Accepted Manuscript

Harnessing CRISPR/Cas systems for programmable transcriptional and post-transcriptional regulation

Ahmed Mahas, C. Neal Stewart, Magdy M. Mahfouz

PII: S0734-9750(17)30150-7
DOI: doi:[10.1016/j.biotechadv.2017.11.008](https://doi.org/10.1016/j.biotechadv.2017.11.008)
Reference: JBA 7174
To appear in: *Biotechnology Advances*
Received date: 5 September 2017
Revised date: 3 November 2017
Accepted date: 27 November 2017

Please cite this article as: Ahmed Mahas, C. Neal Stewart, Magdy M. Mahfouz , Harnessing CRISPR/Cas systems for programmable transcriptional and post-transcriptional regulation. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. Jba(2017), doi:[10.1016/j.biotechadv.2017.11.008](https://doi.org/10.1016/j.biotechadv.2017.11.008)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Harnessing CRISPR/Cas Systems for Programmable Transcriptional and Post-Transcriptional Regulation

Ahmed Mahas¹, C. Neal Stewart Jr², and Magdy M. Mahfouz^{2,*}

¹ Laboratory for Genome Engineering, Division of Biological Sciences, 4700 King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia, and ² Department of plant sciences, University of Tennessee, 252 Ellington Plant Sciences, 2431 Joe Johnson Dr., Knoxville, TN 37996, USA

*** Corresponding author:** Magdy M. Mahfouz (magdy.mahfouz@kaust.edu.sa)

Abstract

Genome editing has enabled broad advances and novel approaches in studies of gene function and structure; now, emerging methods aim to precisely engineer post-transcriptional processes. Developing precise, efficient molecular tools to alter the transcriptome holds great promise for biotechnology and synthetic biology applications. Different approaches have been employed for targeted degradation of RNA species in eukaryotes, but they lack programmability and versatility, thereby limiting their utility for diverse applications. The CRISPR/Cas9 system has been harnessed for genome editing in many eukaryotic species and, using a catalytically inactive Cas9 variant, the CRISPR/dCas9 system has been repurposed for transcriptional regulation. Recent studies have used other CRISPR/Cas systems for targeted RNA degradation and RNA-based manipulations. For example, Cas13a, a Type VI-A endonuclease, has been identified as an RNA-guided RNA ribonuclease and used for manipulation of RNA. Here, we discuss different modalities for targeted RNA interference with an emphasis on the potential applications of CRISPR/Cas systems as programmable transcriptional regulators for broad uses, including functional biology, biotechnology, and synthetic biology applications.

Keywords: Gene editing, CRISPR/Cas9, site-specific endonucleases, functional genomics, transcriptional regulation, synthetic biology, bioengineering, Cas13, dCas9

1. Introduction:

Transcription and translation are coupled and occur at the cytoplasm in prokaryotes, but in eukaryotic cells, these two processes occur in separate cellular compartments. This enables eukaryotes to regulate gene expression at multiple stages, starting at the level of DNA transcription in the nucleus and continuing to extensive post-transcriptional processing of pre- and mature mRNAs, which occurs in the nucleus and the cytoplasm (Orphanides and Reinberg, 2002). The regulation of gene expression is critical for many essential cellular processes, ranging from genome replication and repair, to cell division, developmental processes, cell fate decisions, and cell death. The ability to manipulate gene expression in a specific and precise manner has important implications for many areas of therapeutics and research.

Significant progress has been made in developing genome-engineering technologies that can efficiently produce DNA modifications that were previously believed to be beyond the reach of technology. However, precise manipulation of transcription and post-transcriptional processes has remained challenging. To manipulate transcription, research efforts have focused on developing synthetic transcription factors that can be engineered to precisely alter the level of specific transcripts. Naturally existing transcription factors typically consist of DNA-binding domains and transcriptional activation or repression domains that can turn the expression of target genes on or off, respectively. DNA binding domains of diverse natural transcriptional repressors, such as the TetR family, the Lac operator repressor, and LexA proteins, have been used to regulate eukaryotic genes (Brent and Ptashne, 1985; Cronin et al., 2001; Gossen and Bujard, 1992). However, an alternative and elegant strategy has emerged that relies on engineering DNA binding domains to bind to any user-defined DNA sequence. Two classes of proteins have been shown to be amenable for engineering customizable DNA-binding proteins: zinc finger arrays (ZFs) and transcription activator-like effectors (TALEs) (Boch et al., 2009; Bogdanove and Voytas, 2011; Kim and Kim, 2014; Klug, 2010; Moscou and Bogdanove, 2009; Wolfe et al., 2000). Conceptually, engineering a transcription factor to bind to a specific DNA sequence was the key to developing synthetic and programmable transcriptional regulators. Fusions of transcriptional regulatory domains (e.g. transcriptional activators or repressors) to the DNA-binding domains of ZFs and TALEs made it feasible to generate synthetic programmable

transcriptional regulators for targeted gene regulation (Beerli and Barbas, 2002; Beerli et al., 1998; Li et al., 2012; Maeder et al., 2013b; Perez-Pinera et al., 2013b; Zhang et al., 2011). However, as ZFs and TALEs recognize specific DNA sequences through protein–DNA interactions, each new target requires arduous protein design and construction, which has proven to be challenging and resource intensive (Hsu et al., 2014). The requirement for protein engineering for every single gene target and the inability to simultaneously target multiple genes have been major limitations in the usefulness and application of ZFs and TALEs.

In prokaryotes, clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) systems serve as adaptive molecular immunity mechanisms to fend off invading nucleic acids, including phages and conjugative plasmids (Garneau et al., 2010; Sorek et al., 2013). Practical interest in CRISPR/Cas systems arose from their ability to exploit RNA:DNA base pair complementarity between the guide RNA (gRNA) and the target sequence for target recognition and binding, together with the built-in nuclease activity of Cas nucleases that mediate sequence specific target cleavage, thus offering a readily programmable genome editing tool. The simplicity and effectiveness of the two-component Cas9 and single-guide RNA (sgRNA) system led to it becoming the most powerful genome engineering tool yet developed, offering great promise for functional biology, biotechnology, and genomic medicine.

DNA editing leads to permanent modifications within the genome; therefore, controlling gene expression at the transcriptional and post-transcriptional level represents an attractive and potentially safer alternative, especially for clinical and therapeutic applications. Inactivating mutations in the RuvC and HNH domains of Cas9 generate a nuclease-deficient Cas9 (dCas9) (Fig. 1A), that is unable to cleave target DNA but retains its ability to bind DNA in a sequence-specific manner (Jinek et al., 2012). Unlike ZFs and TALE-based transcriptional regulators, fusing dCas9 to effector domains such as transcriptional activators (CRISPRa) or repressors (CRISPRi) has enabled the repurposing of the system for scalable and robust transcriptome manipulation (Dominguez et al., 2016).

Given that RNA is involved in nearly all cellular processes, *in vivo* manipulation of endogenous RNAs would allow for a broader understanding of cellular functions and applications in

biotechnology, synthetic biology, and therapeutic research. Indeed, insights into the significance of RNA in gene regulation and sequence-specific gene silencing has led to the emergence of innovative approaches based on RNAs as therapeutic agents and targets (Burnett and Rossi, 2012; Dorsett and Tuschl, 2004; Kole et al., 2012; Lewin and Hauswirth, 2001; Rupaimoole and Slack, 2017; Zhou and Rossi, 2017). These approaches have shown promise in treating human diseases including genetic disorders, viral infections, and various cancers (Burnett and Rossi, 2012; Esteller, 2011; Kim and Rossi, 2007). For many years, different antisense-based technologies have been developed and used for the targeted inhibition of gene expression at the post-transcriptional level. Since the development of the original antisense oligonucleotide technology in 1978 (Zamecnik and Stephenson, 1978), various antisense technology derivatives such as ribozymes, DNAzymes, and co-suppression or RNA interference (RNAi) have been developed (Scherer and Rossi, 2003), with RNAi being the current gold standard for gene silencing (Bertrand et al., 2002). RNAi has been a key approach to elucidating gene function *in vitro* and *in vivo* across eukaryotic species (Agrawal et al., 2003; Elbashir et al., 2001). However, such techniques lack reproducibility and specificity, thereby limiting their utility (Jackson et al., 2006; Sigoillot et al., 2012).

Technologies for targeting specific transcripts for post-transcriptional manipulation, such as engineered RNA binding proteins, have also been difficult to design and utilize in a programmable and scalable manner. The potential of CRISPR/Cas systems for genome and transcriptome manipulation has expanded beyond DNA targeting. Recent studies have repurposed the DNA targeting CRISPR/SpCas9 for RNA targeting (RCas9) (O'Connell et al., 2014). In addition, other CRISPR/Cas9 orthologs, such as the *Francisella novicida* Cas9 (FnCas9), can target RNA in a sequence-specific manner and have been used to target RNA in mammalian cells (Price et al., 2015). More recently, investigation of the CRISPR world has revealed an entirely new class of Cas proteins, Cas13 systems, which have RNase but not DNase domains, and which can be harnessed for robust RNA detection and degradation. Additionally, the RNA-targeting CRISPR/Cas13 can be used as a catalytically dead Cas13 (dCas13) variant, which can be coupled to various effector domains, enabling myriad post-transcriptional manipulations of RNA substrates in living cells (Abudayyeh et al., 2016). As CRISPR/Cas systems have revolutionized the field of DNA targeting and editing, harnessing CRISPR/Cas

systems for transcriptional and post-transcriptional regulation will open many new avenues for potential applications in functional biology, biotechnology, medicine, and synthetic biology, promising to advance both basic and translational research.

In this review, we describe advances in the most widely adopted genome engineering tools, particularly those based on CRISPR/Cas systems, for regulating gene expression at the transcriptional and post-transcriptional levels. We highlight approaches for targeting and regulating endogenous RNA substrates, and discuss recent promising developments of RNA-targeted CRISPR/Cas systems, which have expanded the possibilities of RNA targeting and manipulation. We also present potential future applications of RNA-targeting CRISPR for functional biology, biotechnology, and synthetic biology for a new era of functional genomics and RNA manipulation, as well as the limitations of these technologies and areas for future advances.

2. Transcriptional regulation at the DNA level:

2.1. CRISPR/Cas systems:

The continuous arms race between bacteria and their viral parasites led to the evolution of natural defense mechanisms whose discovery has revolutionized molecular biology. Among these mechanisms are the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) genes that are used by most Archaea and many bacteria as adaptive immune systems to protect against invading phages and conjugative plasmids (Barrangou et al., 2007; Barrangou and Marraffini, 2014; Wiedenheft et al., 2012). The CRISPR/Cas adaptive immunity mechanism involves three main phases. The first phase is the acquisition of the molecular record, or memory, from the invading plasmid or virus. The sequences of these molecular records are first inserted into the CRISPR array (acquisition phase). The second phase is the transcription of the CRISPR array and maturation of the guide RNA molecules (transcription and maturation phase). In the final phase, the mature crRNAs guide an endonuclease or ribonuclease to bind the complementary sequence in the target nucleic acids,

and subsequently cleave or degrade the DNA or RNA, respectively (interference phase) (Makarova et al., 2011; van der Oost et al., 2014). Although the establishment of adaptive immunity among CRISPR/Cas systems follows these distinct steps, there is extensive variability in the nature of crRNA maturation and effector molecules (Barrangou, 2015). To date, all of the known CRISPR/Cas systems are classified into two main classes that are further subdivided into different types and subtypes based on the organization of their loci and signature proteins (Makarova et al., 2015; Shmakov et al., 2015). Class I CRISPR/Cas systems (encompassing type I, III, and IV) are the most abundant CRISPR/Cas systems found in bacteria and Archaea, and employ multi-subunit effector complexes (Makarova et al., 2015). Class II CRISPR/Cas systems (including type II, V, and VI) are less common and found mostly in bacteria, although they were recently found in Archaea (Burstein et al., 2017). These systems use a single, RNA-guided, multi-domain Cas protein to recognize and cleave target sequences. Due to the simplicity of reprogramming them to target any user-defined sequence, Class II CRISPR systems have attracted enormous interest as genome editing tools (Jinek et al., 2012; Shmakov et al., 2017).

2.2. CRISPR/Cas9:

Key research efforts on the type II CRISPR/Cas9 components of *Streptococcus pyogenes* and *Streptococcus thermophilus* and their underlying molecular mechanisms established CRISPR/Cas9 as a powerful genome engineering platform (Cong et al., 2013; Jinek et al., 2012). Cas9 is an endonuclease that is guided by two small RNA molecules, CRISPR RNA (crRNA) and the *trans*-activating crRNA (tracrRNA), to generate a double-stranded break at a specific site in the target DNA. Cas9 has two nuclease domains, HNH and RuvC, which it uses to cleave and degrade invading genetic elements. Cas9 can be reprogrammed to target and cleave any DNA sequence of interest through the simple engineering of the 20-nucleotide targeting sequence that guides Cas9 via Watson-Crick base pairing to the targeted locus of interest, which must be immediately followed by a 5'-NGG-3' protospacer-adjacent motif (PAM) that is essential for SpCas9 activity. Fusion of the crRNA to the tracrRNA led to the generation of a single, functional guide RNA (sgRNA) harboring the 20-nucleotide targeting sequence (spacer) at the 5' end (Hsu et al., 2014; Wright et al., 2016). The engineering of the sgRNA significantly simplified the system to a minimal set of two components; the Cas9 protein and the sgRNA,

creating a very simple, efficient genome engineering tool (Fig. 1 B, C). Notably, because the targeting specificity of this system depends on RNA:DNA base pairing complementarity between the engineered sgRNA and the target DNA, the ease of engineering the sgRNA overcame the inherent difficulty of protein engineering required for ZFs and TALENs (Eid and Mahfouz, 2016; Mahfouz et al., 2014). For this reason, development of the CRISPR/Cas9 system dramatically simplified, and thereby revolutionized, the field of genome engineering across prokaryotic and eukaryotic species. Subsequently, the CRISPR/Cas9 system has been adopted as a superior tool to efficiently edit the genomes of mammals (including humans), fungi, and plants, for a variety of applications.

2.3. CRISPR/dCas9:

Because of the unique features of the CRISPR/Cas system, it was possible to harness the system for further genome engineering purposes beyond targeted genome editing. Introducing catalytically inactivating mutations into the two nuclease domains of the Cas9 endonuclease, RuvC and HNH, resulted in the generation of a catalytically inactive variant of Cas9 (dCas9) that lacks DNA cleavage activity but retains the ability to specifically bind to target DNA sequences complementary to the sgRNA (Fig. 1A) (Jinek et al., 2012). The CRISPR/dCas9 system has been used to create highly efficient, programmable, genome-wide scale transcriptional regulators (Gilbert et al., 2014; Gilbert et al., 2013; Konermann et al., 2015). For example, in its simplest form, dCas9 can be targeted to bind to specific regions of gene promoters for use as a synthetic transcriptional repressor, referred to as CRISPR interference (CRISPRi). Consequently, binding of dCas9 would interfere with transcription factor or RNA polymerase binding, thus hindering transcriptional initiation and/or elongation and leading to the repression of gene expression (Qi et al., 2013). Furthermore, the DNA-specific binding ability of dCas9 enables the recruitment of various transcriptional regulators (when tethered to dCas9) to user-defined genomic loci for precise regulation of gene expression. Transcriptional repressors, such as Kruppel-associated box (KRAB) and Max-interacting protein 1 (Mxi1), have been fused to dCas9 for transcriptional repression in mammalian cells and yeast, respectively (Gilbert et al., 2013; Lawhorn et al., 2014). Additionally, transcriptional activation of endogenous or reporter genes has been achieved by fusing activation domains, such as the transcriptional activator VP64 or the

activation domain (AD) of p65, to dCas9 (Maeder et al., 2013a; Perez-Pinera et al., 2013a). We have recently employed the CRISPR/dCas9 system for targeted gene repression and activation in plant cells (Piatek et al., 2015). Specifically, we fused the SRDX repression domain to dCas9 to generate a reprogrammable synthetic repressor; we also fused the EDLL domain and the TAL activation domain (TAD) to dCas9 to generate synthetic transcriptional activators. In contrast to ZF- and TALE-based transcriptional regulators, the simplicity and effectiveness of this platform has enabled the multiplexed targeting of multiple transcriptional regulators to either the same gene, for improved regulatory activity due to the synergistic effect of multiple regulators (Maeder et al., 2013a; Perez-Pinera et al., 2013a), or to different genes, for the simultaneous regulation of entire gene networks (Zalatan et al., 2015). Importantly, such multiplexed recruitment of different transcriptional regulators will facilitate sophisticated transcriptional reprogramming by simultaneously activating and repressing different sets of genes in the same cell, thus enabling a broad set of applications for customized modulation of transcription (Puchta, 2016).

3. Post-transcriptional regulation at the RNA level, targeting the RNA message:

The aforementioned technologies represent robust methods for regulating gene expression at the level of DNA transcription. DNA-based transcriptional regulation must overcome several hurdles to achieve the desired outcome. For example, the compact structure of eukaryotic genomes and the presence of DNA-bound proteins may inhibit dCas9 binding to the targeted gene. Furthermore, since alternative splicing is crucial for the diversity of the eukaryotic transcriptome and proteome, repressing gene expression at the DNA level leads to the elimination of all possible splicing isoforms, thus hampering efforts to probe the functions of certain splice variants and causing unpredicted phenotypic consequences. Given the widespread functions of RNA in different cellular processes, it is increasingly important to manipulate endogenous RNAs in a specific and precise manner. Effective technologies for targeting and controlling the activity of cellular transcripts at the post-transcriptional level are presently lacking. However, emerging CRISPR/Cas-based platforms and technologies hold enormous potential to transform the field of RNA biology.

3.1. RNA interference:

More than two decades ago, the natural phenomenon of RNA silencing was discovered. Originally, a surprising gene-silencing phenomenon, termed co-suppression, was observed in plants. The introduction of transgenic copies of the *Chalcone Synthase (CHS)* gene in petunia flowers led to silencing of both the transgene and the endogenous gene (Napoli et al., 1990; van der Krol et al., 1990). Later, a similar phenomenon, “quelling,” was reported in *Neurospora crassa* (Romano and Macino, 1992) and subsequently in many other organisms (Gaudet et al., 1996; Pal-Bhadra et al., 1997; Ruiz et al., 1998). In 1998, RNA interference (RNAi) was first revealed in nematodes as a biological response to double-stranded RNA (dsRNA), which leads to sequence-specific gene silencing (Fire et al., 1998). Working with *Caenorhabditis elegans*, Fire and Mello demonstrated that dsRNAs mediate robust and specific silencing of homologous genes, as opposed to previously used sense or antisense ssRNAs (Fire et al., 1998). RNAi is an innate mechanism that protects against endogenous and exogenous parasitic genetic elements and regulates the expression of endogenous genes (Hannon, 2002). The silencing relies on small RNAs to guide cellular machinery to homologous sequences for degradation. These small interfering RNAs (siRNAs) originate from long precursor dsRNAs that can be derived from endogenously encoded sequences (e.g.: microRNAs), or from exogenous sources such as viruses or experimentally introduced nucleic acids (Agrawal et al., 2003; Hannon, 2002; Pumplin and Voinnet, 2013; Voinnet, 2005). In the cytoplasm, dsRNAs associate with the RNA-induced silencing complex (RISC) where the two strands are separated and one strand directs the complex to homologous regions in target messenger RNAs, leading to the degradation or inhibition of mRNA translation and specific gene silencing (Fig. 2) (Meister and Tuschl, 2004). The discovery of RNAi opened up a new frontier in RNA regulation and has been exploited as an invaluable technology for gene silencing in various fields and applications (Agrawal et al., 2003; Dorsett and Tuschl, 2004). The substantial advantage of RNAi technology lies in its specificity, which is due to Watson-Crick base pairing interactions between the experimentally introduced dsRNA and the target RNA. This specificity can be controlled and reprogrammed, making the technology both simple to use and effective against an almost unlimited choice of targets. Although RNAi has long been used as a powerful tool for specific gene silencing,

concerns about its stability, efficiency, specificity, and reproducibility have limited its use in many biological systems and downstream applications (Deng et al., 2014; Jackson et al., 2003). Moreover, the RNAi machinery must be conserved in the biological system of interest.

The development of synthetic programmable RNA targeting modules with a high degree of modularity, binding affinity, and specificity, as well as straightforward and predictable targeting will potentially facilitate many new applications in fields such as synthetic biology and genomic medicine, from endogenous RNA detection and trafficking, to manipulation and analysis. The significant potential of such synthetic programmable RNA targeting regulators has recently attracted much interest in developing customizable RNA-targeting modules.

3.2. Sequence-specific RNA-binding proteins:

In cells, RNAs are usually associated with diverse, yet specific, RNA-binding proteins (RBPs), forming ribonucleoproteins. These proteins are important for various RNA processes within cells, such as pre-mRNA processing, RNA stability, biogenesis, transport, function, cellular localization, and catalytic activities (Glisovic et al., 2008). RBPs bind RNA with different affinities and specificities by RNA-binding domains (RBDs), including K homology domains (KH), RNA recognition motif domains (RRM), pentatricopeptide repeats (PPR), cold shock domains (CSDs), zinc-finger domains (ZnF), and Pumilio/FBF repeats (PUF) (Auweter et al., 2006). Elucidation of the mechanisms that dictate the binding specificity of RBPs to their RNA substrates has prompted research efforts towards the development of programmable RNA regulators for sequence-specific RNA manipulation. These efforts have been based on engineering the specificity of the RNA-binding domains to generate modular RBPs with customized sequence specificity (Chen and Varani, 2013). PUF proteins, for instance, have been successfully engineered for the targeted manipulation of a variety of transcripts (Cheong and Hall, 2006; Wang et al., 2002). PUF proteins contain sequence-specific RNA-binding domains that each recognize a single nucleotide. Therefore, by combining multiple domains, PUF proteins can be designed to recognize any user-defined RNA target sequence with high affinity and specificity (Cheong and Hall, 2006; Filipovska et al., 2011; Wang et al., 2002). Furthermore,

ongoing research efforts aimed at elucidating the mechanisms governing RNA binding specificity of other RNA-binding domains, such as the PPR, RRM, and KH domains, suggest a potential use for these RNA-binding domains as programmable RNA-binding proteins (Chen and Varani, 2013; Coquille et al., 2014; Wei and Wang, 2015).

Similar to the utilization of ZFs and TALEs as programmable DNA-binding scaffolds, engineered RNA binding domains can be harnessed as programmable RNA-binding scaffolds to recruit effector domains for the targeted manipulation of endogenous RNAs. Designer RBPs, such as PUF proteins, have been fused to various effector domains generating chimeric artificial RBPs with a broad range of activities, including modulation of alternative splicing (Dong et al., 2011; Wang et al., 2009), site-specific RNA cleavage (site-specific RNA endonucleases) (Choudhury et al., 2012), targeted translational regulation (Cooke et al., 2011), and RNA imaging (Tilsner et al., 2009).

Although the development of designer RBPs has provided new insight into functional RNA biology for the manipulation and study of RNA, a few complications have constrained and limited the utility of these RBPs. RNA-binding domains commonly recognize only short sequences, enabling them to target a wide spectrum of endogenous RNA targets and limiting their specificity (Maris et al., 2005). This limits the ability of RNA-binding domains to target unique RNAs within the transcriptome. For example, PUF proteins can only recognize eight adjacent bases, which leads them to recognize a wide range of transcripts (Galgano et al., 2008; Morris et al., 2008). Nevertheless, the specificity of RBDs has been increased by tethering multiple domains with various structural arrangements together, generating proteins with high RNA binding affinity and specificity (Filipovska et al., 2011; Lunde et al., 2007). Because the specificity of the RBDs is based on protein–RNA interactions, analogous to the interaction of ZFs and TALEs with DNA, it is both challenging and laborious to engineer RBDs that bind to specific, user-defined RNA sequences. Therefore, it remains difficult to design and synthesize large protein libraries with varying sequence specificities for new targets. Additionally, the challenges associated with assembling modular RNA-binding proteins prohibits their use in the multiplex targeting of multiple RNA targets for large-scale screens and gene and/or protein network interrogation. Thus, there is still an enormous need for the development of synthetic

RNA regulators that are easy to program and versatile enough for both single and multiplexed targeting of RNAs.

3.3. CRISPR/Cas-based RNA targeting:

Due to the ability of CRISPR/Cas systems to rely on Watson-Crick base pairing for target recognition, they have overcome the significant drawbacks of purely protein-based genome editing tools, democratizing the field of genome engineering (Hsu et al., 2014). Similarly, programmable RNA targeting proteins based on nucleic acid complementarity can abolish the need to redesign and synthesize libraries of RNA-binding proteins. Thus, harnessing CRISPR/Cas systems to target RNA in a sequence-specific manner will enable the regulation of specific genes and precise transcriptome engineering.

3.3.1. Type III CRISPR/Cas systems:

While most of the currently known CRISPR/Cas systems target DNA, the Class I type III CRISPR/Cas systems, which can be further subdivided into subtype III-A and subtype III-B systems, can target invading RNAs in addition to DNA (Hale et al., 2012; Jiang et al., 2016; Samai et al., 2015; Staals et al., 2013; Staals et al., 2014; Tamulaitis et al., 2014). Detailed reconstitution and characterization of the effector complexes of the type III CRISPR/Cas systems from *Pyrococcus furiosus* and *Thermus thermophilus* have been shown *in vitro* (Hale et al., 2009; Sinkunas et al., 2013). However, unlike Class II CRISPR/Cas systems that rely on a single multi-domain effector protein to mediate the interference, type III CRISPR/Cas systems rely on complexes of multiple Cas proteins (Hale et al., 2009; Makarova et al., 2015). These multicomponent complexes would be cumbersome to reconstitute *in vivo* and harness as synthetic programmable RNA targeting platforms.

3.3.2. The RCas9 system:

Although Cas9 from *Streptococcus pyogenes* (SpCas9) has been widely utilized to enable genome editing by recognizing and targeting dsDNA, the ease with which it can be reprogrammed and its robust catalytic activity have encouraged researchers to repurpose it for

RNA targeting and manipulation. In 2014, work by O'Connell et al. demonstrated that SpCas9 can bind and cleave ssRNA *in vitro* (O'Connell et al., 2014). An essential requirement for SpCas9 recognition and cleavage activity is the presence of the 5'-NGG-3' PAM juxtaposed with the DNA target sequence on the non-target strand (Jinek et al., 2012; Mojica et al., 2009; Sternberg et al., 2014). When the PAM sequence was exogenously provided in specially designed ssDNA oligonucleotides (named PAMmers) complementary to the sequence adjacent to the target sequence in the ssRNA target, SpCas9 in association with its cognate targeting crRNA was successfully reprogrammed to bind and cleave specific target RNAs with high affinity and specificity (O'Connell et al., 2014). Intriguingly, the system exclusively targeted RNA, avoiding the corresponding DNA *in vitro* (Fig. 1D). The importance of the deoxyribonucleotide-based PAM was confirmed by the ability of SpCas9 to cleave an RNA–DNA hybrid but not dsRNA. In addition, this work also described the utilization of this so-called RNA targeting Cas9 (RCas9) as a specific programmable RNA binding platform. By using the nuclease-inactive RCas9 system, dCas9:gRNA with a PAMmer, researchers were able to isolate and pull down endogenous, untagged *Glyceraldehyde-3-phosphate dehydrogenase* mRNA from HeLa cell lysate (O'Connell et al., 2014). Interestingly, in a recent study, Nelles et al. demonstrated the activity of the RCas9 system *in vivo* by successfully applying the system to image and track the trafficking of specific endogenous mRNAs in living cells using a chimeric dCas9-GFP fusion (Nelles et al., 2016).

These studies shed light on the unexpected, yet exciting ability of SpCas9 to target RNA in addition to DNA, illustrating the promise of harnessing CRISPR/Cas systems as synthetic programmable RNA targeting modules. Nevertheless, a few concerns could impede the utilization of the RCas9 system for RNA biology-related applications in living organisms. The activity of the RCas9 system relies on three components (SpCas9, gRNA, and PAMmer); efficient delivery and complex formation of these components could prove problematic. Furthermore, chemical modifications are needed to stabilize the PAMmers and prevent their degradation by RNase H in living cells (Nelles et al., 2016; O'Connell et al., 2014). Therefore, such specially designed oligonucleotides could be costly and difficult to deliver into target cells for large-scale manipulation. Moreover, the ability of RCas9 to target and cleave RNA would make it feasible to target this activity to the nucleus of living eukaryotic cells by taking

advantage of U3 or U6 promoter-transcribed sgRNAs. However, while the *in vivo* study described the activity of only the nuclease-deficient Cas9 for RNA binding purposes (Nelles et al., 2016), and RCas9 seems not to cleave DNA *in vitro* (O'Connell et al., 2014), the native DNA cleavage activity of NLS-tagged, nuclease-active SpCas9 could lead to off-target effects, potentially resulting in deleterious mutations.

3.3.3. FnCas9:

FnCas9, an orthologue of SpCas9, is a type II CRISPR-associated protein effector encoded by the bacterial pathogen *Francisella novicida* (Hirano et al., 2016). Similar to other Cas9 orthologues, such as SpCas9 (Jinek et al., 2012) and the Cas9 of *Staphylococcus aureus* (SaCas9) (Ran et al., 2015), FnCas9 in association with a crRNA:tracrRNA or a synthetic gRNA has been shown to mediate DNA cleavage by recognizing a 5'-NGG-3' PAM immediately downstream of the target sequence, and was successfully harnessed for genome engineering in mouse zygotes, but not in human cells (Hirano et al., 2016). Interestingly, Sampson *et al.* reported an innate RNA-targeting activity of FnCas9 in bacteria (Sampson et al., 2013). To evade the bacterial lipoprotein (BLP)-triggered-host immune response and invade eukaryotic host cells, *F. novicida* is able to employ the CRISPR/Cas9 system to target and degrade its own *BLP* mRNA, leading to suppression of *BLP* expression and promoting virulence (Jones et al., 2012; Sampson et al., 2014). Besides the crRNA and tracrRNA, the *F. novicida* CRISPR/Cas9 system also contains a small RNA termed small CRISPR/Cas associated RNA (scaRNA), which can base pair with the tracrRNA to form a heteroduplex RNA. In contrast to the crRNA:tracrRNA-dependent DNA cleavage by FnCas9, the scaRNA:tracrRNA duplex promotes RNA targeting when associated with FnCas9 (Sampson et al., 2013). This novel activity of FnCas9 led to it being used to target viral RNA in eukaryotic cells. As a proof of principle, Price and colleagues used FnCas9 to target the positive-sense single-stranded RNA genome of the hepatitis C virus (HCV) in human hepatocellular carcinoma cells (Price et al., 2015). By engineering the tracrRNA and scaRNA to generate a single RNA-targeting guide RNA (rgRNA), a FnCas9:rgRNA complex targeted to either the highly conserved 5' or 3' untranslated regions of the HCV genome resulted in an ~60% reduction in viral protein production. The RNA-targeting activity of FnCas9:rgRNA was PAM independent, in contrast to the PAM dependent RNA

targeting by the RCas9 system (Price et al., 2015). These studies highlight the potential of the FnCas9 CRISPR system to efficiently and specifically target RNA *in vivo*.

However, the mechanism by which FnCas9 targets and cleaves RNAs remains largely unknown (Hirano et al., 2016; Sampson et al., 2013). For instance, although wild-type FnCas9 was used for viral RNA targeting, the inhibition of HCV was presumably caused by the binding of FnCas9 to the viral RNA genome and blocking viral translation and replication machineries, rather than degradation of the HCV genome, as endonuclease-inactive FnCas9 produced a similar inhibitory effect (Price et al., 2015). Thus, further investigations are needed to elucidate the FnCas9 RNA-targeting activity. Moreover, like the RCas9 system, the ability of FnCas9 to target and cleave endogenous genomic DNA will restrict its utility for the manipulation of RNA in nuclei.

3.3.4. CRISPR/Cas13 systems:

The great potential offered by the efficiency and simplicity of the CRISPR/Cas system as a genome engineering tool has inspired researchers to improve the existing CRISPR/Cas tools, and to develop new methodologies that enable them to delve deeper into the microbial world and find other undiscovered CRISPR/Cas variants that can be harnessed and added to the genome-engineering toolbox (Burstein et al., 2017; Makarova et al., 2015; Shmakov et al., 2015). In a search for previously unexplored Class II CRISPR/Cas systems using microbial genome data mining, along with computational and bioinformatic prediction approaches, Shmakov et al. discovered novel Class II CRISPR systems, including C2c1, C2c2, and C2c3 (Shmakov et al., 2015). C2c1 and C2c3 contain RuvC-like endonucleases similar to Cpf1, and were therefore classified as Class II type V CRISPR/Cas systems (Shmakov et al., 2015), (now V-B Cas12b and V-C Cas12c, respectively) (Shmakov et al., 2017). By contrast, the Class II candidate 2 (C2c2) exhibited unique features not present in any known CRISPR protein, leading to it being classified as a new Class II subtype, Class II type VI (Shmakov et al., 2015) (now VI-A Cas13a)(Shmakov et al., 2017).

3.3.4.1. Cas13a

Analysis of the Cas13a protein sequence revealed the presence of two Higher Eukaryotes and Prokaryotes Nucleotide-binding Domains (HEPN), which are exclusively associated with RNase activity (Anantharaman et al., 2013). These distinct characteristics of Cas13a raised the enticing possibility that Cas13a might work as a single effector RNA-guided RNA-targeting protein.

A pioneering study characterizing the functionality of Cas13a showed that the single effector Cas13a protein is, indeed, a programmable RNA-guided ssRNA ribonuclease (Abudayyeh et al., 2016). In this study, the RNA targeting and interference activity of *Leptotrichia shahii* Cas13a (LshCas13a) was first explored by performing a bacteriophage interference screen, which demonstrated the ability of the Cas13a protein and phage-genome specific 28-nt spacer sequences to defend *E. coli* against infection by the ssRNA phage MS2. Besides identifying the sequences of the gRNAs that interfered with the ssRNA phage genome most efficiently, the phage interference screen results hinted at the presence of an H (non-G) protospacer flanking site (PFS) immediately following the targeted protospacers (Fig. 3C) (Abudayyeh et al., 2016). The non-G PFS has turned out to be crucial for maintaining the interaction between the Cas13a protein and its cognate crRNA (Liu et al., 2017b). Further characterization of the RNA cleavage activity of Cas13a using purified Cas13a protein for *in vitro* cleavage assays showed that Cas13a is specific for ssRNA targets. Moreover, Cas13a tends to preferentially cleave uracil residues at multiple sites in exposed regions of the secondary structure formed by the ssRNA. Additionally, mutating the putative histidine and arginine catalytic residues within the two HEPN domains abolished the cleavage activity of the protein, indicating that the HEPN domains are responsible for the RNA cleavage activity and resulting in the generation of a catalytically inactive version of the Cas13a enzyme (dCas13a) (Fig. 3A). However, like dCas9, dCas13a retains its ability to specifically bind to the target RNA, resulting in an RNA-guided RNA-binding protein. Cas13a can tolerate single- but not double-base mismatches in the middle of the spacer-protospacer base pairing, pointing to the presence of a seed sequence. Importantly, experiments that successfully knocked down the expression of the RFP protein demonstrated the ability of this system to be reprogrammed to target specific, non-phage RNAs *in vivo* (Abudayyeh et al., 2016). The fundamental findings of this study have raised the possibility that Cas13a is the long-awaited programmable, RNA-guided ssRNA targeting CRISPR system.

A subsequent study investigating the biochemical features of LshCas13a and other Cas13a homologs, such as LbuCas13a (from *Leptotrichia buccalis*) and LseCas13a (from *Listeria seeligr*), has revealed another enzymatic activity, adding to the novelty of Cas13a (East-Seletsky et al., 2016). This study found that Cas13a possesses a ribonuclease activity that is responsible for the processing of precursor crRNA (pre-crRNA) to generate its cognate mature crRNA, demonstrating that Cas13a is a dual ribonuclease (Fig. 4) (East-Seletsky et al., 2016). The crystal structural of LshCas13a revealed that the two RNase activities, the pre-crRNA processing and crRNA-guided RNA cleavage activity, are executed by two distinct and physically separated catalytic sites. A group of positively charged residues located in the N-terminal helical-1 domain within the REC lobe are responsible for the pre-crRNA biogenesis activity. By contrast, the RNA-guided target RNA cleavage takes place in a catalytic site formed between the two conserved HEPN domains (Liu et al., 2017b). A recent study by East-Seletsky et al. (East-Seletsky et al., 2017) found that the pre-crRNA processing activity is highly conserved among diverse type VI-A CRISPR/Cas13a family proteins. However, it appears that while the pre-crRNA processing activity of Cas13a improves ssRNA targeting, it is not essential (East-Seletsky et al., 2017). In addition, functional and biochemical characterization of various Cas13a proteins has shown that Cas13a proteins can be classified into two functional subfamilies that can recognize distinct crRNAs, resulting in differing substrate preferences. These findings expand the potential utility of Cas13a proteins for a wide range of RNA manipulations (East-Seletsky et al., 2017).

Intriguingly, upon activation of the Cas13a enzyme by binding to a crRNA-complementary ssRNA target, non-specific trans-ssRNA cleavage was observed; the collateral degradation of other ssRNAs (Abudayyeh et al., 2016; East-Seletsky et al., 2016). Analysis of the crystal structure of LbuCas13a revealed that binding of the target RNA to the Cas13–crRNA complex, and formation of the crRNA–target RNA duplex, induces a significant conformational change in the Cas13a protein, leading to the activation of the HEPN catalytic site (Liu et al., 2017a). The induced conformational change results in the formation of a guide–target RNA duplex binding channel with the Cas13a protein. The resulting binding channel is located far from the HEPN catalytic site, suggesting that short target RNAs cannot be cleaved by the HEPN catalytic site in *cis*. Importantly, the structural analysis also showed that the HEPN catalytic site of the activated

Cas13 protein is exposed to the surface, making it available to RNAs in solution, and thus explaining the non-specific cleavage of collateral RNAs *in trans* observed upon target RNA binding (Liu et al., 2017a). This phenomenon suggests a natural mechanism to sense invasive viral RNAs and induce programmed cell death (PCD) or dormancy of the host cell, thus preventing the spread of the viral infection (Fig. 4). The East-Seletsky et al. study (East-Seletsky et al., 2017) demonstrated that the collateral cleavage activity of Cas13a could be used to sense and detect the presence of specific transcripts. Building on this idea, a study by Gootenberg and colleagues exploited the promiscuous RNase activity of Cas13a upon target recognition to develop a diagnostic tool for the *in vitro* detection of DNA and RNA with a single-base mismatch specificity and attomolar sensitivity, demonstrating a wide range of potential utility in diagnostic and basic research applications (Fig. 5H) (Gootenberg et al., 2017).

This series of studies illuminating the functional mechanisms of different Cas13a orthologues in prokaryotes has been recently followed by the first experimental evidence demonstrating the amenability of CRISPR/Cas13a to adaptation for RNA targeting in eukaryotic cells (Abudayyeh et al., 2017). By assessing 15 different Cas13a proteins, Abudayyeh and colleagues identified LwaCas13a from *Leptotrichia wadei* as having the greatest interference activity and specificity relative to other Cas13a orthologues, including the previously characterized LshCas13a. Heterologous expression of LwaCas13a and the mature crRNA produced a robust knockdown of either reporter or endogenous transcripts in mammalian and plant cells. Cas13a exhibited a comparable interference activity to that produced by RNAi, but with greater specificity and dramatically reduced off-target effects. The study also tested the ability of LwaCas13a to process pre-crRNAs to generate multiple functional crRNAs that would allow multiplexed targeting. Upon delivery of a CRISPR array consisting of five different guide RNAs that targeted five different endogenous genes, the authors observed levels of knockdown for each gene comparable to those produced using single guide RNAs, indicating that Cas13b can process pre-crRNA for generation of functional crRNAs and multiplexed targeting in mammalian cells (Abudayyeh et al., 2017).

3.3.4.2. Cas13b

Another recently discovered and characterized RNA-targeting CRISPR/Cas system is Cas13b, a member of the Class II subtype VI-B CRISPR/Cas systems (Smargon et al., 2017). Although the

Cas13b effector protein was found to have a novel protein sequence that differs significantly from Cas13a, the sequence showed two predicted HEPN domains, suggesting that the protein targets and cleaves RNA, similar to subtype VI-A (Cas13a). Functional characterization of Cas13b by Smargon et al. revealed a sequence-specific single-stranded RNA targeting of Cas13b *in vitro* and *in vivo*. The RNA targeting of Cas13b is similar to the activity of Cas13a in many ways, including the ability to target and cleave ssRNA but not dsRNA, to process its own pre-crRNA to generate mature crRNAs, and to cause non-specific collateral RNA damage of non-target RNAs upon activation by binding to target RNA. In addition, a nuclease-deficient variant of Cas13b, dCas13b, can bind specifically to a target sequence. However, unlike Cas13a, the RNA targeting of Cas13b requires a double-sided PFS on both ends of the protospacer target sequence, which differ in their nucleotide sequence from the PFS of Cas13a, thus expanding the sequence targeting constraints of these RNA-targeted CRISPR/Cas systems. Another interesting finding was the ability of Cas13b to interact with two small and novel proteins, Csx27 and Csx28, where *in vivo* investigation of these small proteins showed that Csx27 represses the RNA targeting and cleavage of Cas13b and Csx28 enhances it (Smargon et al., 2017).

Notably, a very recent study reported that Cas13b ribonucleases can specifically knock down endogenous transcripts in mammalian cells (Cox et al., 2017). Although the previous study identified LwaCas13a as the most effective Cas13a orthologue for targeted knockdown of transcripts in mammalian cells (Abudayyeh et al., 2017), Cox et al. sought to identify a more robust and specific Cas13 ribonuclease for mammalian cell applications (Cox et al., 2017). Therefore, they evaluated the RNA interference activity of a subset of Cas13 enzymes, including 21 orthologs of Cas13a, 15 of Cas13b, and 7 of Cas13c, and found that PspCas13b (from *Prevotella* sp. P5-125) was the most efficient Cas13 orthologue. Apart from the demonstrated RNA editing ability of the dCas13b-ADAR fusions (which we discuss under the potential applications of Cas13 systems), PspCas13b exhibited a robust and increased level of specificity and knockdown activity compared to the previously characterized LwaCas13a (Cox et al., 2017).

The ability of Cas13b to exhibit specific and robust targeting of RNA substrates represents a useful addition to the RNA-targeted CRISPR/Cas suite, and indicates that additional potential

RNA-targeting CRISPR/Cas systems await, to be found and harnessed for targeted RNA manipulation and interference.

4. Advantages and potential applications of Cas13 systems for post-transcriptional regulation:

Collectively, the findings of these studies suggest that Cas13 is a flexible, RNA-guided, RNA targeting CRISPR systems that hold great potential for precise, robust, and scalable RNA-guided transcriptional regulation applications. The significance of modulating transcriptional regulation led to the emergence of various approaches for transcriptional and post-transcriptional regulation (Brophy and Voigt, 2014; Mohr et al., 2014). However, the development of Cas13 as the first RNA-specific CRISPR technology might provide the long-sought tool for versatile and efficient RNA targeting that will open a new realm in the field of post-transcriptional regulation and offer new strategies for diverse RNA manipulations.

The CRISPR/Cas13 systems demonstrate many unprecedented advantages that render the systems superior to other previously developed RNA-targeting strategies. They consist of only two components, the Cas13 protein and the crRNA that guides Cas13 to its target RNA sequence (Abudayyeh et al., 2016), facilitating its delivery and assembly in most organisms (Fig. 3B). The targeting specificity of the system relies on the spacer sequence in the guide crRNA, thereby offering the versatility and scalability of RNAi technology alongside with the modularity of the RNA-binding proteins. In addition, the capability of Cas13 to process its own crRNA from precursor crRNA transcripts could be utilized to facilitate its employment for important cytoplasmic RNA manipulations, such as mRNA imaging, localization, and translational regulation. For example, the utilization of polymerase II promoters for guide RNA expression will allow the transport of pre-crRNA with the desired targeting spacers into the cytoplasm where they can be processed by Cas13 to generate mature crRNAs, sidestepping the limitations of using the small nuclear RNA (snoRNA) promoters, such as U6 or U3, for applications in the nucleus. Additionally, the fact that Cas13 can process its own crRNA will facilitate the design and expression of multiple gRNAs for targeting and modulating multiple transcripts (Abudayyeh et al., 2017; East-Seletsky et al., 2016) (Table1).

An obvious application of Cas13 is to use its innate endoribonuclease activity to knock down gene expression by cleaving specific transcripts (Abudayyeh et al., 2017) (Fig. 5A). Unlike RNAi, Cas13 gene knockdown is independent of the host's RNAi machinery, enabling precise gene knockdown in organisms or cellular compartments where RNAi machinery is not present. The simple engineering of the guide (crRNA) RNA will also allow the generation of large gRNA libraries to downregulate expression of multiple genes simultaneously, facilitating large-scale screens and the interrogation of gene networks. Additionally, Cas13 will be valuable for targeting and degrading specific splicing isoforms, leading to isoform-specific gene silencing. Previous studies have developed restriction enzyme-like site-specific RNA endonucleases by combining RNA cleavage domains with site specific RNA binding domains in attempts to generate ribonucleases to specifically silence endogenous genes (Choudhury et al., 2012). Such artificial ribonucleases have proven to be efficient in silencing specific pathogenic mRNAs in human cells (Zhang et al., 2014). The simplicity and single-base mismatch specificity of Cas13 (Gootenberg et al., 2017) can now allow for the precise targeting of certain aberrant or pathogenic mRNAs and specific splicing isoforms in plants and animals without affecting the wild-type transcripts. Nevertheless, the usefulness of Cas13 would be severely limited if collateral RNA cleavage is also elicited in eukaryotic cells. Surprisingly, a recent study (Abudayyeh et al., 2017) did not find evidence of Cas13a-induced collateral activity, finding that the growth of mammalian cells expressing active LwaCas13a was unaffected, and thus suggesting that the promiscuous activity of Cas13a might be absent or negligible in eukaryotic cells. The structural studies of Cas13 may provide insight into engineering better versions of Cas13 that have more specificity and lack promiscuous cleavage activity. In addition, ongoing research may lead to the identification of different Cas13 variants with controlled and robust catalytic activity, as other Cas13 homologs, such as LwaCas13a and PspCas13b, were shown to have greater RNase activity than LshCas13a (Abudayyeh et al., 2017; Cox et al., 2017; Gootenberg et al., 2017).

Beyond the RNA cleavage activity of this system, the catalytically inactive Cas13 (dCas13) represents an attractive programmable, sequence-specific RNA binding platform that can be tethered to various functional effectors for a wide range of important molecular manipulations.

Intracellular RNA visualization and tracking in living cells can provide valuable information on the dynamics of endogenous RNAs. Several RNA-imaging approaches have been developed based on the fusion of fluorescent proteins to sequence specific RNA binding domains, such as the MS2 coat protein (MCP) approach (Bertrand et al., 1998), Pumilio homology domains (PUM-HD) (Ozawa et al., 2007; Yamada et al., 2011), and recently, the RCas9 system (Nelles et al., 2016). However, fusion of dCas13 to a fluorescent protein will generate a platform that is highly versatile and easier to program for visualizing the localization and trafficking of specific RNAs in living cells (Fig. 5B). In fact, the recent study of Abudayyeh et al. harnessed the programmable binding ability of dCas13 to engineer a visualization platform capable of tracking the translocation of endogenous transcripts from the nucleus to the cytoplasm, thus providing an efficient programmable platform for live transcript imaging and tracking in mammalian cells (Abudayyeh et al., 2017). Another aspect of post-transcriptional RNA modulation and spatial control of gene expression is RNA localization, which is involved in various important developmental and cellular physiological processes (Martin and Ephrussi, 2009; St Johnston, 2005). By tethering a trafficking agent to dCas13, CRISPR/dCas13 can be engineered to transport targeted RNAs to a desired cellular location. Another way of controlling gene expression at the post-transcriptional level is through the regulation of RNA translation and stability. Translational activators, such as GLD2 and eIF4E, and repressors, such as CAF1, have previously been combined with designer RNA binding proteins, PUF proteins for instance, to specifically recognize and regulate the translation of targeted RNAs (Cao et al., 2014; Cooke et al., 2011; Quenault et al., 2011). Therefore, dCas13 could be combined with translational regulators to boost or repress gene expression without affecting the abundance of the endogenous transcript (Fig. 5C). Another elegant potential application of the CRISPR/dCas13 system would be to identify RNA binding proteins by targeting and pulling down their RNA substrates, RNA immunoprecipitation sequencing (RIP-seq), using tagged-dCas13 (Fig. 5D). Such studies would aid in defining and understanding the function of many RNA regulators.

The composition of endogenous RNAs can also be edited at the pre-mRNA level. For example, alternative splicing of precursor mRNAs is one of the critical steps in post-translational gene regulation in eukaryotic cells, and RNA mis-splicing can lead to severe consequences in humans and plants (Scotti and Swanson, 2016; Syed et al., 2012). Thus, the ability to precisely

manipulate alternative splicing of pre-mRNAs would be valuable for many therapeutic and biotechnological applications. Targeting splicing factors to certain sequences in the pre-mRNA has been demonstrated to alter the splicing pattern of specific mRNAs. For example, splicing activators, such as Arginine/Serine (RS)-rich domains, or suppressors, such as the Glycine (Gly)-rich domain of hnRNP, have been fused to designer RNA proteins like PUF and MS2 for targeted splicing alteration (Graveley and Maniatis, 1998; Wang et al., 2009). The ease of reprogramming Cas13 along with the ability to perform multiplex targeting will facilitate guiding different splicing factors fused to dCas13 to exonic or intronic sequences of interest in different pre-mRNAs simultaneously, reversing the splicing defect associated with pre-mRNA mis-splicing (Fig. 5E). In addition, similar to the newly-developed CRISPR/Cas9-based base editors comprising dCas9 or Cas9 nickase fused to cytidine deaminase for targeted single nucleotide change in the DNA (Kim, K. et al., 2017; Kim, Y.B. et al., 2017; Komor et al., 2016; Nishida et al., 2016; Shimatani et al., 2017; Zong et al., 2017), targeted RNA editing such as adenosine-to-inosine (A-to-I) and cytosine-to-uracil (C-to-U) modification of mature mRNAs could be achievable by engineering dCas13-adenosine deaminase or dCas13-cytidine deaminase fusions guided by crRNAs to the ssRNA targets, leading to targeted mutagenesis at the RNA level without permanent modification of the genome (Fig. 5F) (Montiel-Gonzalez et al., 2013). This was recently accomplished by fusing the catalytic deamination domain of the adenosine deaminase acting on RNA (ADAR_{DD}) protein to catalytically inactive PspCas13b to generate an dCas13b-ADAR_{DD} fusion capable of programmable RNA targeting and editing (Cox et al., 2017). The newly engineered system, named REPAIR (RNA Editing for Programmable A to I Replacement), produced precise and efficient A to I edits on endogenous transcripts as well as pathogenic mutations, demonstrating the promise of programmable RNA editing as a new therapeutic platform. The successful engineering of the REPAIR system represents an insightful example of the tremendously exciting future of developing versatile Cas13-based RNA manipulation technologies. With the ability of dCas13 to specifically bind to any sequence of interest and by fusing the appropriate functional domains to dCas13, diverse tools capable of sequence-specific RNA manipulation can be developed to modulate RNA function at multiple steps of the RNA life cycle. This will enable a broad range of potential applications in basic research, biotechnology, and therapeutics.

4.1. RNA virus interference:

The impressive success of Cas9 in targeting DNA for genome engineering in eukaryotic cells has inspired researchers to exploit this technology as a programmable antiviral defense strategy to confer resistance to many eukaryotic viruses, including human viruses (Price et al., 2016). In plants, CRISPR/Cas9 has proven to efficiently confer viral resistance to host plants (Zaidi et al., 2016). We and others have shown that transforming *Nicotiana benthamiana* plants with CRISPR/Cas9 resulted in strong interference of various DNA geminiviruses, including *Tomato yellow leaf curl virus* (TYLCV), *Beet curly top virus* (BCTV), *Merremia mosaic virus* (MeMV), *Bean yellow dwarf virus* (BeYDV), and *Beet severe curly top virus* (BSCTV), demonstrating the enormous potential of CRISPR/Cas9 as a promising strategy against plant geminiviruses (Ali et al., 2015; Ali et al., 2016; Baltes, 2015; Ji et al., 2015). Nevertheless, the fact that many severe human viruses and the majority of plant viruses have RNA genomes constrains the utility of CRISPR/Cas9-mediated inhibition of RNA viruses. The significance of engineering resistance against eukaryotic RNA viruses has led to several attempts to develop efficient means of targeting them. As outlined above, FxCas9 was used to target the single-stranded RNA genome of hepatitis C virus (HCV) in eukaryotic cells (Price et al., 2015). In plants, instead of directly targeting viral nucleic acids, CRISPR/Cas9 was harnessed to generate resistance to plant RNA viruses via disrupting host genes critical for the viral life cycle, such as the eukaryotic translation initiation factors 4E (eIF4E) and eIF(iso)4E (Chandrasekaran et al., 2016; Pyott et al., 2016).

The remarkable efficiency and simplicity of the CRISPR/Cas13 system to confer immunity against RNA viruses in bacteria (Abudayyeh et al., 2016) may constitute the next generation of RNA antiviral strategies in eukaryotic systems. Combining transient or stable expression of Cas13 along with rational design of crRNAs targeting conserved regions in viral RNA genomes could lead to immunity against eukaryotic RNA viruses. Additionally, the CRISPR/Cas13 system can also be used to confer resistance to DNA viruses that have an RNA stage in their life cycle (Fig. 6). Similar to the natural function of CRISPR/Cas13, it is possible that multiplexing crRNAs in transgenic plants or animals could enable resistance to multiple pathogenic viruses simultaneously.

4.2. Induction of programmed cell death in response to specific transcripts:

The promiscuous RNase activity of Cas13, if portable to eukaryotic cells, could be exploited and harnessed for important *in vivo* applications. In cancer therapy, the ‘Holy Grail’ is to reduce the cytotoxicity of anticancer drugs by selectively targeting and killing tumor cells without affecting normal ones. Many cancers can be characterized by the expression of unique, aberrant biomarkers that distinguish them from healthy tissues. Accordingly, several cancer therapeutic strategies that selectively target these biomarkers have been designed. For instance, prodrugs (inert derivatives of pharmacologically active agents that undergo transformation *in vivo* to release the active drug (Rautio et al., 2008)) such as targeted protein toxins have been developed to target cancer specific biomarkers and kill cells by inducing cell death (Bachran et al., 2014; Boland et al., 2014). By targeting Cas13 to unique and aberrant transcripts expressed in tumor cells, one can envision that the collateral cleavage activity of Cas13 could lead to the induction of programmed cell death, resulting in the selective killing of only the cancerous cells (Fig. 5G).

Another conceivable application of the promiscuous RNase activity of Cas13 is to mediate virus interference in plants. The primary response of resistance gene-mediated resistance to viral infection includes PCD of the cells at the initial site of infection (localized necrotic lesion phenotype), known as the hypersensitive response, which confines the infection to limited lesions and prevents the spread into adjacent non-infected cells (Martin et al., 2003; Soosaar et al., 2005). By stably expressing Cas13 with multiple crRNAs targeting different viral genomes, Cas13 can be used to sense diverse plant RNA viruses. Consequently, once Cas13 recognizes the cognate virus target, its promiscuous activity could lead to localized cell death, leading to a similar localized necrotic phenotype at the infected region.

Acknowledgement

This study was supported by King Abdullah University of Science and Technology (KAUST).

References:

- Abudayyeh, O.O., Gootenberg, J.S., Essletzbichler, P., Han, S., Joung, J., Belanto, J.J., Verdine, V., Cox, D.B.T., Kellner, M.J., Regev, A., Lander, E.S., Voytas, D.F., Ting, A.Y., Zhang, F., 2017. RNA targeting with CRISPR-Cas13. *Nature* 550(7675), 280-284.
- Abudayyeh, O.O., Gootenberg, J.S., Konermann, S., Joung, J., Slaymaker, I.M., Cox, D.B., Shmakov, S., Makarova, K.S., Semenova, E., Minakhin, L., Severinov, K., Regev, A., Lander, E.S., Koonin, E.V., Zhang, F., 2016. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* 353(6299), aaf5573.
- Agrawal, N., Dasaradhi, P.V., Mohammed, A., Malhotra, P., Bhatnagar, R.K., Mukherjee, S.K., 2003. RNA interference: biology, mechanism, and applications. *Microbiol Mol Biol Rev* 67(4), 657-685.
- Ali, Z., Abulfaraj, A., Idris, A., Ali, S., Tashkandi, M., Mahfouz, M.M., 2015. CRISPR/Cas9-mediated viral interference in plants. *Genome Biol* 16, 238.
- Ali, Z., Ali, S., Tashkandi, M., Zaidi, S.S., Mahfouz, M.M., 2016. CRISPR/Cas9-Mediated Immunity to Geminiviruses: Differential Interference and Evasion. *Sci Rep* 6, 26912.
- Anantharaman, V., Makarova, K.S., Burroughs, A.M., Koonin, E.V., Aravind, L., 2013. Comprehensive analysis of the HEPN superfamily: identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing. *Biol Direct* 8, 15.
- Auweter, S.D., Oberstrass, F.C., Allain, F.H., 2006. Sequence-specific binding of single-stranded RNA: is there a code for recognition? *Nucleic Acids Res* 34(17), 4943-4959.
- Bachran, C., Hasikova, R., Leysath, C.E., Sastalla, I., Zhang, Y., Fattah, R.J., Liu, S., Leppla, S.H., 2014. Cytolethal distending toxin B as a cell-killing component of tumor-targeted anthrax toxin fusion proteins. *Cell Death Dis* 5, e1003.
- Baltes, N.J., 2015. Conferring resistance to geminiviruses with the CRISPR-Cas prokaryotic immune system. *Nature plants* 1(10), 15145.
- Barrangou, R., 2015. Diversity of CRISPR-Cas immune systems and molecular machines. *Genome Biol* 16, 247.
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., Horvath, P., 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315(5819), 1709-1712.
- Barrangou, R., Marraffini, L.A., 2014. CRISPR-Cas systems: Prokaryotes upgrade to adaptive immunity. *Mol Cell* 54(2), 234-244.
- Beerli, R.R., Barbas, C.F., 3rd, 2002. Engineering polydactyl zinc-finger transcription factors. *Nat Biotechnol* 20(2), 135-141.
- Beerli, R.R., Segal, D.J., Dreier, B., Barbas, C.F., 3rd, 1998. Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc Natl Acad Sci U S A* 95(25), 14628-14633.
- Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S.M., Singer, R.H., Long, R.M., 1998. Localization of ASH1 mRNA particles in living yeast. *Mol Cell* 2(4), 437-445.
- Bertrand, J.R., Pottier, M., Vekris, A., Opolon, P., Maksimenko, A., Malvy, C., 2002. Comparison of antisense oligonucleotides and siRNAs in cell culture and in vivo. *Biochem Biophys Res Commun* 296(4), 1000-1004.
- Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., Bonas, U., 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326(5959), 1509-1512.
- Bogdanove, A.J., Voytas, D.F., 2011. TAL effectors: customizable proteins for DNA targeting. *Science* 333(6051), 1843-1846.
- Boland, E.L., Van Dyken, C.M., Duckett, R.M., McCluskey, A.J., Poon, G.M., 2014. Structural complementation of the catalytic domain of pseudomonas exotoxin A. *J Mol Biol* 426(3), 645-655.
- Brent, R., Ptashne, M., 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* 43(3 Pt 2), 729-736.
- Brophy, J.A., Voigt, C.A., 2014. Principles of genetic circuit design. *Nat Methods* 11(5), 508-520.
- Burnett, J.C., Rossi, J.J., 2012. RNA-based therapeutics: current progress and future prospects. *Chem Biol* 19(1), 60-71.
- Burstein, D., Harrington, L.B., Strutt, S.C., Probst, A.J., Anantharaman, K., Thomas, B.C., Doudna, J.A., Banfield, J.F., 2017. New CRISPR-Cas systems from uncultivated microbes. *Nature* 542(7640), 237-241.
- Cao, J., Arha, M., Sudrik, C., Schaffer, D.V., Kane, R.S., 2014. Bidirectional regulation of mRNA translation in mammalian cells by using PUF domains. *Angew Chem Int Ed Engl* 53(19), 4900-4904.
- Chandrasekaran, J., Brumin, M., Wolf, D., Leibman, D., Klap, C., Pearlsman, M., Sherman, A., Arazi, T., Gal-On, A., 2016. Development of broad virus resistance in non-transgenic cucumber using CRISPR/Cas9 technology. *Mol Plant Pathol* 17(7), 1140-1153.
- Chen, Y., Varani, G., 2013. Engineering RNA-binding proteins for biology. *FEBS J* 280(16), 3734-3754.
- Cheong, C.G., Hall, T.M., 2006. Engineering RNA sequence specificity of Pumilio repeats. *Proc Natl Acad Sci U S A* 103(37), 13635-13639.
- Choudhury, R., Tsai, Y.S., Dominguez, D., Wang, Y., Wang, Z., 2012. Engineering RNA endonucleases with customized sequence specificities. *Nat Commun* 3, 1147.
- Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., Zhang, F., 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339(6121), 819-823.
- Cooke, A., Prigge, A., Opperman, L., Wickens, M., 2011. Targeted translational regulation using the PUF protein family scaffold. *Proc Natl Acad Sci U S A* 108(38), 15870-15875.

- Coquille, S., Filipovska, A., Chia, T., Rajappa, L., Lingford, J.P., Razif, M.F., Thore, S., Rackham, O., 2014. An artificial PPR scaffold for programmable RNA recognition. *Nat Commun* 5, 5729.
- Cox, D.B.T., Gootenberg, J.S., Abudayyeh, O.O., Franklin, B., Kellner, M.J., Joung, J., Zhang, F., 2017. RNA editing with CRISPR-Cas13. *Science*.
- Cronin, C.A., Gluba, W., Scrable, H., 2001. The lac operator-repressor system is functional in the mouse. *Genes Dev* 15(12), 1506-1517.
- Deng, Y., Wang, C.C., Choy, K.W., Du, Q., Chen, J., Wang, Q., Li, L., Chung, T.K., Tang, T., 2014. Therapeutic potentials of gene silencing by RNA interference: principles, challenges, and new strategies. *Gene* 538(2), 217-227.
- Dominguez, A.A., Lim, W.A., Qi, L.S., 2016. Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nat Rev Mol Cell Biol* 17(1), 5-15.
- Dong, S., Wang, Y., Cassidy-Amstutz, C., Lu, G., Bigler, R., Jezyk, M.R., Li, C., Hall, T.M., Wang, Z., 2011. Specific and modular binding code for cytosine recognition in Pumilio/FBF (PUF) RNA-binding domains. *J Biol Chem* 286(30), 26732-26742.
- Dorsett, Y., Tuschl, T., 2004. siRNAs: applications in functional genomics and potential as therapeutics. *Nat Rev Drug Discov* 3(4), 318-329.
- East-Seletsky, A., O'Connell, M.R., Burstein, D., Knott, G.J., Doudna, J.A., 2017. RNA Targeting by Functionally Orthogonal Type VI-A CRISPR-Cas Enzymes. *Mol Cell* 66(3), 373-383 e373.
- East-Seletsky, A., O'Connell, M.R., Knight, S.C., Burstein, D., Cate, J.H., Tjian, R., Doudna, J.A., 2016. Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. *Nature* 538(7624), 270-273.
- Eid, A., Mahfouz, M.M., 2016. Genome editing: the road of CRISPR/Cas9 from bench to clinic. *Exp Mol Med* 48(10), e265.
- Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., Tuschl, T., 2001. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 411(6836), 494-498.
- Esteller, M., 2011. Non-coding RNAs in human disease. *Nat Rev Genet* 12(12), 861-874.
- Filipovska, A., Razif, M.F., Nygard, K.K., Rackham, O., 2011. A universal code for RNA recognition by PUF proteins. *Nat Chem Biol* 7(7), 425-427.
- Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., Mello, C.C., 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391(6669), 806-811.
- Galgano, A., Forrer, M., Jaskiewicz, L., Kanitz, A., Zavolan, M., Gerber, A.P., 2008. Comparative analysis of mRNA targets for human PUF-family proteins suggests extensive interaction with the miRNA regulatory system. *PLoS One* 3(9), e3164.
- Garneau, J.E., Dupuis, M.E., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadan, A.H., Moineau, S., 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468(7320), 67-71.
- Gaudet, J., VanderElst, I., Spence, A.M., 1996. Post-transcriptional regulation of sex determination in *Caenorhabditis elegans*: widespread expression of the sex-determining gene *fem-1* in both sexes. *Mol Biol Cell* 7(7), 1107-1121.
- Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., Ploegh, H.L., Bassik, M.C., Qi, L.S., Kampmann, M., Weissman, J.S., 2014. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* 159(3), 647-661.
- Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., Lim, W.A., Weissman, J.S., Qi, L.S., 2013. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 154(2), 442-451.
- Glisovic, T., Bachorik, J.L., Yong, J., Dreyfuss, G., 2008. RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett* 582(14), 1977-1986.
- Gootenberg, J.S., Abudayyeh, O.O., Lee, J.W., Essletzbichler, P., Dy, A.J., Joung, J., Verdine, V., Donghia, N., Daringer, N.M., Freije, C.A., Myhrvold, C., Bhattacharyya, R.P., Livny, J., Regev, A., Koonin, E.V., Hung, D.T., Sabeti, P.C., Collins, J.J., Zhang, F., 2017. Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science*.
- Gossen, M., Bujard, H., 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci U S A* 89(12), 5547-5551.
- Graveley, B.R., Maniatis, T., 1998. Arginine/serine-rich domains of SR proteins can function as activators of pre-mRNA splicing. *Mol Cell* 1(5), 765-771.
- Hale, C.R., Majumdar, S., Elmore, J., Pfister, N., Compton, M., Olson, S., Resch, A.M., Glover, C.V., 3rd, Graveley, B.R., Terns, R.M., Terns, M.P., 2012. Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNAs. *Mol Cell* 45(3), 292-302.
- Hale, C.R., Zhao, P., Olson, S., Duff, M.O., Graveley, B.R., Wells, L., Terns, R.M., Terns, M.P., 2009. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* 139(5), 945-956.
- Hannon, G.J., 2002. RNA interference. *Nature* 418(6894), 244-251.
- Hirano, H., Gootenberg, J.S., Horii, T., Abudayyeh, O.O., Kimura, M., Hsu, P.D., Nakane, T., Ishitani, R., Hatada, I., Zhang, F., Nishimasu, H., Nureki, O., 2016. Structure and Engineering of *Francisella novicida* Cas9. *Cell* 164(5), 950-961.
- Hsu, P.D., Lander, E.S., Zhang, F., 2014. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157(6), 1262-1278.

- Jackson, A.L., Bartz, S.R., Schelter, J., Kobayashi, S.V., Burchard, J., Mao, M., Li, B., Cavet, G., Linsley, P.S., 2003. Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol* 21(6), 635-637.
- Jackson, A.L., Burchard, J., Schelter, J., Chau, B.N., Cleary, M., Lim, L., Linsley, P.S., 2006. Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. *RNA* 12(7), 1179-1187.
- Ji, X., Zhang, H., Zhang, Y., Wang, Y., Gao, C., 2015. Establishing a CRISPR-Cas-like immune system conferring DNA virus resistance in plants. *Nat Plants* 1, 15144.
- Jiang, W., Samai, P., Marraffini, L.A., 2016. Degradation of Phage Transcripts by CRISPR-Associated RNases Enables Type III CRISPR-Cas Immunity. *Cell* 164(4), 710-721.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., Charpentier, E., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337(6096), 816-821.
- Jones, C.L., Napier, B.A., Sampson, T.R., Llewellyn, A.C., Schroeder, M.R., Weiss, D.S., 2012. Subversion of host recognition and defense systems by *Fransisella* spp. *Microbiol Mol Biol Rev* 76(2), 383-404.
- Kim, D.H., Rossi, J.J., 2007. Strategies for silencing human disease using RNA interference. *Nat Rev Genet* 8(3), 173-184.
- Kim, H., Kim, J.S., 2014. A guide to genome engineering with programmable nucleases. *Nat Rev Genet* 15(5), 321-334.
- Kim, K., Ryu, S.M., Kim, S.T., Baek, G., Kim, D., Lim, K., Chung, E., Kim, S., Kim, J.S., 2017. Highly efficient RNA-guided base editing in mouse embryos. *Nat Biotechnol* 35(5), 435-437.
- Kim, Y.B., Komor, A.C., Levy, J.M., Packer, M.S., Zhao, K.T., Liu, D.R., 2017. Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nat Biotechnol* 35(4), 371-376.
- Klug, A., 2010. The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annu Rev Biochem* 79, 213-231.
- Kole, R., Krainer, A.R., Altman, S., 2012. RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nat Rev Drug Discov* 11(2), 125-140.
- Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., Liu, D.R., 2016. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533(7603), 420-424.
- Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., Nureki, O., Zhang, F., 2015. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* 517(7536), 583-588.
- Lawhorn, I.E., Ferreira, J.P., Wang, C.L., 2014. Evaluation of sgRNA target sites for CRISPR-mediated repression of TP53. *PLoS One* 9(11), e113232.
- Lewin, A.S., Hauswirth, W.W., 2001. Ribozyme gene therapy: applications for molecular medicine. *Trends Mol Med* 7(5), 221-228.
- Li, L., Piatek, M.J., Atef, A., Piatek, A., Wibowo, A., Fang, X., Sabir, J.S., Zhu, J.K., Mahfouz, M.M., 2012. Rapid and highly efficient construction of TALE-based transcriptional regulators and nucleases for genome modification. *Plant Mol Biol* 78(4-5), 407-416.
- Liu, L., Li, X., Ma, J., Li, Z., You, L., Wang, J., Wang, M., Zhang, X., Wang, Y., 2017a. The Molecular Architecture for RNA-Guided RNA Cleavage by Cas13a. *Cell* 170(4), 714-726 e710.
- Liu, L., Li, X., Wang, J., Wang, M., Chen, P., Yin, M., Li, J., Sheng, G., Wang, Y., 2017b. Two Distant Catalytic Sites Are Responsible for C2c2 RNase Activities. *Cell* 168(1-2), 121-134 e112.
- Lunde, B.M., Moore, C., Varani, G., 2007. RNA-binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol* 8(6), 479-490.
- Maeder, M.L., Linder, S.J., Casado, V.M., Fu, Y., Ho, Q.H., Joung, J.K., 2013a. CRISPR RNA-guided activation of endogenous human genes. *Nat Methods* 10(10), 977-979.
- Maeder, M.L., Linder, S.J., Reyon, D., Angstman, J.F., Fu, Y., Sander, J.D., Joung, J.K., 2013b. Robust, synergistic regulation of human gene expression using TALE activators. *Nat Methods* 10(3), 243-245.
- Mahfouz, M.M., Piatek, A., Stewart, C.N., Jr., 2014. Genome engineering via TALENs and CRISPR/Cas9 systems: challenges and perspectives. *Plant Biotechnol J* 12(8), 1006-1014.
- Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F.J., Wolf, Y.I., Yakunin, A.F., van der Oost, J., Koonin, E.V., 2011. Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* 9(6), 467-477.
- Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., Barrangou, R., Brouns, S.J., Charpentier, E., Haft, D.H., Horvath, P., Moineau, S., Mojica, F.J., Terns, R.M., Terns, M.P., White, M.F., Yakunin, A.F., Garrett, R.A., van der Oost, J., Backofen, R., Koonin, E.V., 2015. An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol* 13(11), 722-736.
- Maris, C., Dominguez, C., Allain, F.H., 2005. The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J* 272(9), 2118-2131.
- Martin, G.B., Bogdanove, A.J., Sessa, G., 2003. Understanding the functions of plant disease resistance proteins. *Annu Rev Plant Biol* 54, 23-61.
- Martin, K.C., Ephrussi, A., 2009. mRNA localization: gene expression in the spatial dimension. *Cell* 136(4), 719-730.
- Meister, G., Tuschl, T., 2004. Mechanisms of gene silencing by double-stranded RNA. *Nature* 431(7006), 343-349.

- Mohr, S.E., Smith, J.A., Shamu, C.E., Neumuller, R.A., Perrimon, N., 2014. RNAi screening comes of age: improved techniques and complementary approaches. *Nat Rev Mol Cell Biol* 15(9), 591-600.
- Mojica, F.J., Diez-Villasenor, C., Garcia-Martinez, J., Almendros, C., 2009. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155(Pt 3), 733-740.
- Montiel-Gonzalez, M.F., Valledillo-Viejo, I., Yudowski, G.A., Rosenthal, J.J., 2013. Correction of mutations within the cystic fibrosis transmembrane conductance regulator by site-directed RNA editing. *Proc Natl Acad Sci U S A* 110(45), 18285-18290.
- Morris, A.R., Mukherjee, N., Keene, J.D., 2008. Ribonomic analysis of human Pum1 reveals cis-trans conservation across species despite evolution of diverse mRNA target sets. *Mol Cell Biol* 28(12), 4093-4103.
- Moscou, M.J., Bogdanove, A.J., 2009. A simple cipher governs DNA recognition by TAL effectors. *Science* 326(5959), 1501.
- Napoli, C., Lemieux, C., Jorgensen, R., 1990. Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell* 2(4), 279-289.
- Nelles, D.A., Fang, M.Y., O'Connell, M.R., Xu, J.L., Markmiller, S.J., Doudna, J.A., Yeo, G.W., 2016. Programmable RNA Tracking in Live Cells with CRISPR/Cas9. *Cell* 165(2), 488-496.
- Nishida, K., Arazoe, T., Yachie, N., Banno, S., Kakimoto, M., Tabata, M., Mochizuki, M., Miyabe, A., Arai, M., Hara, K.Y., Shimatani, Z., Kondo, A., 2016. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* 353(6305).
- O'Connell, M.R., Oakes, B.L., Sternberg, S.H., East-Seletsky, A., Kaplan, M., Doudna, J.A., 2014. Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature* 516(7530), 263-266.
- Orphanides, G., Reinberg, D., 2002. A unified theory of gene expression. *Cell* 108(4), 439-451.
- Ozawa, T., Natori, Y., Sato, M., Umezawa, Y., 2007. Imaging dynamics of endogenous mitochondrial RNA in single living cells. *Nat Methods* 4(5), 413-419.
- Pal-Bhadra, M., Bhadra, U., Birchler, J.A., 1997. Cosuppression in Drosophila: gene silencing of Alcohol dehydrogenase by white-Adh transgenes is Polycomb dependent. *Cell* 90(3), 479-490.
- Perez-Pinera, P., Kocak, D.D., Vockley, C.M., Adler, A.F., Kabadi, A.M., Polstein, L.R., Thakore, P.I., Glass, K.A., Ousterout, D.G., Leong, K.W., Guilak, F., Crawford, G.E., Reddy, T.E., Gersbach, C.A., 2013a. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat Methods* 10(10), 973-976.
- Perez-Pinera, P., Ousterout, D.G., Brunger, J.M., Farin, A.M., Glass, K.A., Guilak, F., Crawford, G.E., Hartemink, A.J., Gersbach, C.A., 2013b. Synergistic and tunable human gene activation by combinations of synthetic transcription factors. *Nat Methods* 10(3), 239-242.
- Piatek, A., Ali, Z., Baazim, H., Li, L., Abulfaraj, A., Al-Shareef, S., Aouida, M., Mahfouz, M.M., 2015. RNA-guided transcriptional regulation in planta via synthetic dCas9-based transcription factors. *Plant Biotechnol J* 13(4), 578-589.
- Price, A.A., Grakoui, A., Weiss, D.S., 2016. Harnessing the Prokaryotic Adaptive Immune System as a Eukaryotic Antiviral Defense. *Trends Microbiol* 24(4), 294-306.
- Price, A.A., Sampson, T.R., Ratner, H.K., Grakoui, A., Weiss, D.S., 2015. Cas9-mediated targeting of viral RNA in eukaryotic cells. *Proc Natl Acad Sci U S A* 112(19), 6164-6169.
- Puchta, H., 2016. Using CRISPR/Cas in three dimensions: towards synthetic plant genomes, transcriptomes and epigenomes. *Plant J* 87(1), 5-15.
- Pumplin, N., Voinnet, O., 2013. RNA silencing suppression by plant pathogens: defence, counter-defence and counter-counter-defence. *Nat Rev Microbiol* 11(11), 745-760.
- Pyott, D.E., Sheehan, E., Molnar, A., 2016. Engineering of CRISPR/Cas9-mediated potyvirus resistance in transgene-free Arabidopsis plants. *Mol Plant Pathol* 17(8), 1276-1288.
- Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., Lim, W.A., 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152(5), 1173-1183.
- Quenault, T., Lithgow, T., Traven, A., 2011. PUF proteins: repression, activation and mRNA localization. *Trends Cell Biol* 21(2), 104-112.
- Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., Zetsche, B., Shalem, O., Wu, X., Makarova, K.S., Koonin, E.V., Sharp, P.A., Zhang, F., 2015. In vivo genome editing using Staphylococcus aureus Cas9. *Nature* 520(7546), 186-191.
- Rautio, J., Kumpulainen, H., Heimbach, T., Oliyai, R., Oh, D., Jarvinen, T., Savolainen, J., 2008. Prodrugs: design and clinical applications. *Nat Rev Drug Discov* 7(3), 255-270.
- Romano, N., Macino, G., 1992. Quelling: transient inactivation of gene expression in Neurospora crassa by transformation with homologous sequences. *Mol Microbiol* 6(22), 3343-3353.
- Ruiz, F., Vayssie, L., Klotz, C., Sperling, L., Madeddu, L., 1998. Homology-dependent gene silencing in Paramecium. *Mol Biol Cell* 9(4), 931-943.
- Rupaimoole, R., Slack, F.J., 2017. MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. *Nat Rev Drug Discov* 16(3), 203-222.
- Samai, P., Pyenson, N., Jiang, W., Goldberg, G.W., Hatoum-Aslan, A., Marraffini, L.A., 2015. Co-transcriptional DNA and RNA Cleavage during Type III CRISPR-Cas Immunity. *Cell* 161(5), 1164-1174.

- 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., Merlin, D., Zhou, P., Endtz, H.P., Weiss, D.S., 2014. A CRISPR-Cas system enhances envelope integrity mediating antibiotic resistance and inflammasome evasion. *Proc Natl Acad Sci U S A* 111(30), 11163-11168.
- Sampson, T.R., Saroj, S.D., Llewellyn, A.C., Tzeng, Y.L., Weiss, D.S., 2013. A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. *Nature* 497(7448), 254-257.
- Scherer, L.J., Rossi, J.J., 2003. Approaches for the sequence-specific knockdown of mRNA. *Nat Biotechnol* 21(12), 1457-1465.
- Scotti, M.M., Swanson, M.S., 2016. RNA mis-splicing in disease. *Nat Rev Genet* 17(1), 19-32.
- Shimatani, Z., Kashojiya, S., Takayama, M., Terada, R., Arazoe, T., Ishii, H., Teramura, H., Yamamoto, T., Komatsu, H., Miura, K., Ezura, H., Nishida, K., Ariizumi, T., Kondo, A., 2017. Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. *Nat Biotechnol* 35(5), 441-443.
- Shmakov, S., Abudayyeh, O.O., Makarova, K.S., Wolf, Y.I., Gootenberg, J.S., Semenova, E., Minakhin, L., Joung, J., Konermann, S., Severinov, K., Zhang, F., Koonin, E.V., 2015. Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems. *Mol Cell* 60(3), 385-397.
- Shmakov, S., Smargon, A., Scott, D., Cox, D., Pyzocha, N., Yan, W., Abudayyeh, O.O., Gootenberg, J.S., Makarova, K.S., Wolf, Y.I., Severinov, K., Zhang, F., Koonin, E.V., 2017. Diversity and evolution of class 2 CRISPR-Cas systems. *Nat Rev Microbiol* 15(3), 169-182.
- Sigoillot, F.D., Lyman, S., Huckins, J.F., Adamson, B., Chung, E., Quattrochi, B., King, R.W., 2012. A bioinformatics method identifies prominent off-targeted transcripts in RNAi screens. *Nat Methods* 9(4), 363-366.
- Sinkunas, T., Gasiunas, G., Waghmare, S.P., Dickman, M.J., Barrangou, R., Horvath, P., Siksnys, V., 2013. In vitro reconstitution of Cas9-mediated CRISPR immunity in *Streptococcus thermophilus*. *EMBO J* 32(3), 385-394.
- Smargon, A.A., Cox, D.B., Pyzocha, N.K., Zheng, K., Slaymaker, I.M., Gootenberg, J.S., Abudayyeh, O.A., Essletzbichler, P., Shmakov, S., Makarova, K.S., Koonin, E.V., Zhang, F., 2017. Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNase Differentially Regulated by Accessory Proteins Csx27 and Csx28. *Mol Cell* 65(4), 618-630 e617.
- Soosaar, J.L., Burch-Smith, T.M., Dinesh-Kumar, S.P., 2005. Mechanisms of plant resistance to viruses. *Nat Rev Microbiol* 3(10), 789-798.
- Sorek, R., Lawrence, C.M., Wiedenheft, B., 2013. CRISPR-mediated adaptive immune systems in bacteria and archaea. *Annu Rev Biochem* 82, 237-266.
- St Johnston, D., 2005. Moving messages: the intracellular localization of mRNAs. *Nat Rev Mol Cell Biol* 6(5), 363-375.
- Staals, R.H., Agari, Y., Maki-Yonekura, S., Zhu, Y., Taylor, D.W., van Duijn, E., Barendregt, A., Vlot, M., Koehorst, J.J., Sakamoto, K., Masuda, A., Dohmae, N., Schaap, P.J., Doudna, J.A., Heck, A.J., Yonekura, K., van der Oost, J., Shinkai, A., 2013. Structure and activity of the RNA-targeting Type III-B CRISPR-Cas complex of *Thermus thermophilus*. *Mol Cell* 52(1), 135-145.
- Staals, R.H., Zhu, Y., Taylor, D.W., Kornfeld, J.E., Sharma, K., Barendregt, A., Koehorst, J.J., Vlot, M., Neupane, N., Varosieau, K., Sakamoto, K., Suzuki, T., Dohmae, N., Yokoyama, S., Schaap, P.J., Urlaub, H., Heck, A.J., Nogales, E., Doudna, J.A., Shinkai, A., van der Oost, J., 2014. RNA targeting by the type III-A CRISPR-Cas Csm complex of *Thermus thermophilus*. *Mol Cell* 56(4), 518-530.
- Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C., Doudna, J.A., 2014. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* 507(7490), 62-67.
- Syed, N.H., Kalyna, M., Marquez, Y., Barta, A., Brown, J.W., 2012. Alternative splicing in plants—coming of age. *Trends Plant Sci* 17(10), 616-623.
- Tamulaitis, G., Kazlauskienė, M., Manakova, E., Venclovas, C., Nwokeoji, A.O., Dickman, M.J., Horvath, P., Siksnys, V., 2014. Programmable RNA shredding by the type III-A CRISPR-Cas system of *Streptococcus thermophilus*. *Mol Cell* 56(4), 506-517.
- Tilsner, J., Linnik, O., Christensen, N.M., Bell, K., Roberts, I.M., Lacomme, C., Oparka, K.J., 2009. Live-cell imaging of viral RNA genomes using a Pumilio-based reporter. *Plant J* 57(4), 758-770.
- van der Krol, A.R., Mur, L.A., Beld, M., Mol, J.N., Stuitje, A.R., 1990. Flavonoid genes in petunia: a addition of a limited number of gene copies may lead to a suppression of gene expression. *Plant Cell* 2(4), 291-299.
- van der Oost, J., Westra, E.R., Jackson, R.N., Wiedenheft, B., 2014. Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat Rev Microbiol* 12(7), 479-492.
- Voinnet, O., 2005. Induction and suppression of RNA silencing: insights from viral infections. *Nat Rev Genet* 6(3), 206-220.
- Wang, X., McLachlan, J., Zamborelli, P.D., Hall, T.M., 2002. Modular recognition of RNA by a human pumilio-homology domain. *Cell* 110(4), 501-512.
- Wang, Y., Cheong, C.G., Hall, T.M., Wang, Z., 2009. Engineering splicing factors with designed specificities. *Nat Methods* 6(11), 825-830.
- Wei, H., Wang, Z., 2015. Engineering RNA-binding proteins with diverse activities. *Wiley Interdiscip Rev RNA* 6(6), 597-613.
- Wiedenheft, B., Sternberg, S.H., Doudna, J.A., 2012. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482(7385), 331-338.
- Wolfe, S.A., Nekudova, L., Pabo, C.O., 2000. DNA recognition by Cys2His2 zinc finger proteins. *Annu Rev Biophys Biomol Struct* 29, 183-212.
- Wright, A.V., Nunez, J.K., Doudna, J.A., 2016. Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering. *Cell* 164(1-2), 29-44.

- Yamada, T., Yoshimura, H., Inaguma, A., Oza wa, T., 2011. Visualization of nonengineered single mRNAs in living cells using genetically encoded fluorescent probes. *Anal Chem* 83(14), 5708-5714.
- Zaidi, S.S., Tashkandi, M., Mansoor, S., Mahfouz, M.M., 2016. Engineering Plant Immunity: Using CRISPR/Cas9 to Generate Virus Resistance. *Front Plant Sci* 7, 1673.
- Zalatan, J.G., Lee, M.E., Almeida, R., Gilbert, L.A., Whitehead, E.H., La Russa, M., Tsai, J.C., Weissman, J.S., Dueber, J.E., Qi, L.S., Lim, W.A., 2015. Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell* 160(1-2), 339-350.
- Zamecnik, P.C., Stephenson, M.L., 1978. Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. *Proc Natl Acad Sci U S A* 75(1), 280-284.
- Zhang, F., Cong, L., Lodato, S., Kosuri, S., Church, G.M., Arlotta, P., 2011. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol* 29(2), 149-153.
- Zhang, W., Wang, Y., Dong, S., Choudhury, R., Jin, Y., Wang, Z., 2014. Treatment of type 1 myotonic dystrophy by engineering site-specific RNA endonucleases that target (CUG)(n) repeats. *Mol Ther* 22(2), 312-320.
- Zhou, J., Rossi, J., 2017. Aptamers as targeted therapeutics: current potential and challenges. *Nat Rev Drug Discov* 16(3), 181-202.
- Zong, Y., Wang, Y., Li, C., Zhang, R., Chen, K., Ran, Y., Qiu, J.L., Wang, D., Gao, C., 2017. Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. *Nat Biotechnol* 35(5), 438-440.

Table 1. Comparison of CRISPR/Cas13 to RNAi and engineered RNA-binding proteins.

	RNAi	RNA-binding proteins	CRISPR/Cas13
Components	Requires host endogenous RNA-induced silencing complex (RISC), and delivery of dsRNA	Requires delivery of engineered RBPs	Requires delivery of Cas13 protein and gRNAs
Specificity	Dictated by ~21-nt sequence of the siRNA	Dictated by engineered protein domains	Dictated by 28–36 nt within the gRNA
Cleavage activity	Requires RISC's nuclease activity	Requires fusion of a nuclease domain to the RBD	Possesses an innate ribonuclease activity
Localization	Cytoplasmic system	Cytoplasmic and nuclear system	Cytoplasmic and nuclear system
Off-target effects	High	Low	Low
Cost of experiment	Low	High	Low
Multiplex targeting	Feasible with multiple siRNAs	Requires the design and synthesis of large protein libraries	Feasible with multiple gRNAs
Modularity	RNA binding and degradation only	Diverse potential functions depending on the fused functional domain	Diverse potential functions depending on the fused functional domain

Figure Legends:

Figure 1: CRISPR/SpCas9 and RCas9 systems

A: Schematic representation of the domain structure and organization of the SpCas9 protein. The two nuclease domains HNH and RuvC-like domains are shown with the positions of the domain mutations for generation of dead Cas9 (dCas9), D10A, Asp¹⁰→Ala¹⁰; H840A; His⁸⁴⁰→Ala⁸⁴⁰.

B: Expression of Cas9 and its cognate gRNA. Engineering an expression system for eukaryotic expression of a species-specific codon-optimized Cas9 protein and its guide RNA. Expression of Cas9 protein is driven by a constitutive promoter and fused to nuclear localization signals (NLS) that enable its import into the nucleus. Whereas the guide RNA is engineered to include the 20-nt complementary region (spacer) for specific DNA binding (yellow) and the gRNA hairpin or scaffold (blue) and is driven by a Pol III promoter such as U3 or U6 promoters.

C: CRISPR/Cas9 effector on the target DNA. Cas9 endonuclease forms the effector complex with the gRNA, which then screens DNA for sequence complementarity to the 20 nt spacer sequence in the gRNA. Cas9 binding and activity require the presence of a protospacer adjacent motif (PAM) of 5'-NGG-3', which N can be any nucleotide, downstream of the target sequence on the DNA. Upon recognition and binding of Cas9 to the target DNA sequence, Cas9 unwinds the target DNA duplex and cleaves the target sequence 3 nt upstream of the PAM.

D: RNA-targeting Cas9 (RCas9) system as a divergent of CRISPR/Cas9. CRISPR/Cas9 can be repurposed to recognize and target single stranded RNA (ssRNA) while avoiding the recognition of genomic DNA region complementary to the spacer sequence of the gRNA. The strategy of targeting RNA with Cas9 relies on the dependency of Cas9 on PAM sequence for target binding. When Cas9-gRNA complex is programmed to target a specific ssRNA sequence complementary to the gRNA spacer sequence, and by providing a mismatched PAM in a separate antisense DNA oligonucleotide (PAMmer) that can hybridize to the target RNA and positions the PAM adjacent to the target sequence, Cas9 is able to recognize and cleave target RNAs, while avoiding PAM-lacking genomic DNA sequences complementary to the gRNA.

Figure 2: The mechanism of RNA interference

The mechanism of RNA interference and strategies for generating siRNAs. The RNA silencing mechanism relies on small interfering RNAs (siRNAs) to guide cellular machinery to

homologous sequences, leading to their degradation and silencing. siRNAs originate from long dsRNA precursors that can be derived from endogenous or exogenous sources. Endogenous micro-RNAs (miRNAs) are initially transcribed as primary miRNAs (not shown) that are processed in the nucleus into pre-miRNAs, and subsequently exported to the cytoplasm, where they are processed by the Dicer enzyme (yellow) into 21–23 nt double-stranded siRNAs. Alternatively, siRNAs can be supplied exogenously as short (~21-mer) dsRNAs or long dsRNAs via experimental manipulation or viral RNA. Similar to the pre-miRNAs, long dsRNAs are processed in the cytoplasm by Dicer, generating siRNAs. siRNAs are incorporated into the multicomponent RNA-induced silencing complex (RISC) (blue), where the siRNA duplex is unwound and the guide (anti-sense) RNA strand assembles into the RISC complex. The guide siRNA directs RISC to bind to and degrade complementary mRNA targets through the activity of Argonaute 2 (AGO2) proteins (pink), leading to specific silencing of the target gene.

Figure 3: CRISPR/Cas13a System:

A: Schematic representation of the domain structure and organization of the LshCas13a protein. The two HEPN nuclease domains are shown with the positions of the domain mutations for generation of dead Cas13a (dCas13a) indicated with the red stars. In addition, the position of the catalytic site for pre-crRNA processing is highlighted with the blue star.

B: Expression of Cas13a and its cognate gRNA. Engineering an expression system for in eukaryotic expression of a species-specific codon optimized Cas13a protein and its guide RNA. Expression of Cas13a protein is driven by a constitutive promoter and can be either fused to nuclear localization signals (NLS) that enable its import into the nucleus for in nucleus activity, or without NLS to be utilized for in cytoplasm activity. Whereas the guide RNA is engineered to comprise (in case of LshCas13a homologue) 28 nt complementary region (spacer) for specific RNA binding (brown) and the gRNA hairpin or scaffold (blue) and is driven by a Pol III promoter such as U3 or U6 promoters to be used for in the nucleus. However, in case CRISPR/Cas13a is intended to be used in the cytoplasm, a Pol II promoter can be used to drive the expression of the pre-crRNA transcript for subsequent processing in the cytoplasm by Cas13a to generate mature crRNAs.

C: CRISPR/Cas13a effector on the target RNA. Cas13a ribonuclease forms the effector complex with the gRNA, which then seeks for RNA sequence complementarity to the 28 nt

spacer sequence in the gRNA. Cas13a activity requires the presence of a protospacer flanking site (PFS) of H, which H can be any nucleotide except G, downstream of the target sequence on the RNA. Upon recognition and binding of Cas13a to the target RNA sequence, Cas13a cleaves the target RNA non-specifically, leading to target RNA degradation.

Figure 4: Natural activity of the CRISPR/Cas13a system

The CRISPR/Cas13a system evolved to provide adaptive immunity in bacteria and Archaea against invading RNA elements, such as RNA viruses, by mediating the detection of invading RNAs and their subsequent degradation. The mechanism of interference involves three main stages: 1) In the adaptation or spacer acquisition phase, protospacers are acquired from the genome of the invading viruses and integrated as new spacers into the CRISPR array by an unknown mechanism. 2) In the expression and maturation stage, the CRISPR array is transcribed into precursor CRISPR RNA transcripts (pre-crRNA) composed of spacers interspersed with repeats. The pre-crRNA transcripts are processed to generate mature cr-RNAs. Unlike CRISPR/Cas9 and other systems that require accessory factors for processing the pre-crRNA, Cas13a can process its own pre-crRNA transcripts, generating mature crRNAs which contain a transcribed spacer and part of the repeat sequence that associates with Cas13a to form an inactive Cas13a-crRNA complex. 3) In the interference stage, the mature crRNAs guide the Cas13a ribonuclease to bind the complementary sequence of the target RNA. Binding of Cas13a to the target sequence leads to the activation of its innate, non-specific RNase activity, resulting in promiscuous cleavage of the target and non-target RNAs in the cell. This promiscuous cleavage of host endogenous RNAs may suggest a natural mechanism to sense invasive viral RNAs, and induce programmed cell death (PCD) or host cell dormancy, preventing the spread of the viral infection.

Figure 5: Potential Applications of CRISPR/Cas13a system.

The ability of CRISPR/Cas13a to recognize untagged, endogenous RNA substrates through simple RNA:RNA base pairing complementarity represents a major advance in the field of RNA targeting and manipulation. In eukaryotes, RNAs undergo various processing steps, including alternative splicing in the nucleus, nuclear export, base or backbone modification, and translation. The simplicity and multiplexed targeting afforded by CRISPR/Cas13a could support

the development of programmable tools for RNA recognition, facilitating targeted manipulation of these processes in living cells with broad applications in basic and applied biology, biotechnology, and genomic medicine.

A: Targeted RNA cleavage and gene knockdown. Harnessing the innate ribonuclease activity of Cas13a with its high level of specificity and affinity to RNAs may allow more efficient gene knockdown and enable knockdown in compartments or organisms in which RNAi-mediated gene silencing is not possible.

B: Imaging and tracking specific endogenous transcripts. dCas13a tethered to a fluorescent protein such as GFP could allow for the imaging of unlabeled RNAs, facilitating the tracking and localization of endogenous transcripts inside cells.

C: Enhancing RNA stabilization and gene expression. dCas13a fused to an RNA stabilizing factor could prevent degradation of target RNAs. In addition, tethering an initiation factor such as GLD2 or eIF4E to dCas13a can be used to drive translation of specific endogenous mRNAs, resulting in the enhancement of gene expression of target genes without affecting the abundance of endogenous transcripts.

D: Specific isolation and identification of RNA and its associated RNA binding proteins (RBPs). dCas13a fused to beads could allow for the specific isolation of endogenous transcripts or RNA–protein complexes for analysis, including RNA sequencing, identification of RNA-binding proteins, and the identification of protein–protein interactions between RNA-binding proteins.

E: Manipulation of alternative splicing of specific pre-mRNA. Pre-mRNA splicing could be altered by fusing dCas13 to a splicing factor. dCas13a can be programmed to recruit the splicing factor to a specific splicing regulatory signal, promoting the desired alternative splicing outcome.

F: RNA base editing. Fusing an RNA editing enzyme, such as RNA deaminase, to dCas13a could enable site-specific A-to-I or C-to-U editing of RNA. The ability of such RNA base editors to modify the composition of endogenous transcripts, particularly coding mRNAs, could be used to correct harmful mutations without affecting the genomic DNA sequence.

G: Induction of programmed cell death (PCD). The collateral cleavage activity of Cas13a upon binding to target RNA could be harnessed to induce PCD in response to a specific transcript. For example, when Cas13a is targeted to a unique or aberrant transcript expressed only in tumor cells, the collateral cleavage activity of Cas13a activated due to binding to the

target transcript will result in the collateral cleavage of endogenous transcripts, leading to the induction of PCD and selective killing of cancerous cells.

H: CRISPR/Cas13-based diagnostic platform. The high specificity and target RNA binding-based collateral cleavage activity of Cas13 can be harnessed to detect the presence of specific nucleic acids in a pool of non-target transcripts. By using reporter RNAs that release a fluorescent signal upon cleavage, the binding of active Cas13 to the target RNA will activate the collateral cleavage activity of Cas13, thereby cleaving the reporter RNAs and releasing of the signal, indicating the presence of the target RNA. Such a highly sensitive detection tool would have widespread use in diagnostics and research.

Figure 6: CRISPR/Cas13a-mediated RNA and DNA virus interference.

CRISPR/Cas13a could be utilized to confer RNA as well as DNA virus resistance in eukaryotic cells. By engineering host genome to express CRISPR/Cas13a machinery components (Cas13a and gRNA), NLS-Cas13a-gRNA complex can be reprogrammed to target and interfere with viral RNAs either transcribed inside the nucleus from DNA or reverse transcribed RNA viruses, or imported directly into the nucleus such as in Orthomyxovirus. In addition, since most RNA viruses replicate in the cytoplasm with no DNA intermediate, the ability of CRISPR/Cas13a to process pre-crRNA to generate mature crRNAs facilitates its applicability for targeting RNA viral genomes in the cytoplasm, offering the potential to be developed to provide specific and efficacious acquired resistance against a wide spectrum of pathogenic viruses.

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6