the type V systems with tracrRNAs, the first reported example came from a bioinformatics search for novel class 2 CRISPR-Cas systems (8, 96). Characterization of one system later classified as type V-B revealed a

a

Subtype/ variant	Nuclease	tracrRNA?
V-A	Cpf1/Cas12a	N
V-B	C2c1/Cas12b	Y
V-C	C2c3/Cas12c	Y
V-D	CasY/Cas12d	Y
V-E	CasX/Cas12e	Y
V-F1	Cas14a/Cas12f1	Y
V-F2	Cas14b/Cas12f2	NA
V-F3	Cas14c/Cas12f3	NA
V-G	Cas12g	Y
V-H	Cas12h	N
V-I	Cas12i	N
V-K	Cas12k	Y
V-U1	C2c4	NA
V-U2	C2c8	NA
V-U4	C2c9	NA

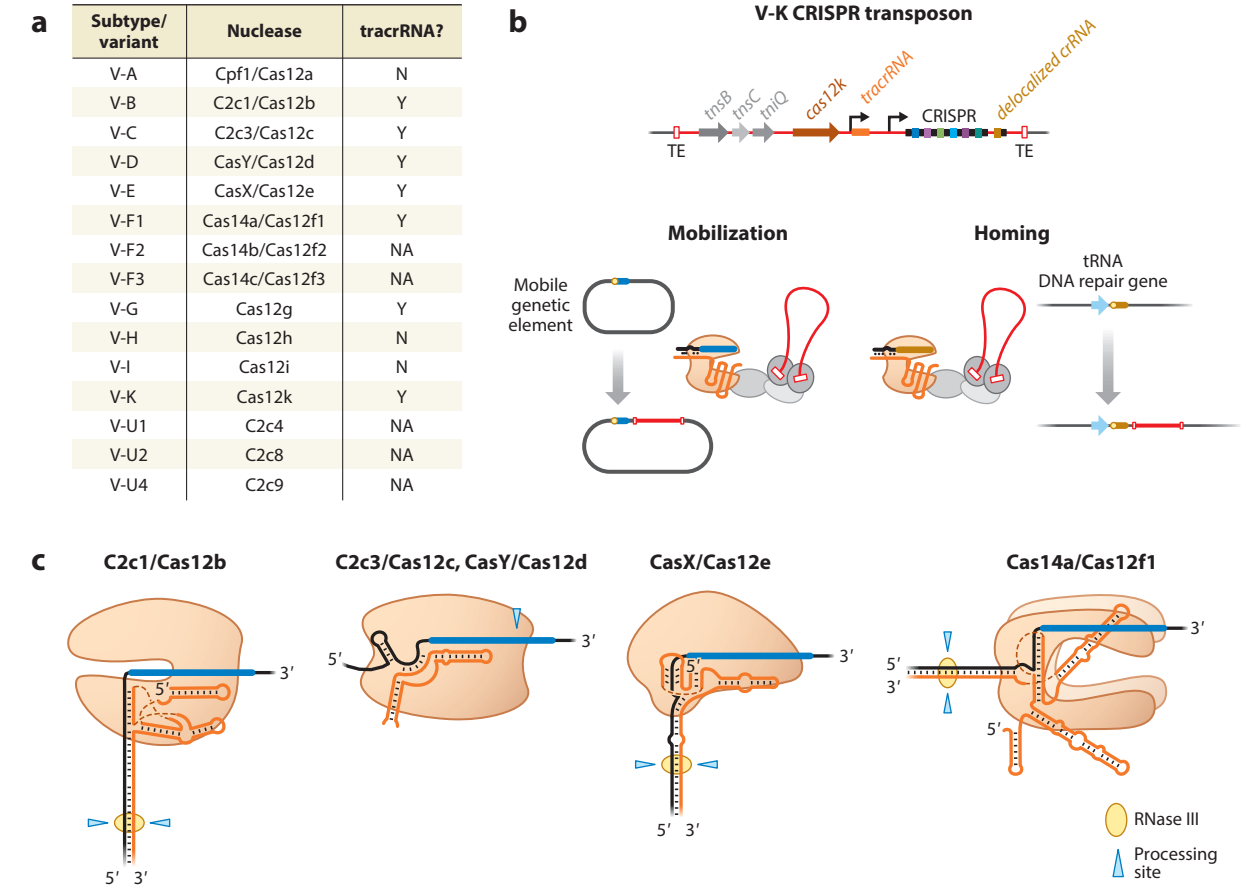


Figure 4

tracrRNAs associated with type V CRISPR-Cas systems. (*a*) A list of currently classified type V subtypes and variants, along with the name of the effector nuclease and whether the system does (indicated by Y) or does not (indicated by N) include a tracrRNA. NA indicates systems where no tracrRNAs are predicted, but the systems have not undergone experimental characterization. (*b*) tracrRNA-dependent mobilization and homing by type V-K CRISPR transposons. crRNAs derived from the CRISPR array direct mobilization through insertion into mobile genetic elements, while a delocalized crRNA directs insertion into a conserved chromosomal site flanking tRNA or DNA repair genes. (*c*) Structure and processing for crRNA:tracrRNA complexes bound to the cognate Cas12 nuclease. The structures associated with Cas12b, Cas12e, and Cas12f1 are based on crystal structures, while the structure for Cas12c/Cas12d is based on folding predictions and experimental evidence. Cas12f1 forms a dimer that binds one crRNA:tracrRNA complex. The crRNA:tracrRNA duplex associated with Cas12b, Cas12e, and Cas12f1 is presumed to undergo processing by RNase III (*yellow oval*). Processing of the crRNA guide by Cas12c/Cas12d is established for Cas12c but remains circumstantial for Cas12d. A structure for Cas12g bound to a crRNA:tracrRNA duplex was reported after acceptance of this review (63). Abbreviations: Cas, CRISPR-associated; CRISPR, clustered regularly interspaced short palindromic repeats; crRNA, CRISPR RNA; TE, transposon end; tracrRNA, *trans*-activating CRISPR RNA; tRNA, transfer RNA.

highly abundant small RNA encoded between *cas2* and the CRISPR array. The 3′ end of the small RNA exhibited extensive complementarity to the CRISPR repeat, underwent processing within the repeat:antirepeat duplex, and was indispensable for DNA targeting by the C2c1/Cas12b nuclease. This study also identified C2c3/Cas12c from type V-C systems but without a tracrRNA. Later work showed that a tracrRNA was present and required for DNA targeting, although limited complementarity between the crRNA repeat and tracrRNA antirepeat prevented initial

annotation (37, 114). Similar issues prevented the identification of tracrRNAs in type V-D systems, although later work showed that CasY/Cas12d relied on a tracrRNA for DNA targeting (37). Separate searches for novel class 2 nucleases within metagenomic sequences and uncultivated microorganisms identified the smaller CRISPR nuclease CasX/Cas12e representing type V-E systems and the smallest type V nuclease Cas14a/Cas12f1 representing type V-F1 systems, and both required a tracrRNA for DNA targeting (36, 67). Cas12f1 nucleases preferentially cleave ssDNA but can also weakly cleave dsDNA flanked by a PAM (36, 54). A broader investigation of Cas12 nucleases revealed the tracrRNA-dependent Cas12g nuclease from type V-G systems, the only type V subtype to date shown to target RNA (114). Other subtypes and variants within type V systems have been identified through bioinformatics searches but remain uncharacterized experimentally, leaving open the possibility that additional type V systems encode tracrRNAs (70).

Beyond the Cas12 nucleases associated with CRISPR-Cas immune systems, a separate tracrRNA-dependent Cas12 nuclease called Cas12k forms part of the newly discovered mobile genetic elements termed CRISPR-associated transposons (CASTs) (58, 82, 100). Cas12k interacts with three Tn7-like transposon genes (*tnsB*, *tnsC*, *tniQ*) to mediate crRNA-directed transposition of the CAST. The crRNAs can derive from not only an encoded CRISPR array but also a delocalized crRNA often positioned downstream of the CRISPR array (90) (**Figure 4*b***). The repeat portion of the delocalized crRNA heavily deviates from the consensus repeat in the CRISPR array, although both repeats can direct DNA transposition with the help of the tracrRNA. Intriguingly, the CRISPR array and delocalized crRNA appear to fulfill distinct biological roles, with the CRISPR array driving mobilization via integration into mobile genetic elements and the delocalized crRNA driving homing via integration next to transfer RNA (tRNA) or DNA mismatch repair genes conserved across bacterial genomes.

All of these type V subtypes rely on a tracrRNA that hybridizes through its 3′ end to the crRNA repeat, although few additional shared properties exist (**Figure 4*c***). For instance, the length of the processed tracrRNA varies from ~78 (for V-B) to ~216 nt (for V-K), while the extent of predicted base-pairing between the crRNA repeat and unprocessed tracrRNA antirepeat varies from 5 (for V-C/V-D) to 20 bp (for V-E) (37, 67, 96). The types of structures formed between the crRNA repeat and tracrRNA antirepeat and how these structures are processed to form the final mature crRNA range widely and often deviate from the paradigm set by type II systems. Type V-B, V-E, V-F1, and V-G systems represent the closest analogs to type II systems, as the tracrRNA for these systems forms an extended duplex with the crRNA repeat that presumably undergoes processing by RNase III (63, 67, 96, 114). All represent notable deviations from the crRNA:tracrRNA duplex associated with Cas9, though, because the repeat:antirepeat duplex associated with these nucleases is divided into two regions of pairing. For V-B, V-F1, and V-G systems, the two regions are separated by an intervening pseudoknot and, for V-F1 systems, additional stem-loops. The crRNA:tracrRNA duplex for V-F1 systems may also be processed by a host ribonuclease besides RNase III, as the reconstructed host genome lacks this gene and no duplex processing by RNase III was observed with Cas12f1 in vitro (36). Within type V-E systems, the repeat:antirepeat duplex includes a triple helix formed between two strands of the tracrRNA and one strand of the crRNA repeat (67). V-C and V-D systems represent the largest deviation from type II systems to date. For these systems, the 5′ end of the tracrRNA is predicted to form a large hairpin with a bulge near the base of the stem (37). Mutational analyses and sequence conservation within the crRNA and tracrRNA suggested that a short stretch within this bulge base-pairs with the crRNA repeat to form a pseudoknot, although further structural analyses are needed. Cas12c was further shown to process the crRNA repeat a fixed distance from the proposed crRNA:tracrRNA pseudoknot. This mode of processing reflects nucleus-mediated processing of the repeat by Cas12a and Cas12i, while the tracrRNA aids in crRNA recognition by the nuclease (37). In vitro processing of the

crRNA by Cas12d was inconclusive. Based on the unique proposed structure of the tracrRNA and its involvement in crRNA processing, the authors proposed the tracrRNA for these systems instead be called a short-complementarity untranslated RNA (scoutRNA) (37). How the crRNA:tracrRNA duplex is formed and processed remains to be elucidated for V-K systems.

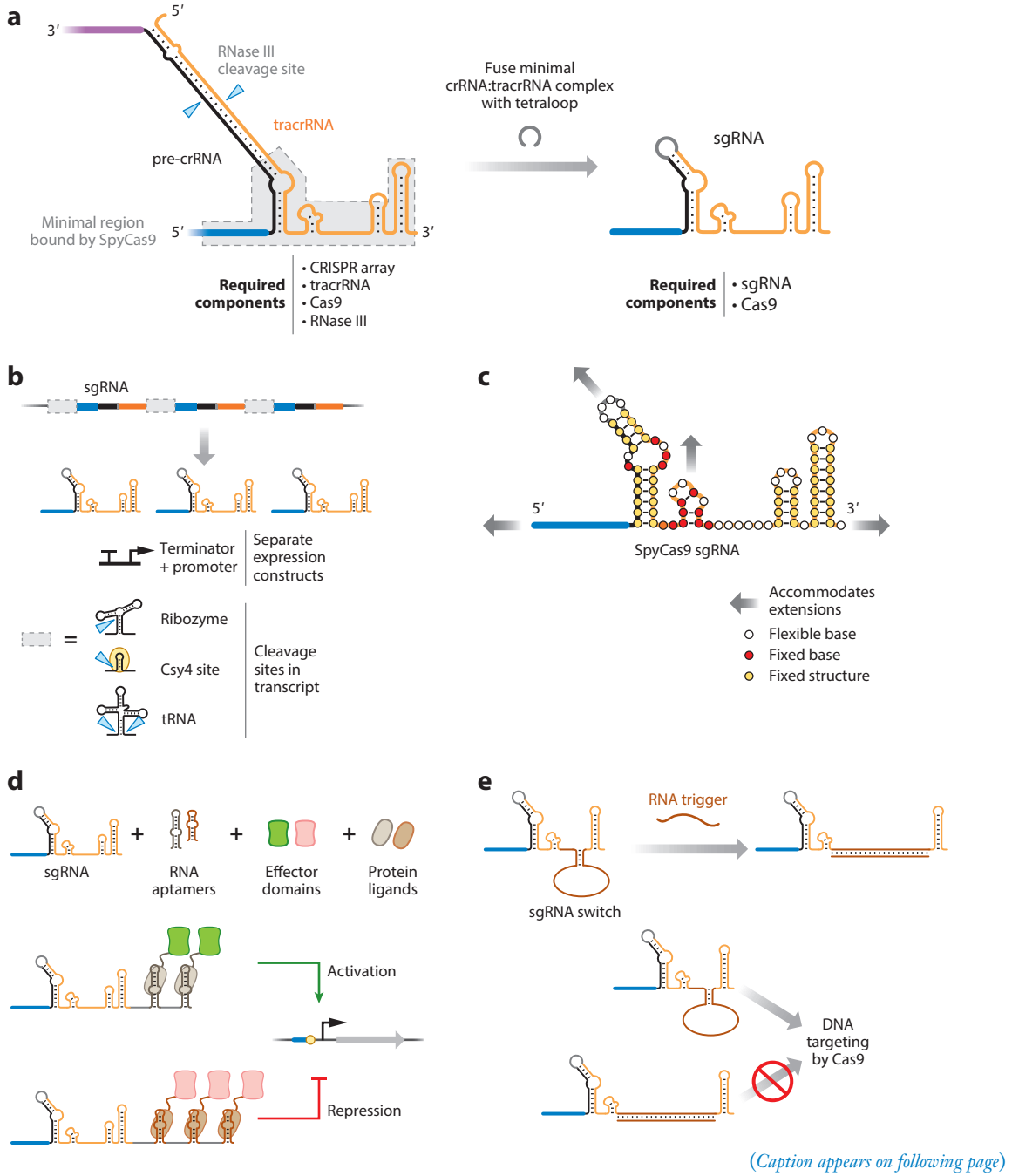
The few available crystal structures of Cas12 nucleases bound to a guide RNA have begun to reveal how these nucleases uniquely interact with the crRNA:tracrRNA complex. To date, structures are available for Cas12b, Cas12e, and Cas12f1 bound to an sgRNA and target DNA (67, 68, 103, 115), while a structure was reported for Cas12g after acceptance of this review (63). Cas12b consists of two lobes that both interact extensively with the crRNA:tracrRNA complex, reflecting sequence-specific as well as nonspecific interactions (115). Most of the interactions occur through the region of the repeat:antirepeat stem adjacent to the guide as well as two of the stem-loops within the rest of the tracrRNA. The pseudoknot within the tracrRNA minimally contacts the nuclease and instead likely helps form the structure of the crRNA:tracrRNA complex recognized by Cas12b. Cas12e interacts with the triple helix formed between the crRNA repeat and tracrRNA as well as with a stem-loop within the rest of the tracrRNA (67). Beyond the triple helix, though, the traditional duplex formed between the crRNA repeat and the tracrRNA projects away from the nuclease without any direct contacts. Finally, Cas12f1, as a miniature version of other Cas12 nucleases, dimerizes to form the typical bilobed architecture that recognizes one crRNA:tracrRNA complex (103, 111). Both lobes interact extensively with the sgRNA but to varying degrees. For instance, the pseudoknot interacts with only one nuclease monomer, while the stem-loop sandwiched between the divided regions of the crRNA:tracrRNA duplex interacts with both monomers. These few structures highlight the diversity of mechanisms Cas12 nucleases employ to recognize the crRNA:tracrRNA complex, and new modalities likely await discovery as other Cas12 nucleases are crystallized.

THE tracrRNA IN CRISPR TECHNOLOGIES

Beyond its natural functions, CRISPR-Cas systems have been a seemingly unlimited source of CRISPR technologies. The technological potential of these immune systems was clear as the original mechanisms emerged around 2010. However, much of the early characterization work focused on type I and III systems that required three to eight proteins along with designed crRNAs for any application (70). Studies of type II systems instead offered the most expedient path to CRISPR technologies, as these systems rely on one Cas protein for immune defense, and the next discovered single-effector nuclease (Cas12a) came much later (117). Utilizing Cas9 initially required two RNAs as well as RNase III (or an equivalent ribonuclease). While processed crRNA:tracrRNA complexes have been widely used with delivered ribonucleoprotein complexes, expressing these various components in cells represented a notable but surmountable hurdle (17). Fortunately, a simple workaround was immediately apparent: fusing the processed repeat:antirepeat duplex with a stable tetraloop (50) (**Figure 5*a***). The resulting sgRNA still contained the final processed crRNA:tracrRNA complex bound by Cas9 but as one RNA species. With this addition, RNA-directed DNA targeting by Cas9 could be achieved with only two components. The Cas9-sgRNA combination quickly became the standard for implementing CRISPR technologies. The concept of fusing the processed crRNA:tracrRNA complex through its duplex also readily applied to the tracrRNA-dependent Cas12 nucleases, allowing many of these nucleases to also be harnessed as two-component CRISPR technologies (54, 67, 99, 105).

While the critical contribution of sgRNAs to advancing CRISPR technologies cannot be overstated, their implementation came with an immediate drawback: loss of the multiplexing inherent to CRISPR arrays. CRISPR arrays encode and produce large numbers of crRNAs from a singular compact locus. For type II and many type V systems, the tracrRNA serves as the

processing element needed to convert a long pre-crRNA into individual crRNAs. As a result, fusing the crRNA:tracrRNA duplex prevents the architecture of the CRISPR array from being readily used to express multiple sgRNAs. This challenge gave way to numerous workarounds that could achieve multiplexing with sgRNAs (see 72 for an in-depth review) (**Figure 5*b***). One



(*Caption appears on following page*)

Figure 5 (*Figure appears on preceding page*)

tracrRNA technologies. (*a*) Creation of the sgRNA. The minimal region of the crRNA:tracrRNA duplex is fused with a short tetraloop. The resulting sgRNA only needs to be paired with Cas9 to achieve targeted DNA binding and cleavage. In contrast, using Cas9 and a CRISPR array also requires the tracrRNA and RNase III. (*b*) Coexpressing multiple sgRNAs. sgRNAs lose the inherent multiplexing capabilities of CRISPR arrays because a tracrRNA is fused to each crRNA. However, different workarounds have been developed to coexpress multiple sgRNAs, including arrayed expression constructs and single transcripts with multiple sgRNAs and intervening cleavage domains. (*c*) Extensively modifying the sgRNA sequence. Almost all of the sgRNA can be altered as long as the overall secondary structure is maintained. These sequence alterations can enhance DNA targeting by improving folding or allowing the synthesis of linear DNA that encode extremely long sgRNA arrays. Different domains within the sgRNA can also be extended without interfering with Cas9 binding or DNA targeting. (*d*) Extending the tracrRNA to localize different effector domains to Cas9 targets. Aptamer domains are fused to the 3′ end of the tracrRNA, and the cognate protein ligands are fused to the effector domains. The guide sequence therefore determines which effector domains are localized to a specific DNA locus. (*e*) sgRNA switches. Switches engineered by extending the tracrRNA tail are shown, although other configurations have been reported. Binding of the sensed RNA trigger (*brown*) drives the sgRNA switch into a conformation incompatible with Cas9 recognition and subsequent DNA targeting. Abbreviations: Cas, CRISPR-associated; CRISPR, clustered regularly interspaced short palindromic repeats; crRNA, CRISPR RNA; pre-crRNA, precursor CRISPR RNA; sgRNA, single-guide RNA; SpyCas9, *Streptococcus pyogenes* Cas9; tracrRNA, *trans*-activating CRISPR RNA, tRNA, transfer RNA.

common approach has been fusing sgRNA expression constructs, with each construct encoding a promoter, sgRNA, and transcriptional terminator (53, 91). The resulting constructs were much larger than the standard spacer-repeat subunit of a CRISPR array, although the sgRNA expression constructs could be spliced together using established cloning approaches. An alternative and more compact approach involved separating multiple sgRNAs with RNA cleavage sites. Different types of cleavage sites have been employed, including ribozymes that undergo self-cleavage without accessory factors (32), the repeat hairpin from I-F CRISPR-Cas systems cleaved by the Csy4/Cas6 processing enzyme (106), and immature tRNAs processed by RNase P and RNase Z (84, 112). While these approaches have become standard for CRISPR-based multiplexing in eukaryotes, CRISPR arrays and the tracrRNA are still regularly employed in bacteria (18, 47, 62).

The ability to multiplex with an expanding collection of CRISPR technologies spurred interest in implementing multiple technologies in the same cellular environment. However, using different versions of the same Cas nuclease (e.g., Cas9 and dCas9) faced immediate complications, as the two nucleases cannot differentiate between sgRNAs intended for different functions. Insights into the orthogonal recognition of the crRNA:tracrRNA complex became the immediate solution. By utilizing two phylogenetically distant Cas9 nucleases, more than one function could be implemented at one time. Examples have included achieving DNA cleavage and gene regulation in the same cell (28, 64), enhancing editing through proximal binding (12), and performing combinatorial genetic screens (76). Later work established other innovative workarounds centered around engineering the sgRNA so one Cas nuclease could be used for multiple functions, such as expressing sgRNAs with full-length or truncated guides to respectively edit DNA (with DNA cleavage) and regulate gene expression (with DNA binding but no cleavage) (20, 57, 66). Even with these more recent workarounds, the use of orthogonal Cas nucleases remains a popular approach.

THE tracrRNA AS AN ENGINEERING PLATFORM

The original suite of CRISPR technologies relied heavily on sgRNAs directly derived from the fused crRNA:tracrRNA complex. However, new generations of these technologies treated the sgRNA not as a fixed scaffold for guide sequences but as an engineering platform that could be extensively modified. One common insight from structural studies and mutational analyses was that the crRNA:tracrRNA complex is bound by few sequence restrictions as long as the general secondary structure is maintained and all domains are present (5, 42, 80). One study combined these insights along with machine learning and three rounds of testing to create design rules for

synthetic sgRNA scaffolds (88). The resulting set of scaffolds maintained recognition and DNA targeting by SpyCas9 but bore little resemblance to each other (**Figure 5*c***). By combining different sgRNA scaffolds, researchers could synthesize a large set of sgRNA expression constructs as a single piece of DNA—a feat that would have been impossible with unmodified sgRNAs given the repetitiveness of the resulting DNA. In some cases, altered scaffolds exhibited enhanced DNA targeting for sgRNAs as well as for crRNA:tracrRNA duplexes (21, 88, 93). Similar approaches are now being applied to guide RNA scaffolds utilized by other Cas nucleases (23).

Beyond modifying the sequence of the sgRNA scaffold, the scaffold can accommodate extensions to imbue the sgRNA with new functionalities. These extensions could be added to locations on the crRNA or tracrRNA portions of the sgRNA represented by highly flexible regions or extending beyond the minimal region bound by Cas9, including the 5′ end of the sgRNA guide, the upper stem of the repeat:antirepeat duplex, the top of the nexus, and the 3′ end of the tracrRNA tail (**Figure 5*c***). Most extensions had little impact on DNA targeting (24, 69, 79, 116), although extending the 5′ end of the sgRNA guide even a few nucleotides has been associated with reduced targeting activity (74). The most widely adopted approach involves extending the sgRNA with RNA aptamers that recruit proteins to a target DNA locus (**Figure 5*d***). The proteins principally involve the fusion of an aptamer's peptide ligand to an effector module, such as an epigenetic or transcriptional modifier for gene regulation or a fluorescent protein for DNA or RNA imaging (24, 69, 116). Because different aptamers can be fused with different copy numbers and combinations, individual sgRNAs can dictate functionally and quantitatively different effects at a given locus, even with the same Cas nuclease. Beyond introducing aptamers, researchers have extended sgRNAs to render DNA targeting by Cas9 dependent on the presence of another RNA (31, 35, 49, 97) (**Figure 5*e***). This RNA trigger base-pairs with the extended portion of the engineered sgRNA switch, driving a conformational change that either allows or prevents Cas9 recognition or DNA targeting. A similar concept was applied to fuse sgRNAs with aptazymes, rendering DNA targeting dependent on a small molecule (60, 104).

These examples highlight opportunities to engineer the tracrRNA for CRISPR technologies, although they all rely on an sgRNA or crRNA:tracrRNA complex. In contrast, the discovery of ncrRNAs (**Figure 3*c***) offered a distinct opportunity for tracrRNA engineering (5). Specifically, reprogramming the tracrRNA antirepeat to base-pair with a selected cellular RNA drove the generation of an ncrRNA and subsequent sequence-specific DNA targeting by Cas9 (48) (**Figure 6*a***). These reprogrammed tracrRNAs, or Rptrs, open numerous applications in which a CRISPR technology can be rendered fully dependent on the presence of a selected RNA. The most immediate application involved cleaving a matching DNA target as a specific readout of the corresponding RNA in a sample, forming a multiplexable diagnostic platform called LEOPARD (leveraging engineered tracrRNAs and on-target DNAs for parallel RNA detection) (**Figure 6*b***). LEOPARD was used to sense multiple respiratory viruses, including SARS-CoV-2, while the sequence-specificity of Cas9 allowed LEOPARD to readily differentiate a single point mutation, distinguishing SARS-CoV-2 and the D614G variant that dominated early in the COVID-19 pandemic in patient samples.

tracrRNA BIOLOGY AND TECHNOLOGIES GOING FORWARD

Clearly, our understanding of tracrRNA biology and its use in CRISPR technologies has advanced tremendously since the tracrRNA's discovery roughly a decade ago. At the same time, there are ample opportunities to further explore the biology of tracrRNAs and how they can be further harnessed for CRISPR technologies. One opportunity is within crRNA biogenesis. While the basic steps were elucidated with the original discovery of the tracrRNA, there are still fundamental

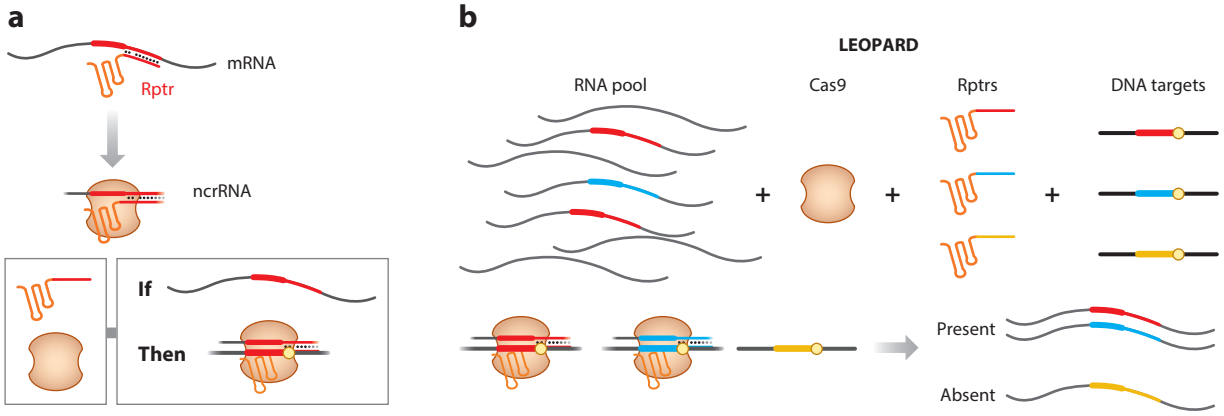


Figure 6

Reprogramming tracrRNAs for RNA detection. (*a*) Reprogramming the tracrRNA to create RNA-derived ncrRNAs. The reprogrammed Rptrs contain an altered antirepeat domain designed to base-pair with an RNA of interest, recreating the standard crRNA:tracrRNA duplex. Cas9 recognizes the complex and utilizes the resulting ncrRNAs for DNA targeting. As a result, DNA targeting occurs only if the RNA of interest is present. (*b*) The Rptr-based platform for multiplexable RNA detection. The platform, called LEOPARD, combines Rtprs and associated DNA targets with an RNA pool, and the DNA targets are monitored to determine which of the RNAs of interest are present in the sample. Abbreviations: Cas, CRISPR-associated; CRISPR, clustered regularly interspaced short palindromic repeats; crRNA, CRISPR RNA; LEOPARD, leveraging engineered tracrRNAs and on-target DNAs for parallel RNA detection; mRNA, messenger RNA, ncrRNA, noncanonical CRISPR RNA; Rptrs, reprogrammed tracrRNA; tracrRNA, *trans*-activating CRISPR RNA.

aspects that remain poorly understood. For instance, while the tracrRNA, crRNA, Cas9, RNase III, and other host RNases participate in the formation of the final effector complex, the order of events remains largely unknown. The abundance of crRNAs from an array also varies widely, likely impacting the extent of immune defense conferred by each crRNA. However, the underlying mechanisms remain poorly explored (19, 65, 87). Another opportunity centers around the structures of Cas9 and Cas12 nucleases bound to the crRNA:tracrRNA complex. The few available structures (119) (**Supplemental Table 1**) have revealed varying shapes and folds that allow the nuclease to selectively distinguish the crRNA:tracrRNA complex from the rest of the RNA milieu. These structures cannot be readily predicted with standard RNA folding algorithms, requiring additional structural studies. The presence of pseudoknots, triple helices, and other complex structures associated with crRNA:tracrRNA complexes bound by Cas12 nucleases further suggests an incredible range of structures await discovery within the diversity of Cas9 nucleases (33). Finally, the combination of RNA targeting by Cas9, the regulatory role of the scaRNA in *Francisella*, and the generation of ncrRNAs in *Campylobacter* suggests that the tracrRNA could participate in roles extending beyond immune defense (109). Beyond further uncovering tracrRNA biology, continued research efforts could translate into new and improved CRISPR technologies given the inherent importance of the tracrRNA in sgRNAs. The engineering of Rptrs opens up additional opportunities beyond multiplexable RNA detection and could be extended to the growing set of tracrRNA-dependent Cas12 nucleases. Overall, the study and engineering of tracrRNAs remain a fruitful source for biological discovery and beneficial technologies.

Supplemental Material >

DISCLOSURE STATEMENT

C.L.B. is a cofounder and member of the scientific advisory board for Locus Biosciences and a member of the scientific advisory board for Benson Hill. C.L.B. has filed multiple patent applications related to CRISPR technologies.

ACKNOWLEDGMENTS

We thank Natalia Luise Peeck for help illustrating the Cas12 structures. This work was supported in part through a European Research Council Consolidator Award (865973 to C.L.B.) and the SPP2141 priority program of the Deutsche Forschungsgemeinschaft (BE 6703/1-1 to C.L.B.).

LITERATURE CITED

1. Acharya S, Mishra A, Paul D, Ansari AH, Azhar M, et al. 2019. *Francisella novicida* Cas9 interrogates genomic DNA with very high specificity and can be used for mammalian genome editing. *PNAS* 116(42):20959–68
2. Anders C, Niewoehner O, Duerst A, Jinek M. 2014. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* 513(7519):569–73
3. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, et al. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315(5819):1709–12
4. Batra R, Nelles DA, Pirie E, Blue SM, Marina RJ, et al. 2017. Elimination of toxic microsatellite repeat expansion RNA by RNA-targeting Cas9. *Cell* 170(5):899–912.e10
5. Briner AE, Donohoue PD, Gomaa AA, Selle K, Slorach EM, et al. 2014. Guide RNA functional modules direct Cas9 activity and orthogonality. *Mol. Cell* 56(2):333–39
6. Briner AE, Henriksen ED, Barrangou R. 2016. Prediction and validation of native and engineered Cas9 guide sequences. *Cold Spring Harb. Protoc.* 2016(7):628–634
7. Brouns SJJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJH, et al. 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321(5891):960–64
8. Burstein D, Harrington LB, Strutt SC, Probst AJ, Anantharaman K, et al. 2017. New CRISPR-Cas systems from uncultivated microbes. *Nature* 542(7640):237–41
9. Carte J, Pfister NT, Compton MM, Terns RM, Terns MP. 2010. Binding and cleavage of CRISPR RNA by Cas6. *RNA* 16(11):2181–88
10. Carte J, Wang R, Li H, Terns RM, Terns MP. 2008. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev.* 22(24):3489–96
11. Chang N, Sun C, Gao L, Zhu D, Xu X, et al. 2013. Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. *Cell Res.* 23(4):465–72
12. Chen F, Ding X, Feng Y, Seebeck T, Jiang Y, Davis GD. 2017. Targeted activation of diverse CRISPR-Cas systems for mammalian genome editing via proximal CRISPR targeting. *Nat. Commun.* 8:14958
13. Chen JS, Ma E, Harrington LB, Da Costa M, Tian X, et al. 2018. CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. *Science* 360(6387):436–39
14. Cho SW, Kim S, Kim JM, Kim J-S. 2013. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* 31(3):230–32
15. Chylinski K, Le Rhun A, Charpentier E. 2013. The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA Biol.* 10(5):726–37
16. Chyou T-Y, Brown CM. 2019. Prediction and diversity of tracrRNAs from type II CRISPR-Cas systems. *RNA Biol.* 16(4):423–34
17. Cong L, Ran FA, Cox D, Lin S, Barretto R, et al. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339(6121):819–23
18. Cress BF, Toparlak ÖD, Guleria S, Lebovich M, Stieglitz JT, et al. 2015. CRISPathBrick: modular combinatorial assembly of type II-A CRISPR arrays for dCas9-mediated multiplex transcriptional repression in *E. coli*. *ACS Synth. Biol.* 4(9):987–1000
19. Creutzburg SCA, Wu WY, Mohanraju P, Swartjes T, Alkan F, et al. 2020. Good guide, bad guide: spacer sequence-dependent cleavage efficiency of Cas12a. *Nucleic Acids Res.* 48(6):3228–43
20. Dahlman JE, Abudayyeh OO, Joung J, Gootenberg JS, Zhang F, Konermann S. 2015. Orthogonal gene knockout and activation with a catalytically active Cas9 nuclease. *Nat. Biotechnol.* 33(11):1159–61
21. Dang Y, Jia G, Choi J, Ma H, Anaya E, et al. 2015. Optimizing sgRNA structure to improve CRISPR-Cas9 knockout efficiency. *Genome Biol.* 16(1):280

22. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, et al. 2011. CRISPR RNA maturation by *trans*-encoded small RNA and host factor RNase III. *Nature* 471(7340):602–7
23. DeWeirdt PC, Sanson KR, Sangree AK, Hegde M, Hanna RE, et al. 2021. Optimization of AsCas12a for combinatorial genetic screens in human cells. *Nat. Biotechnol.* 39(1):94–104
24. Dong C, Fontana J, Patel A, Carothers JM, Zalatan JG. 2018. Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria. *Nat. Commun.* 9(1):2489
25. Dooley SK, Baken EK, Moss WN, Howe A, Young JK. 2020. Identification and evolution of Cas9 tracrRNAs. *CRISPR J.* 4(3):438–47
26. Dugar G, Herbig A, Förstner KU, Heidrich N, Reinhardt R, et al. 2013. High-resolution transcriptome maps reveal strain-specific regulatory features of multiple *Campylobacter jejuni* isolates. *PLOS Genet.* 9(5):e1003495
27. Dugar G, Leenay RT, Eisenbart SK, Bischler T, Aul BU, et al. 2018. CRISPR RNA-dependent binding and cleavage of endogenous RNAs by the *Campylobacter jejuni* Cas9. *Mol. Cell* 69(5):893–905.e7
28. Esvelt KM, Mali P, Braff JL, Moosburner M, Yaung SJ, Church GM. 2013. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat. Methods* 10(11):1116–21
29. Fonfara I, Le Rhun A, Chylinski K, Makarova KS, Lécrivain A-L, et al. 2014. Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. *Nucleic Acids Res.* 42(4):2577–90
30. Fonfara I, Richter H, Bratovič M, Le Rhun A, Charpentier E. 2016. The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature* 532(7600):517–21
31. Galizi R, Duncan JN, Rostain W, Quinn CM, Storch M, et al. 2020. Engineered RNA-interacting CRISPR guide RNAs for genetic sensing and diagnostics. *CRISPR J.* 3(5):398–408
32. Gao Y, Zhao Y. 2014. Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. *J. Integr. Plant Biol.* 56(4):343–49
33. Gasiunas G, Young JK, Karvelis T, Kazlauskas D, Urbaitis T, et al. 2020. A catalogue of biochemically diverse CRISPR-Cas9 orthologs. *Nat. Commun.* 11(1):5512
34. Gratz SJ, Cummings AM, Nguyen JN, Hamm DC, Donohue LK, et al. 2013. Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics* 194(4):1029–35
35. Hanewich-Hollatz MH, Chen Z, Hochrein LM, Huang J, Pierce NA. 2019. Conditional guide RNAs: programmable conditional regulation of CRISPR/Cas function in bacterial and mammalian cells via dynamic RNA nanotechnology. *ACS Cent. Sci.* 5(7):1241–49
36. Harrington LB, Burstein D, Chen JS, Paez-Espino D, Ma E, et al. 2018. Programmed DNA destruction by miniature CRISPR-Cas14 enzymes. *Science* 362(6416):839–42
37. Harrington LB, Ma E, Chen JS, Witte IP, Gertz D, et al. 2020. A scoutRNA is required for some type V CRISPR-Cas systems. *Mol. Cell* 79(3):416–24.e5
38. Haurwitz RE, Jinek M, Wiedenheft B, Zhou K, Doudna JA. 2010. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* 329(5997):1355–58
39. Heler R, Samai P, Modell JW, Weiner C, Goldberg GW, et al. 2015. Cas9 specifies functional viral targets during CRISPR-Cas adaptation. *Nature* 519(7542):199–202
40. Hirano H, Gootenberg JS, Horii T, Abudayyeh OO, Kimura M, et al. 2016. Structure and engineering of *Francisella novicida* Cas9. *Cell* 164(5):950–61
41. Hoikkala V, Ravantti J, Díez-Villaseñor C, Tiirola M, Conrad RA, et al. 2021. Cooperation between different CRISPR-Cas types enables adaptation in an RNA-targeting system. *mBio* 12(2):e03338-20
42. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, et al. 2013. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* 31(9):827–32
43. Huai C, Li G, Yao R, Zhang Y, Cao M, et al. 2017. Structural insights into DNA cleavage activation of CRISPR-Cas9 system. *Nat. Commun.* 8(1):1375
44. Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, et al. 2013. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* 31(3):227–29
45. Jiang F, Taylor DW, Chen JS, Kornfeld JE, Zhou K, et al. 2016. Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science* 351(6275):867–71
46. Jiang F, Zhou K, Ma L, Gressel S, Doudna JA. 2015. A Cas9-guide RNA complex preorganized for target DNA recognition. *Science* 348(6242):1477–81

47. 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):233–39
48. Jiao C, Sharma S, Dugar G, Peeck NL, Bischler T, et al. 2021. Noncanonical crRNAs derived from host transcripts enable multiplexable RNA detection by Cas9. *Science* 372(6545):941–48
49. Jin M, Garreau de Loubresse N, Kim Y, Kim J, Yin P. 2019. Programmable CRISPR-Cas repression, activation, and computation with sequence-independent targets and triggers. *ACS Synth. Biol.* 8(7):1583–89
50. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337(6096):816–21
51. Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. 2013. RNA-programmed genome editing in human cells. *eLife* 2:e00471
52. Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, et al. 2014. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* 343(6176):1247997
53. Kabadi AM, Ousterout DG, Hilton IB, Gersbach CA. 2014. Multiplex CRISPR/Cas9-based genome engineering from a single lentiviral vector. *Nucleic Acids Res.* 42(19):e147
54. Karvelis T, Bigelyte G, Young JK, Hou Z, Zedaveinyte R, et al. 2020. PAM recognition by miniature CRISPR-Cas12f nucleases triggers programmable double-stranded DNA target cleavage. *Nucleic Acids Res.* 48(9):5016–23
55. Karvelis T, Gasiunas G, Miksys A, Barrangou R, Horvath P, Siksnys V. 2013. crRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus*. *RNA Biol.* 10(5):841–51
56. Karvelis T, Young JK, Siksnys V. 2019. A pipeline for characterization of novel Cas9 orthologs. *Methods Enzymol.* 616:219–40
57. Kiani S, Chavez A, Tuttle M, Hall RN, Chari R, et al. 2015. Cas9 gRNA engineering for genome editing, activation and repression. *Nat. Methods* 12(11):1051–54
58. Klompe SE, Vo PLH, Halpin-Healy TS, Sternberg SH. 2019. Transposon-encoded CRISPR–Cas systems direct RNA-guided DNA integration. *Nature* 571:219–25
59. Koonin EV, Makarova KS. 2013. CRISPR-Cas: evolution of an RNA-based adaptive immunity system in prokaryotes. *RNA Biol.* 10(5):679–86
60. Kundert K, Lucas JE, Watters KE, Fellmann C, Ng AH, et al. 2019. Controlling CRISPR-Cas9 with ligand-activated and ligand-deactivated sgRNAs. *Nat. Commun.* 10(1):2127
61. Leenay RT, Beisel CL. 2017. Deciphering, communicating, and engineering the CRISPR PAM. *J. Mol. Biol.* 429(2):177–91
62. Leenay RT, Vento JM, Shah M, Martino ME, Leulier F, Beisel CL. 2019. Genome editing with CRISPR-Cas9 in *Lactobacillus plantarum* revealed that editing outcomes can vary across strains and between methods. *Biotechnol. J.* 14(3):e1700583
63. Li Z, Zhang H, Xiao R, Han R, Chang L. 2021. Cryo-EM structure of the RNA-guided ribonuclease Cas12g. *Nat. Chem. Biol.* 17:387–93
64. Lian J, HamediRad M, Hu S, Zhao H. 2017. Combinatorial metabolic engineering using an orthogonal tri-functional CRISPR system. *Nat. Commun.* 8(1):1688
65. Liao C, Ttofali F, Slotkowski RA, Denny SR, Cecil TD, et al. 2019. Modular one-pot assembly of CRISPR arrays enables library generation and reveals factors influencing crRNA biogenesis. *Nat. Commun.* 10(1):2948
66. Liu G, Yin K, Zhang Q, Gao C, Qiu J-L. 2019. Modulating chromatin accessibility by transactivation and targeting proximal dsgRNAs enhances Cas9 editing efficiency in vivo. *Genome Biol.* 20(1):145
67. Liu J-J, Orlova N, Oakes BL, Ma E, Spinner HB, et al. 2019. CasX enzymes comprise a distinct family of RNA-guided genome editors. *Nature* 566(7743):218–23
68. Liu L, Chen P, Wang M, Li X, Wang J, et al. 2017. C2c1-sgRNA complex structure reveals RNA-guided DNA cleavage mechanism. *Mol. Cell* 65(2):310–22
69. Ma H, Tu L-C, Naseri A, Chung Y-C, Grunwald D, et al. 2018. CRISPR-Sirius: RNA scaffolds for signal amplification in genome imaging. *Nat. Methods* 15(11):928–31
70. Makarova KS, Wolf YI, Iranzo J, Shmakov SA, Alkhnbashi OS, et al. 2020. Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* 18(2):67–83

71. Mali P, Yang L, Esvelt KM, Aach J, Guell M, et al. 2013. RNA-guided human genome engineering via Cas9. *Science* 339(6121):823–26
72. McCarty NS, Graham AE, Studená L, Ledesma-Amaro R. 2020. Multiplexed CRISPR technologies for gene editing and transcriptional regulation. *Nat. Commun.* 11(1):1281
73. Meeske AJ, Marraffini LA. 2018. RNA guide complementarity prevents self-targeting in type VI CRISPR systems. *Mol. Cell* 71(5):791–801.e3
74. Mullally G, van Aelst K, Naqvi MM, Diffin FM, Karvelis T, et al. 2020. 5′ modifications to CRISPR-Cas9 gRNA can change the dynamics and size of R-loops and inhibit DNA cleavage. *Nucleic Acids Res.* 48(12):6811–23
75. Mullard A. 2020. CRISPR pioneers win Nobel prize. *Nat. Rev. Drug Discov.* 19(12):827
76. Najm FJ, Strand C, Donovan KF, Hegde M, Sanson KR, et al. 2018. Orthologous CRISPR-Cas9 enzymes for combinatorial genetic screens. *Nat. Biotechnol.* 36(2):179–89
77. Nelles DA, Fang MY, O'Connell MR, Xu JL, Markmiller SJ, et al. 2016. Programmable RNA tracking in live cells with CRISPR/Cas9. *Cell* 165(2):488–96
78. Nicholson AW. 2014. Ribonuclease III mechanisms of double-stranded RNA cleavage. *WIREs RNA* 5(1):31–48
79. Nishimasu H, Cong L, Yan WX, Ran FA, Zetsche B, et al. 2015. Crystal structure of *Staphylococcus aureus* Cas9. *Cell* 162(5):1113–26
80. Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, et al. 2014. Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* 156(5):935–49
81. O'Connell MR, Oakes BL, Sternberg SH, East-Seletsky A, Kaplan M, Doudna JA. 2014. Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature* 516(7530):263–66
82. Peters JE, Makarova KS, Shmakov S, Koonin EV. 2017. Recruitment of CRISPR-Cas systems by Tn7-like transposons. *PNAS* 114:E7358–66
83. Pickar-Oliver A, Gersbach CA. 2019. The next generation of CRISPR-Cas technologies and applications. *Nat. Rev. Mol. Cell Biol.* 20(8):490–507
84. Port F, Bullock SL. 2016. Augmenting CRISPR applications in *Drosophila* with tRNA-flanked sgRNAs. *Nat. Methods* 13(10):852–54
85. Price AA, Sampson TR, Ratner HK, Grakoui A, Weiss DS. 2015. Cas9-mediated targeting of viral RNA in eukaryotic cells. *PNAS* 112(19):6164–69
86. Ratner HK, Escalera-Maurer A, Le Rhun A, Jaggavarapu S, Wozniak JE, et al. 2019. Catalytically active Cas9 mediates transcriptional interference to facilitate bacterial virulence. *Mol. Cell* 75(3):498–510.e5
87. Reimann V, Alkhnbashi OS, Saunders SJ, Scholz I, Hein S, et al. 2017. Structural constraints and enzymatic promiscuity in the Cas6-dependent generation of crRNAs. *Nucleic Acids Res.* 45(2):915–25
88. Reis AC, Halper SM, Vezeau GE, Cetnar DP, Hossain A, et al. 2019. Simultaneous repression of multiple bacterial genes using nonrepetitive extra-long sgRNA arrays. *Nat. Biotechnol.* 37(11):1294–301
89. Rousseau BA, Hou Z, Gramelspacher MJ, Zhang Y. 2018. Programmable RNA cleavage and recognition by a natural CRISPR-Cas9 system from *Neisseria meningitidis*. *Mol. Cell* 69(5):906–14.e4
90. Saito M, Ladha A, Strecker J, Faure G, Neumann E, et al. 2021. Dual modes of CRISPR-associated transposon homing. *Cell* 184:2441–53.e18
91. Sakuma T, Nishikawa A, Kume S, Chayama K, Yamamoto T. 2014. Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. *Sci. Rep.* 4:5400
92. Sampson TR, Saroj SD, Llewellyn AC, Tzeng Y-L, Weiss DS. 2013. A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. *Nature* 497(7448):254–57
93. Scott T, Urak R, Soemardy C, Morris KV. 2019. Improved Cas9 activity by specific modifications of the tracrRNA. *Sci. Rep.* 9(1):16104
94. Sharma CM, Vogel J. 2014. Differential RNA-seq: the approach behind and the biological insight gained. *Curr. Opin. Microbiol.* 19:97–105
95. Shen B, Zhang J, Wu H, Wang J, Ma K, et al. 2013. Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. *Cell Res.* 23(5):720–23
96. Shmakov S, Abudayyeh OO, Makarova KS, Wolf YI, Gootenberg JS, et al. 2015. Discovery and functional characterization of diverse Class 2 CRISPR-Cas systems. *Mol. Cell* 60(3):385–97

97. Siu K-H, Chen W. 2019. Riboregulated toehold-gated gRNA for programmable CRISPR-Cas9 function. *Nat. Chem. Biol.* 15(3):217–20
98. Storz G, Vogel J, Wassarman KM. 2011. Regulation by small RNAs in bacteria: expanding frontiers. *Mol. Cell* 43(6):880–91
99. Strecker J, Jones S, Koopal B, Schmid-Burgk J, Zetsche B, et al. 2019. Engineering of CRISPR-Cas12b for human genome editing. *Nat. Commun.* 10(1):212
100. Strecker J, Ladha A, Gardner Z, Schmid-Burgk JL, Makarova KS, et al. 2019. RNA-guided DNA insertion with CRISPR-associated transposases. *Science* 365(6448):48–53
101. Strutt SC, Torrez RM, Kaya E, Negrete OA, Doudna JA. 2018. RNA-dependent RNA targeting by CRISPR-Cas9. *eLife* 7:e32724
102. Sun W, Yang J, Cheng Z, Amrani N, Liu C, et al. 2019. Structures of *Neisseria meningitidis* Cas9 complexes in catalytically poised and anti-CRISPR-inhibited states. *Mol. Cell* 76(6):938–52.e5
103. Takeda SN, Nakagawa R, Okazaki S, Hirano H, Kobayashi K, et al. 2021. Structure of the miniature type V-F CRISPR-Cas effector enzyme. *Mol. Cell* 81(3):558–70.e3
104. Tang W, Hu JH, Liu DR. 2017. Aptazyme-embedded guide RNAs enable ligand-responsive genome editing and transcriptional activation. *Nat. Commun.* 8:15939
105. Teng F, Cui T, Feng G, Guo L, Xu K, et al. 2018. Repurposing CRISPR-Cas12b for mammalian genome engineering. *Cell Discov.* 4:63
106. Tsai SQ, Wyvekens N, Khayter C, Foden JA, Thapar V, et al. 2014. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat. Biotechnol.* 32(6):569–76
107. Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, et al. 2013. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 153(4):910–18
108. Wei Y, Terns RM, Terns MP. 2015. Cas9 function and host genome sampling in Type II-A CRISPR–Cas adaptation. *Genes Dev.* 29(4):356–61
109. Wimmer F, Beisel CL. 2019. CRISPR-Cas systems and the paradox of self-targeting spacers. *Front. Microbiol.* 10:3078
110. Workman RE, Pammi T, Nguyen BTK, Graeff LW, Smith E, et al. 2021. A natural single-guide RNA repurposes Cas9 to autoregulate CRISPR-Cas expression. *Cell* 184(3):675–88.e19
111. Xiao R, Li Z, Wang S, Han R, Chang L. 2021. Structural basis for substrate recognition and cleavage by the dimerization-dependent CRISPR–Cas12f nuclease. *Nucleic Acids Res.* 49:4120–28
112. Xie K, Minkenberg B, Yang Y. 2015. Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *PNAS* 112(11):3570–75
113. Yamada M, Watanabe Y, Gootenberg JS, Hirano H, Ran FA, et al. 2017. Crystal structure of the minimal Cas9 from *Campylobacter jejuni* reveals the molecular diversity in the CRISPR-Cas9 systems. *Mol. Cell* 65(6):1109–21.e3
114. Yan WX, Hunnewell P, Alfonse LE, Carte JM, Keston-Smith E, et al. 2019. Functionally diverse type V CRISPR-Cas systems. *Science* 363(6422):88–91
115. Yang H, Gao P, Rajashankar KR, Patel DJ. 2016. PAM-dependent target DNA recognition and cleavage by C2c1 CRISPR-Cas endonuclease. *Cell* 167(7):1814–28.e12
116. Zalatan JG, Lee ME, Almeida R, Gilbert LA, Whitehead EH, et al. 2015. Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell* 160(1–2):339–50
117. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, et al. 2015. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* 163(3):759–71
118. Zhang Y, Heidrich N, Ampattu BJ, Gunderson CW, Seifert HS, et al. 2013. Processing-independent CRISPR RNAs limit natural transformation in *Neisseria meningitidis*. *Mol. Cell* 50(4):488–503
119. Zhang Y, Zhang H, Xu X, Wang Y, Chen W, et al. 2020. Catalytic-state structure and engineering of *Streptococcus thermophilus* Cas9. *Nat. Catal.* 3(10):813–23
120. Zuo Z, Zolekar A, Babu K, Lin VJ, Hayatshahi HS, et al. 2019. Structural and functional insights into the catalytic state of Cas9 HNH nuclease domain. *eLife* 8:e46500