Genome Editing and CRISPR Technology

Colin KW Lim and Thomas Gaj, Department of Bioengineering, University of Illinois, Urbana, IL, United States

© 2023 Elsevier Inc. All rights reserved.

Introduction 650
CRISPR Basics: From Adaptive Immunity to Genome Editing 650
DNA Targeting by Cas9 651
Non-Homologous End-Joining: A Genome Editing Repair Pathway 652
Homology-Directed Repair: High-fidelity Genome Editing 652
Emerging Technologies for Genome Editing: Base Editing and Prime Editing 653
Alternate CRISPR Systems and Their Applications for Genome Editing 654
Using CRISPR Technology to Regulate Gene Expression 654
Conclusion 655
References 655

Abstract

Genome editing has revolutionized the life sciences by providing scientists with an efficient means for introducing targeted modifications into DNA. Driving these advances are programmable nucleases, which stimulate gene editing by inducing targeted DNA double-strand breaks (DSBs) that in turn activate cellular DNA repair pathways that can mediate the introduction of specific edits. For most genome engineers, technologies derived from naturally occurring CRISPR systems have become the platform of choice as CRISPR is highly versatile and efficient. In this article, we introduce CRISPR technology, describe how it is used to edit DNA and discuss its most common applications.

Key Points

- CRISPR systems are versatile and efficient technologies for genome editing and transcriptome engineering.
- The most commonly used CRISPR system is CRISPR-Cas9, which consists of the Cas9 endonuclease and a guide RNA molecule that directs Cas9 to a target DNA sequence.
- DNA breaks induced by Cas9 activate cellular DNA repair pathways that mediate the introduction of a DNA edit.
- CRISPR systems are naturally diverse and the discovery of unique CRISPR effector proteins is expanding the capabilities of the genome editing and transcriptome engineering toolbox.

Introduction

Genome editing technologies have revolutionized many areas of the life sciences by providing scientists with an efficient means for introducing targeted modifications – or edits – to DNA. The emergence of these tools has made tasks such as creating designer cell lines, generating transgenic animal models, or performing large-scale genetic screens more routine than ever, which in turn, has accelerated the pace of discoveries in the biological sciences. The technology behind these advances are targetable nucleases, which stimulate the DNA editing process by inducing a DNA double-strand break (DSB) at a targeted genomic sequence. This DSB, in turn, activates cellular DNA repair pathways that are exploited to introduce a desired edit at the target sequence. Importantly, while several modalities capable of inducing targeted DSBs have been developed over the past two decades, technologies derived from naturally occurring CRISPR (short for Clustered Regularly Interspaced Short Palindromic Repeat) systems have become the platform of choice for most biologists and bioengineers. CRISPR systems are simple to implement, versatile in their capabilities and efficient, making them advantageous for routine and complex genome-editing procedures alike. In this article, we introduce CRISPR technology. We describe how it is used to edit DNA, its most common applications, its shortcomings and several of the latest versions of the technology that has expanded its capabilities.

CRISPR Basics: From Adaptive Immunity to Genome Editing

CRISPR-based technologies are now widely used for DNA editing in eukaryotic and prokaryotic cells. However, CRISPR systems evolved to provide adaptive immunity to bacteria and archaea against invading mobile genetic elements. In nature, a CRISPR

 doi:10.1016/B978-0-12-821618-7.00104-8

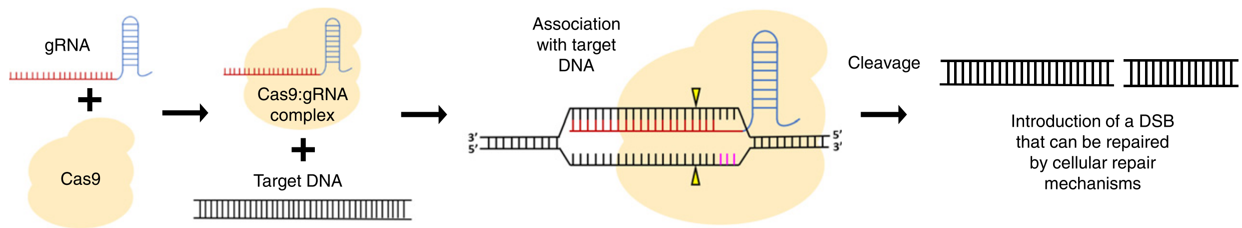

Fig. 1 *CRISPR-Cas9 targeting*. The CRISPR-Cas9 system consists of two components: the Cas9 endonuclease and a guide RNA (gRNA) molecule. When co-expressed in cells these components form a Cas9:gRNA complex. Cas9 scans the genome, searching for compatible PAM sequences (magenta). Hybridization occurs if the spacer in the gRNA (red) is complementary to the target DNA sequence. This results in the cleavage of DNA by Cas9 (yellow arrowheads) and the formation of a double-strand break (DSB).

genomic loci consists of an array of spacer sequences with intervening repeats that are flanked by genes encoding for CRISPR-associated (Cas) proteins, with the number and function of these Cas proteins depending on a particular CRISPR system (Wiedenheft *et al.*, 2012). Within a CRISPR array, each spacer encodes a short sequence that is identical to a segment of a foreign genetic element that previously invaded the cell and was subsequently captured and inserted into the CRISPR locus. Thus, acquisition of spacer sequences to a CRISPR locus is equivalent to immunization in the human immune system.

Following the invasion of a mobile element, the CRISPR array is transcribed into a precursor RNA that is processed to individual CRISPR RNA (crRNA) molecules that each contain the aforementioned spacer sequences (Wiedenheft *et al.*, 2012). At this point, the mechanism for interference against the invading element can vary depending on the type and class of the CRISPR system. In the case of the prototypical CRISPR-Cas locus from *Streptococcus pyogenes*, the crRNA anneals to a trans-activating RNA (tracrRNA) to form a tracrRNA:crRNA complex, which then directs an endonuclease named Cas9 to a DNA sequence complementary to the spacer encoded by the crRNA via RNA-DNA base-pairing (Jinek *et al.*, 2012). Once bound to the target, Cas9 cleaves both strands of the DNA, resulting in the degradation of the foreign genetic element and elimination of the invading threat. Notably, the only requirement for DNA targeting by Cas9 is the presence of a short nucleotide motif, named the protospacer adjacent motif (PAM), immediately downstream of the crRNA sequence and located in the DNA. The PAM is specifically bound by Cas9, which triggers its unwinding of DNA, enabling the base-pairing of the crRNA with the target DNA sequence (of note, while the PAM is a conserved element for targeting by Cas9, the exact sequence required for the PAM can vary depending on the Cas9 ortholog) (Sternberg *et al.*, 2014).

Thus, Cas9 is a programmable endonuclease that cleaves target DNA sequences that are complementary to the spacer sequence in its crRNA. For their central roles in elucidating the mechanism of DNA targeting by Cas9 and for demonstrating its potential for DNA editing, Dr. Emmanuelle Charpentier and Dr. Jennifer Doudna were jointly awarded the 2020 Nobel Prize in Chemistry.

DNA Targeting by Cas9

Though naturally occurring CRISPR systems contain crRNA arrays and Cas proteins that contribute to spacer acquisition and targeting, the CRISPR-Cas9 system most commonly used for genome editing consists of only two components: the Cas9 endonuclease, which induces a DSB at a targeted genomic site, and an engineered guide RNA (gRNA) molecule that directs Cas9 to the genomic site (**Fig. 1**) (Jinek *et al.*, 2012). The gRNA most commonly used for genome editing is a chimeric molecule that combines crRNA and tracrRNA elements and thus also contains the programmable spacer sequence that mediates the targeting of Cas9. For the prototypical Cas9 variant from *S. pyogenes*, referred to as SpCas9, the spacer is typically 20 nucleotides in length (Jinek *et al.*, 2012).

To initiate DNA editing, Cas9 and the gRNA must first be introduced to cells. For cells in culture, this is usually achieved by transient transfection of a plasmid DNA encoding the Cas9 nuclease and the gRNA, though Cas9 and the gRNA can also be delivered to cells as a pre-formed ribonucleoprotein (RNP) complex. Once inside the cell and complexed with gRNA, Cas9 enters the nucleus, where it searches for compatible PAMs through random collisions with genomic DNA (Sternberg *et al.*, 2014). The SpCas9 protein requires a 5′-NGG-3′ PAM, where "N" can be any nucleotide (Jinek *et al.*, 2012).

Once bound to a compatible PAM, Cas9 unwinds the DNA, enabling the gRNA to hybridize to the exposed target strand. Hybridization of the gRNA to a perfectly complementary target sequence then activates the cleavage activity of the Cas9 protein (Sternberg *et al.*, 2014). Using its RuvC and HNH nuclease domains, Cas9 cleaves both strands of the target DNA but then remains tightly bound to DNA until it is displaced by cellular enzymes (**Fig. 1**). Cas9 specificity is thus enforced by RNA-DNA base-pairing and PAM binding. To this end, the ability for Cas9 to tolerate mismatches between the gRNA and target DNA has been intensely investigated, as the cleavage of non-targeted DNA sequences can result in potentially deleterious off-target mutations. Though largely sensitive to base mismatches, studies have nonetheless demonstrated that Cas9 can, in some cases, tolerate one- or two-base mismatches, particularly if they are located at the PAM-distal end of the targeted sequence (Anderson *et al.*, 2015; Fu *et al.*, 2016). Findings such as these have spurred the development of high-fidelity Cas9 variants with a reduced ability to tolerate mismatches (Kleinstiver *et al.*, 2016).

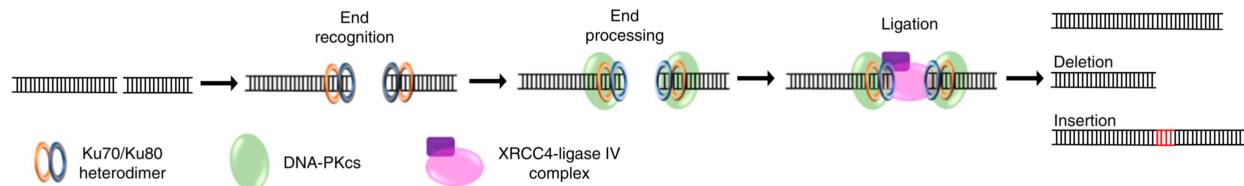

Fig. 2 *Brief description of NHEJ.* NHEJ comprises three steps: end recognition, end processing and ligation. During end recognition, the Ku70 and Ku80 heterodimer binds to free DNA ends and recruits additional accessory factors, including the DNA-dependent protein kinase DNA-PKcs. Recruited proteins contribute to the processing of the DNA ends to enable re-ligation by the XRCC4-ligase IV complex. This pathway is not high-fidelity. Thus, random base insertion and deletion (indels) can be introduced during DNA repair.

Non-Homologous End-Joining: A Genome Editing Repair Pathway

Though induction of a DSB by Cas9 is critical to DNA editing, it is only the first step in the process. Cleavage of the target sequence by Cas9 next activates cellular DNA repair mechanisms that influence the type of edit that will be introduced to the DNA. One of these pathways is non-homologous end joining (NHEJ), an error-prone mechanism that can create base insertions or deletions (indels) at the target site.

In particular, following the induction of a DSB by Cas9, a complex consisting of the proteins Ku70 and Ku80 bind to the ends of the DSB and recruit accessory factors that work together to process the DNA ends for re-ligation (**Fig. 2**) (Chang *et al.*, 2017). However, this pathway is not high-fidelity. During DNA repair, random indel mutations, which typically range in size from one to ten nucleotides, can be introduced at the target sequence (Chang *et al.*, 2017). Importantly, given the random nature of NHEJ, the majority of these indel mutations are expected to result in frameshift mutations, which can lead to a downstream nonsense mutation that can terminate the expression of a target gene. Thus, Cas9 can be utilized to induce a gene knockout, an outcome that can be exploited to generate knockout cell lines. This outcome also has therapeutic applications, as it offers means for disrupting the expression of a mutant gene.

In addition to inducing a gene knockout, NHEJ can be harnessed to disrupt functional elements within a gene sequence. One such example is exon skipping. By using Cas9 to direct indel mutations at splice acceptor or donor sites in introns (thereby disrupting them), a targeted exon can be excluded from its mature mRNA, resulting in the generation of a unique protein isoform (Chen *et al.*, 2018a). Beyond exon skipping, Cas9 can be used to delete specific sequences of DNA from cells, including whole exons and genes. This can be achieved by utilizing two gRNAs to simultaneously introduce two DSBs, which, following NHEJ-mediated re-joining of the distal DNA ends, can result in the deletion of the intervening DNA sequence (Neldeborg *et al.*, 2019).

Finally, in addition to disrupting or deleting genomic sequences, NHEJ can be harnessed to insert or knock-in new sequences at the DSB via a method known as homology-independent targeted integration (Zeng *et al.*, 2020). This is achieved by co-delivering Cas9 and gRNA with either: (1) a linear donor DNA molecule encoding a tag or gene of interest with free DNA ends or (2) a plasmid DNA encoding the tag or gene of interest adjacent to gRNA target site(s), which are converted to free DNA ends following Cas9-mediated cleavage. Following the induction of the targeted DSB by Cas9, free end-containing donor molecules can be ligated to the target sequence, resulting in the knock-in of the tag- or gene-containing sequence. This approach can be used to insert epitope tags or reporter genes to endogenous protein-coding gene sequences that could, for example, enable monitoring of protein trafficking or localization.

Homology-Directed Repair: High-fidelity Genome Editing

Despite its efficiency and wide availability in a range of cell types, NHEJ possesses several limitations that can impact its utility. For instance, NHEJ is a stochastic pathway that creates random indel mutations. Thus, though useful for inducing a gene knockout or creating a deletion, the genome-editing products generated by NHEJ are unpredictable and potentially mutagenic. Additionally, NHEJ does not have the ability to faithfully introduce single- or multi-base edits at a target sequence, limiting its potential for applications that require high-fidelity outcomes, such as the correction of a disease-causing mutation. However, there exists a DNA repair pathway that can be utilized by Cas9 and other genome-editing nucleases to enable a more precise form of genome-editing: homology-directed repair (HDR).

HDR is a homology-dependent repair pathway that can be harnessed by delivering Cas9 and gRNA with a DNA donor template bearing the target edit with flanking homology arms to the target sequence. Following the induction of a DSB by Cas9, the 5′ ends at the DSB are resected by repair enzymes to create 3′ overhangs. These ends then invade the homologous sequence in the donor template to serve as primers to repair the damaged DNA (**Fig. 3**) (Ran *et al.*, 2013b). Thus, by HDR, DNA sequences encoded between the homology arms on a donor template can be faithfully copied into the target sequence. In practice, HDR is typically accomplished with one of several types of donor templates, including a plasmid, a linear viral vector or a single-stranded oligonucleotide (ssODN). For inserting DNA sequences typically less than the 50 base-pairs, ssODNs have emerged as a particularly efficient method, requiring only ~50–80 base-pairs of homology in each arm (Ran *et al.*, 2013b). However,

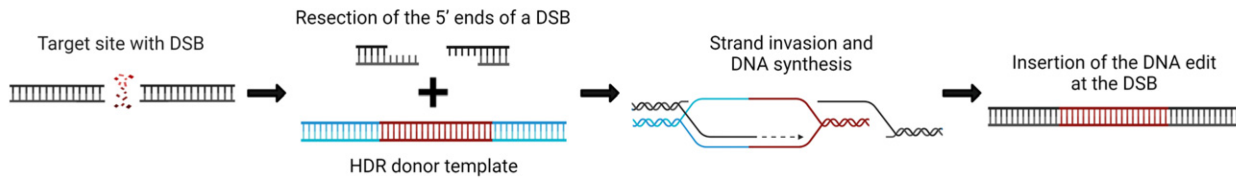

Fig. 3 *Brief description of HDR*. Following the induction of a DSB by Cas9, the 5' ends of the DSB are resected by endogenous exonucleases, enabling 3' overhangs to invade a homologous donor template (blue) for DNA synthesis. This enables the sequence between the homology arms (red) to be inserted to the target sequence.

ssODNs are not as effective for inserting larger sequences or whole transgenes. For larger constructs, plasmid DNA or viral vectors are instead commonly utilized. Importantly, numerous factors can influence the efficiency of insertion, including the length of the homology arms and the position of the Cas9 cleavage site relative to the homology arms. For maximally efficient HDR, homology arms are recommended to be less than 10 base-pairs away from the cleavage site and no farther than 100 base-pairs (Ran *et al.*, 2013b).

While HDR is a useful pathway for applications requiring the seamless introduction of single- or multi-base substitutions, it can also be used to knock-in single or multiple transgenes into a genomic locus more faithfully than with the homology-independent approaches described above. However, despite its fidelity, HDR-based methods do possess limitations. For one, HDR can be inefficient, in some cases requiring the use of selectable markers to isolate gene-modified cells (Ran *et al.*, 2013b). To this end, strategies have been developed for enhancing HDR or, alternatively, inhibiting NHEJ, which can compete with HDR. Furthermore, like NHEJ, HDR traditionally requires the induction of a DSB for its activation, which, as discussed above, can be deleterious to the cell. However, it is important to note that HDR can also be activated by cleaving only one strand of the target DNA (Ran *et al.*, 2013a), thereby overcoming the risks associated with DSBs. This can be achieved using a Cas9 variant with an inactivating mutation in one of its two nuclease domains (known as a Cas9 nickase or nCas9). Finally, and perhaps most importantly, HDR is not available in all cell types. Because the repair proteins involved in HDR are upregulated in the stages of the cell cycle involving DNA replication, non-proliferative cells (which can include neurons and myocytes) are typically unable to incorporate sequences from donor templates at rates that are comparable to actively dividing cells. Thus, HDR-based methods for genome editing are more often employed in proliferating cell types.

Emerging Technologies for Genome Editing: Base Editing and Prime Editing

While NHEJ and HDR provide means for introducing a range of modifications in various cell types, these pathways nonetheless possess several limitations that can prevent their implementation for certain applications. As described above, NHEJ is an error-prone and low-fidelity DNA repair pathway, while HDR is sometimes inefficient and restricted to proliferating cells. Moreover, both of these pathways rely on a DSB for their activation, which by itself, can cause numerous side-effects in cells. Cas9-induced DSBs, for example, can induce karyotypic abnormalities, trigger chromosomal rearrangements and activate the tumor suppressor protein TP53 (Kosicki *et al.*, 2018; Haapaniemi *et al.*, 2018), all of which can compromise the integrity and/or survival of a gene-edited cell. To address this need, new technologies have been developed that harness the flexible targeting capabilities of Cas9 but rely instead on physically tethered functional domains to mediate DNA editing. Importantly, these genome-modifying enzymes can edit DNA in the absence of a DSB. Two such tools are base editors and prime editors.

Among the base editing platforms that have been developed to date, two of the commonly used are cytosine base editors, abbreviated CBEs, which facilitate the conversion of a cytosine base to a thymine base, and adenosine base editors, nicknamed ABEs, which can catalyze the conversion of an adenosine base to a guanosine base. More specifically, CBEs consist of fusions of nCas9 with a cytidine deaminase enzyme (for example the rat APOBEC1 deaminase domain) (Komor *et al.*, 2016), whereas ABEs consist of an nCas9 protein fused to an *Escherichia coli* TadA adenine deaminase domain which, though natively an enzyme that acts on transfer RNA (tRNA), has been engineered to modify DNA substrates (Gaudelli *et al.*, 2017).

Similar to a traditional Cas9 nuclease, base editors are directed to a specific target site via RNA-DNA base complementarity by a gRNA. Then, the tethered cytosine or adenosine deaminase domain binds to individual cytosine or adenosine bases within a narrow editing window – usually a stretch of nucleotides four to eight positions from the beginning of the spacer sequence – to catalyze their transformation to, in the case of CBEs, uracil (a nucleobase recognized by cells as thymidine) or, in the case of ABEs, inosine (a nucleobase recognized by cells as guanosine). The resulting U-G and I-T mismatches are then resolved into the target T-A or G-C base-pairs, respectively, by the action of cellular DNA replication and repair mechanisms that are activated by nCas9-induced nicking of the non-edited strand. Notably, in addition to C-to-T editing, CBEs have also been co-opted to convert a cytosine base to a guanosine, which can be accomplished by fusing CBEs to either an *E. coli*-derived uracil DNA N-glycosylase enzyme or other base excision repair proteins (Zhao *et al.*, 2021) that can excise uracil, thereby creating abasic sites that, after DNA repair, can be converted to a guanosine.

A major advantage of base editors over traditional Cas9 nucleases is that they do not require a DSB to edit DNA, thereby avoiding not only the intrinsic functional limitations of NHEJ and HDR, but also the safety concerns associated with DSBs, which are described above. This latter point is particularly relevant for therapeutic applications of gene-editing (Porto *et al.*, 2020). Much like other genome-editing

platforms, however, base editors possess certain limitations. For one, CBEs can, in some instances, introduce C-to-non-T edits, both at the target base and at non-target bases at the target site and at off-target sites. To address this limitation, base editing technologies with increased product purity – a term used to describe the homogeneity of the edit – have been developed. For example, CBEs have been engineered to contain two copies of uracil glycosylase inhibitors (UGIs) tethered to the C-terminus of the nCas9 protein (Komor *et al.*, 2017). These UGIs can prevent endogenous base excision repair proteins from excising a CBE-induced uracil, which can form abasic sites that can result in the incorporation of non-T bases. Second, because base editors rely on deaminase domains that can also naturally edit RNA, they have the capacity to edit mRNA and potentially other cellular RNAs. To this end, mutations have been introduced into the deaminase domains of base editor proteins to reduce their capacity to bind RNA, which in turn can reduce their RNA editing activity (Zhou *et al.*, 2019). Finally, due to their reliance on cytosine and adenosine deaminases that can carry out only a limited number of base conversions, base editors are limited in the range of edits they can introduce.

However, unlike base editors, prime editors have the capability to introduce not only all 12 possible base-to-base conversions but also a range of base insertions and deletions, all without a DSB. To date, the most commonly used iteration of a prime editor consists of a nCas9 domain fused to a murine leukemia virus reverse transcriptase (RT) domain and a chimeric gRNA named the prime editing gRNA (pegRNA) which not only specifies the target site for the nCas9 protein, but also serves as a template for the RT domain during DNA repair (Anzalone *et al.*, 2019).

Prime editors are first directed to a specific target site by RNA-DNA base complementarity via the pegRNA. This then results in nCas9-induced nicking of the target strand and the subsequent hybridization of the nicked 3′ DNA strand with a complementary primer binding site encoded in the pegRNA. The RT domain then catalyzes the synthesis of a new DNA strand from the template encoded within the pegRNA, which is then introduced into the target site after DNA repair. To date, this modality has been utilized to both model and correct pathogenic mutations in cell culture and in animal models (Liu *et al.*, 2021). Though this technology remains in its infancy, prime editing holds the potential to become a standard tool for many applications because of the range of outcomes it can generate.

Alternate CRISPR Systems and Their Applications for Genome Editing

Competitive co-evolution between microbes and viruses has led to the diversification of CRISPR–Cas loci against invading elements. Thus, there exists broad genetic and functional diversity in CRISPR-Cas systems. In particular, due to the widespread success of CRISPR-Cas9, the mining and characterization of unique CRISPR-Cas systems has become a powerful strategy for identifying Cas variants with unique functionalities that can expand the capabilities of our genome-editing toolbox.

One example of this is Cas12 effectors, which, unlike Cas9, can naturally process a pre-crRNA array. More specifically, Cas12 proteins can cleave the direct repeat (DR) sequences that separate the spacer within a pre-crRNA array, thereby creating multiple mature crRNA molecules (Yan *et al.*, 2019). This feature has been co-opted for genome engineering, as Cas12 can be used to introduce multiplex edits via the delivery of just a single pre-crRNA. This is in contrast to Cas9, which, for multiplexing, requires that each gRNA is expressed individually, a requirement that could pose delivery challenges. In addition to processing pre-crRNA molecules, Cas12 can also collaterally cleave non-specific single-stranded DNA (ssDNA) following its cleavage of a target DNA (Chen *et al.*, 2018b), a functionality that has been exploited for diagnostics (more specifically, this is achieved by using a ssDNA probe that can be designed to emit a fluorescent signal following its trans-cleavage by a Cas12 protein that is unlocked following its cleavage of a target DNA sequence).

In addition to Cas proteins with unique processing and trans-cleaving capabilities, CRISPR systems that are capable of targeting RNA have also been discovered. These elements, known as Cas13 effectors, possess an intrinsic RNase activity that is unlocked following the binding of a target RNA sequence via its crRNA guide molecule (Tang *et al.*, 2021). Analogous to Cas9, this crRNA encodes a programmable spacer that mediates target engagement through complementary base-pairing, though, in the case of Cas13, recognition is mediated by RNA-RNA base complementarity. Upon hybridization with the target RNA, Cas13 undergoes a conformational change that activates its intrinsic RNAse activity, resulting in the cleavage of the complexed RNA (Tang *et al.*, 2021). Interestingly,Cas13 can then cleave non-target RNAs indiscriminately following its activation, a process which may have evolved to help combat a viral infection (Meeske *et al.*, 2019).

Given their ability to cleave RNA in a programmable manner, Cas13 proteins have been adapted to target RNA in mammalian cells, enabling such outcomes as target gene knockdown (Yan *et al.*, 2019). Similar to Cas9, Cas13 proteins have also been repurposed to enable RNA editing. Specifically, catalytically inactivated forms of Cas13 can be tethered to deaminase enzymes, forming single-base RNA editors (Tang *et al.*, 2021). Further, inactivated Cas13 can be directed to bind splice acceptor and donor sites in RNA, where they can then block splicing and induce exon skipping (Konermann *et al.*, 2018). Thus, Cas13 effectors are a versatile platform for transcriptome engineering.

Using CRISPR Technology to Regulate Gene Expression

Methods capable of controlling the expression of a target gene have broad potential for identifying and studying gene function and cellular processes. Due to its ability to bind DNA in a programmable manner, Cas9 can be repurposed for transcriptional regulation through a technique known as CRISPR interference (CRISPRi).

CRISPRi consists of a catalytically inactivated version of the Cas9 protein (named dCas9) and a gRNA molecule, which directs the dCas9 protein to a target sequence. However, by targeting dCas9 protein to sites within a gene sequence or a promoter region, transcription can be blocked through steric hindrance, thereby decreasing the expression of the target gene (Larson *et al.*, 2013).

Thus, CRISPRi provides a sequence-specific means for modulating expression. Importantly, some regions of the genome may be more susceptible to CRISPRi than others. For instance, gRNAs targeting regions with lower nucleosome occupancy or higher chromatin accessibility can be more effectively repressed by CRISPRi (Smith *et al*., 2016). However, in addition to physically blocking transcription, the dCas9 protein can also be physically linked to an effector domain with the ability to repress transcription. One such repressor is the Krüppel associated box (KRAB) domain, which, when fused to dCas9, can silence an endogenous genomic locus by recruiting chromatin remodeling proteins to the target sequence (Gilbert *et al*., 2013).

dCas9, however, can also be used to activate gene expression. Analogous to the dCas9 repressors described above, dCas9 can be tethered to transactivator domains that, when directed to a promoter sequence in a gRNA-dependent manner, can activate a target gene. The simplest of these CRISPR activator (CRISPRa) proteins rely on VP64, a transactivator domain consisting of four tandem copies of the herpes simplex viral protein 16, to mediate transcriptional activation (Perez-Pinera *et al*., 2013). However, since their creation, dCas9 activators have been refined to stimulate gene expression with greater strength than first-generation effectors. These improved activators can consist of fusions of several transactivator domains – such as the VPR protein (Chavez *et al*., 2015) – or rely on an engineered polypeptide scaffold to recruit additional enhancer proteins (Tanenbaum *et al*., 2014). Importantly, because CRISPRi and CRISPRa targeting is coordinated by just a gRNA molecule, both platforms can be used to conduct genome-scale perturbation screens, where they have been used, for example, to discover new drugs targets. Additionally, these techniques hold the potential to advance metabolic engineering by enabling the identification of genes or functional elements whose repression or activation can enhance the production of a desired compound.

Finally, dCas9 can also be harnessed to regulate gene expression by directly influencing the epigenetics of a gene sequence. For instance, dCas9 can be tethered to demethylase proteins, such as DNA methyltransferase 3 alpha (DNMT3A) (Vojta *et al*., 2016) or DNA methyltransferase MQ1 (Lei *et al*., 2017), to coordinate the methylation of specific cytosine bases within a promoter sequence, thereby regulating gene expression. However, dCas9-based effectors can also influence gene expression by modifying histone proteins. For example, dCas9 can be tethered to effector domains that carry histone acetyltransferase activity, which can be leveraged to increase the levels of acetylation of core histone protein within a promoter region (Hilton *et al*., 2015), thereby epigenetically activating transcription.

Thus, the Cas9 protein is a valuable scaffold protein that can be used to modulate target gene expression by physically blocking transcription, recruiting repressor or activator proteins to a promoter sequence or by epigenetically modifying a gene sequence.

Conclusion

CRISPR technology has transformed the life sciences. Due in part to its efficiency and the ease with which it can be configured, CRISPR has emerged as a foundational tool for many biologists, accelerating the pace of scientific discoveries in a wide-ranging number of fields. Importantly, the CRISPR toolbox is continuously expanding, as scientists continue to discover new CRISPR proteins with unique capabilities and engineer new CRISPR modalities with expanded or refined functionalities. Thus, CRISPR technology is poised to continue re-defining the boundaries of genome editing and enabling advances across the life sciences.

References

Anderson, E.M., Haupt, A., Schiel, J.A., *et al*., 2015. Systematic analysis of CRISPR–Cas9 mismatch tolerance reveals low levels of off-target activity. Journal of Biotechnology 211, 56–65.

Anzalone, A.V., Randolph, P.B., Davis, J.R., *et al*., 2019. Search-and-replace genome editing without double-strand breaks or donor DNA. Nature 576, 149–157.

Chang, H.H.Y., Pannunzio, N.R., Adachi, N., Lieber, M.R., 2017. Non-homologous DNA end joining and alternative pathways to double-strand break repair. Nature Reviews Molecular Cell Biology 18, 495–506.

Chavez, A., Scheiman, J., Vora, S., *et al*., 2015. Highly efficient Cas9-mediated transcriptional programming. Nature Methods 12, 326–328.

Chen, D., Tang, J.-X., Li, B., *et al*., 2018a. CRISPR/Cas9-mediated genome editing induces exon skipping by complete or stochastic altering splicing in the migratory locust. BMC Biotechnology 18, 60.

Chen, J.S., Ma, E., Harrington, L.B., *et al*., 2018b. CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. Science 360, 436–439.

Fu, B.X.H., St Onge, R.P., Fire, A.Z., Smith, J.D., 2016. Distinct patterns of Cas9 mismatch tolerance in vitro and in vivo. Nucleic Acids Research 44, 5365–5377.

Gaudelli, N.M., Komor, A.C., Rees, H.A., *et al*., 2017. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. Nature 551, 464–471.

Gilbert, L.A., Larson, M.H., Morsut, L., *et al*., 2013. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell 154, 442–451.

Haapaniemi, E., Botla, S., Persson, J., Schmierer, B., Taipale, J., 2018. CRISPR–Cas9 genome editing induces a p53-mediated DNA damage response. Nature Medicine 24, 927–930.

Hilton, I.B., D'ippolito, A.M., Vockley, C.M., *et al*., 2015. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nature Biotechnology 33, 510–517.

Jinek, M., Chylinski, K., Fonfara, I., *et al*., 2012. A programmable Dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816–821.

Kleinstiver, B.P., Pattanayak, V., Prew, M.S., *et al*., 2016. High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. Nature 529, 490–495.

Komor, A.C., Zhao, K.T., Packer, M.S., *et al*., 2017. Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity. Science Advances 3.eaao4774.

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, 420–424.

Konermann, S., Lotfy, P., Brideau, N.J., *et al*., 2018. Transcriptome engineering with RNA-targeting type VI-D CRISPR effectors. Cell 173, 665–676. e14.

Kosicki, M., Tomberg, K., Bradley, A., 2018. Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. Nature Biotechnology 36, 765–771.

Larson, M.H., Gilbert, L.A., Wang, X., *et al*., 2013. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. Nature Protocols 8, 2180–2196.

Lei, Y., Zhang, X., Su, J., *et al*., 2017. Targeted DNA methylation in vivo using an engineered dCas9-MQ1 fusion protein. Nature Communications 8, 16026.
Liu, P., Liang, S.-Q., Zheng, C., *et al*., 2021. Improved prime editors enable pathogenic allele correction and cancer modelling in adult mice. Nature Communications 12, 2121.
Meeske, A.J., Nakandakari-Higa, S., Marraffini, L.A., 2019. Cas13-induced cellular dormancy prevents the rise of CRISPR-resistant bacteriophage. Nature 570, 241–245.
Neldeborg, S., Lin, L., Stougaard, M., Luo, Y., 2019. Rapid and efficient gene deletion by CRISPR/Cas9. In: Luo, Y. (Ed.), CRISPR Gene Editing: Methods and Protocols. New York, NY: Springer New York.
Perez-Pinera, P., Kocak, D.D., Vockley, C.M., *et al*., 2013. RNA-guided gene activation by CRISPR-Cas9–based transcription factors. Nature Methods 10, 973–976.
Porto, E.M., Komor, A.C., Slaymaker, I.M., Yeo, G.W., 2020. Base editing: advances and therapeutic opportunities. Nature Reviews Drug Discovery, 19, 839–859.
Ran, F.A., Hsu, P.D., Lin, C.-Y., *et al*., 2013a. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell 154, 1380–1389.
Ran, F.A., Hsu, P.D., Wright, J., *et al*., 2013b. Genome engineering using the CRISPR-Cas9 system. Nature Protocols 8, 2281–2308.
Smith, J.D., Suresh, S., Schlecht, U., *et al*., 2016. Quantitative CRISPR interference screens in yeast identify chemical-genetic interactions and new rules for guide RNA design. Genome Biology 17, 45.
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, 62–67.
Tanenbaum, M.E., Gilbert, L.A., Qi, L.S., Weissman, J.S., Vale, R.D., 2014. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. Cell 159, 635–646.
Tang, T., Han, Y., Wang, Y., Huang, H., Qian, P., 2021. Programmable system of Cas13-mediated RNA modification and its biological and biomedical applications. Frontiers in Cell and Developmental Biology 9.
Vojta, A., Dobrinić, P., Tadić, V., *et al*., 2016. Repurposing the CRISPR-Cas9 system for targeted DNA methylation. Nucleic Acids Research 44, 5615–5628.
Wiedenheft, B., Sternberg, S.H., Doudna, J.A., 2012. RNA-guided genetic silencing systems in bacteria and archaea. Nature 482, 331–338.
Yan, F., Wang, W., Zhang, J., 2019. CRISPR-Cas12 and Cas13: The lesser known siblings of CRISPR-Cas9. Cell Biology and Toxicology 35, 489–492.
Zeng, F., Beck, V., Schuierer, S., *et al*., 2020. A simple and efficient CRISPR technique for protein tagging. Cells 9.
Zhao, D., Li, J., Li, S., *et al*., 2021. Glycosylase base editors enable C-to-A and C-to-G base changes. Nature Biotechnology 39, 35–40.
Zhou, C., Sun, Y., Yan, R., *et al*., 2019. Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. Nature 571, 275–278.