INDONESIAN JOURNAL OF LIFE SCIENCES

Volume 6, No. 1, March 2024

REVIEW ARTICLE

CRISPR-Cas9: A Story of Discovery, Innovation, and Revolution in Genome Editing

Rio Hermantara*

Department of Biomedicine, Institut Bio Scientia Internasional Indonesia, Jakarta, Indonesia

**Corresponding author: rio.hermantara@i3l.ac.id*

A B S T R A C T

CRISPR-Cas9 is a powerful and flexible genome editing technology that has transformed the field of life sciences. It has enabled the creation and modification of genomic sequences from various organisms, with wide-ranging implications for biotechnology, agriculture, and medical research. The origin of CRISPR as an adaptive immune system in bacteria, the elucidation of CRISPR-Cas9 biological and molecular mechanisms, and the subsequent engineering and optimization of it as a programmable genome editing tool are remarkable achievements in science. In this review, we summarized this fascinating history and also explored the recent innovations of the CRISPR-Cas9 system, which has extended the genomic toolbox to improve the accuracy and efficiency of CRISPR-Cas9 genome editing.

K E Y W O R D S

CRISPR, Cas9, Genome editing

H I G H L I G H T S

- ❖ CRISPR system is an elaborate adaptive immune system of bacteria, discovered through the 'build up' of scientific discoveries since 1987
- ❖ CRISPR is a guided endonuclease system that targets complementary sequences of the CRISPR RNA.
- ❖ The mostly used CRISPR systems for genome editing, CRISPR Cas9, requires two components, the Cas9 protein and one single guide RNA molecule.
- ❖ Multiple advancements were made to improve Cas9 activity, specificity, and adaptability

INTRODUCTION

At the end of the year 2023, multiple drug regulatory bodies from the United Kingdom, the United States of America, and European countries approved the first CRISPR genome editing-based therapy to treat sickle cell anemia and beta-thalassemia (Sheridan, 2023). This unprecedented event, in which the gene therapy offers a permanent cure to its patients, culminates the remarkable scientific journey of CRISPR-Cas9 as a genome editing tool.

CRISPR, an acronym for Clustered Regularly Interspaced Short Palindromic Repeats, is a natural adaptive immune mechanism of bacteria and archaea (Jansen et al., 2002; Mojica et al., 2005; Pourcel et al., 2005). At its core, CRISPR acts by the action of Cas9, short for CRISPR-associated protein 9, which is an endonuclease. These Cas9 nucleases are guided by short RNA molecules targeting complementary sequences, which are adopted by the system from invading viruses (Marraffini & Sontheimer, 2008). Once bound to the target, Cas9 nuclease-induced DNA double-strand break, which triggers viral clearance or

cellular DNA repair. Doudna and Charpentier's group adapted this system, resulting in the creation of an engineered guide RNA (gRNA) that has developed this revolutionary genome editing tool (Jinek et al., 2012).

Prior to CRISPR, the genetic editing landscape was once dominated by complex, time-consuming techniques, such as homologous recombination, zinc finger nucleases (ZFNs), and transcription-like activation factor nucleases (TALENs) (González Castro et al., 2021). Since its establishment, CRISPR-Cas9 has quickly become the most used genome editing tool by offering simplicity, versatility, and accessibility. This technology has pushed the boundaries of conventional genetic engineering, giving researchers unprecedented access to editing genes in a diverse range of organisms, from bacteria (Yao et al., 2018) to plants (Zhu et al., 2020) and animals (Lin et al., 2021), including humans (Huang et al., 2022; Li et al., 2023).

In this review, we will explore the fundamental principles of CRISPR-Cas9, starting from its discovery and adaptation as a genome editing tool to its advancements in increasing the genome editing toolbox. By delving into the intricacies of this revolutionary genome editing tool, we aim to provide a comprehensive overview of CRISPR-Cas9's history and advancements, contributing to the ongoing discourse and supporting its use and application in life science research.

THE DISCOVERY: CRISPR SYSTEM IN BACTERIA AND ARCHAEA

Back in 1987, while attempting to sequence the *iap* isozyme-converting gene in *Escherichia coli,* Ishino and colleagues observed unusual DNA sequences downstream of the *iap* gene. Five short, highly conserved repeats were spaced with a low homology sequence in between (Ishino et al., 1987). The transcripts of these sequences result in the formation of RNA hairpin loops due to the palindromic conserved sequences. While Ishino and colleagues did not provide any hypotheses on its function, they noted that this unique structure might have a particular biological function. As more sequences were made available, multiple groups reported similar findings in different bacterial species, yet with distinct repeat sequences (Figure 1) (Mojica et al., 2000).

In 2002, Jansen and colleagues performed an *in-silico* analysis on these unique repeats from sequences deposited in the genebank database. With the repeats being unique within a species, they managed to systematically characterize these repeat motifs, such as (i) multiple short repeats were present in all analyzed sequences, (ii) there was the presence of a non-repetitive spacer sequence between repeats, (iii) a common leader sequence was present upstream of the repeat-non-repeat array, (iv) the absence of an open reading frame, which indicated that the sequences were not translated, and (v) the presence of CRISPR-associated (Cas) genes proximal to the loci (Jansen et al., 2002). An extended search found that more than 40 different species shared this unique characteristic sequence motif. The acronym CRISPR (Clustered Regularly Interspaced Palindromic Repeats) was proposed to collectively combine these findings (Jansen et al., 2002).

The widespread conservation of the CRISPR loci and the uniqueness of the sequence structures suggested its important biological function in prokaryotes. Extensive sequencing analysis by multiple groups identified that spacer sequences, which are the non-repetitive sequences between repeats, originate from foreign DNA, such as bacteriophage or external plasmid DNA. This indicates the role of CRISPR in prokaryotic adaptive immunity (Mojica et al., 2005; Pourcel et al., 2005). These suggestions were confirmed by Barrangou and colleagues in 2007. Genomic analysis from bacteriophage-resistant *Staphylococcus thermophilus* mutants revealed that phage-specific sequences were acquired and inserted as spacer sequences in their CRISPR array. It was further demonstrated in the same study that phage-specific spacer sequences direct resistance to those specific phages in *S. thermophilus* (Barrangou et al., 2007).

A deeper insight into the biological mechanism of CRISPR was shown in an experiment by Marraffini and Sontheimer in 2008, using *Staphylococcus epidermis* strains that contain a spacer sequence targeting a

known *Staphylococcus* conjugative plasmid, they revealed that the CRISPR machinery interacts and acts directly on plasmid DNA and prevents transformation. Garneau and colleagues expedite the investigation by observing the fate of bacteriophage DNA in *S. thermophilus* following bacteriophage infection. It was revealed that the cleavage site of the CRISPR system is homologous to the CRISPR spacer sequence (Garneau et al., 2010). Two key *in vitro* studies that followed showcased the mechanism of CRISPR-Cas9 as an RNA-guided DNA endonuclease, with the potential to be used as a programable nuclease for precise genome editing (Gasiunas et al., 2012; Jinek et al., 2012).

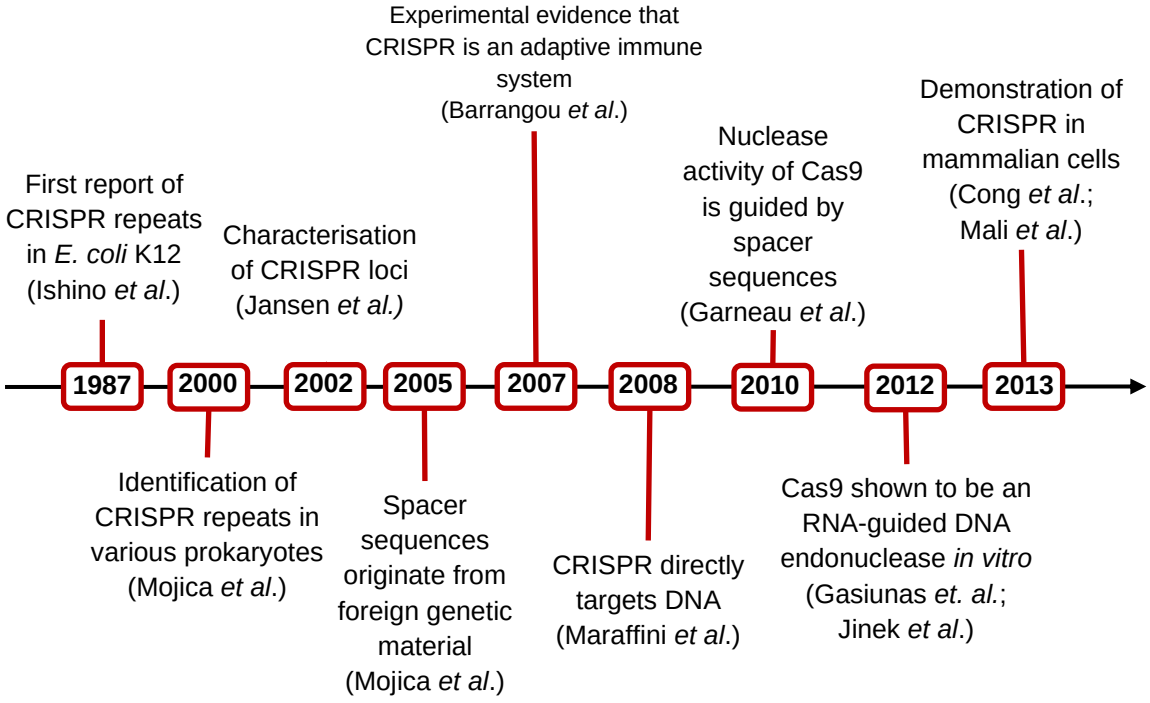

Figure 1. Key events during the discovery and characterization of CRISPR. A schematic timeline of key events of the discovery, characterization, and development of the CRISPR-Cas9 system as a genome editing tool up until the first proof-of-principle of CRISPR-Cas9 genome editing in mammalian cells.

In August-September of 2012, two reports from Jennifer Doudna and Emmanuelle Charpentier's groups and Virginijus Siksnys's group, respectively, were published. Both groups were working independently with different bacteria species, but their findings convey similar information. Collectively, they managed to demonstrate the mechanism of targeted double-stranded DNA (dsDNA) cleavage by CRISPR-Cas9. One major difference between the groups is that Doudna and Charpentier's group engineered a single chimeric RNA that could replace the crRNA:tracrRNA hybrid RNA for targeted nuclease activity. This enhanced the simplicity and versatility of the CRISPR-Cas9 system, designating it to be a programmable nuclease system that precisely cleave a genomic target that is complement to the guide RNA. Subsequently, multiple groups demonstrated the use of the CRISPR-Cas9 system in editing the genome of mammalian cells a year after these in vitro studies (Cong et al., 2013; Jinek et al., 2013; Mali, et al., 2013). Before we cover the applications of CRISPR-Cas9 as a genome editing tool, the biological mechanism of the CRISPR-Cas9 system must first be reviewed.

CRISPR-CAS9: A GUIDED ENDONUCLEASE AS AN ADAPTIVE IMMUNE SYSTEM IN BACTERIA

There are three phases of CRISPR adaptive immunity: (i) spacer adaptation, (ii) crRNA biogenesis or maturation, and (iii) targeted interference (Figure 2). Each stage is reviewed thoroughly, focusing on the biological aspects of the most widely known and used type II CRISPR-Cas9 system.

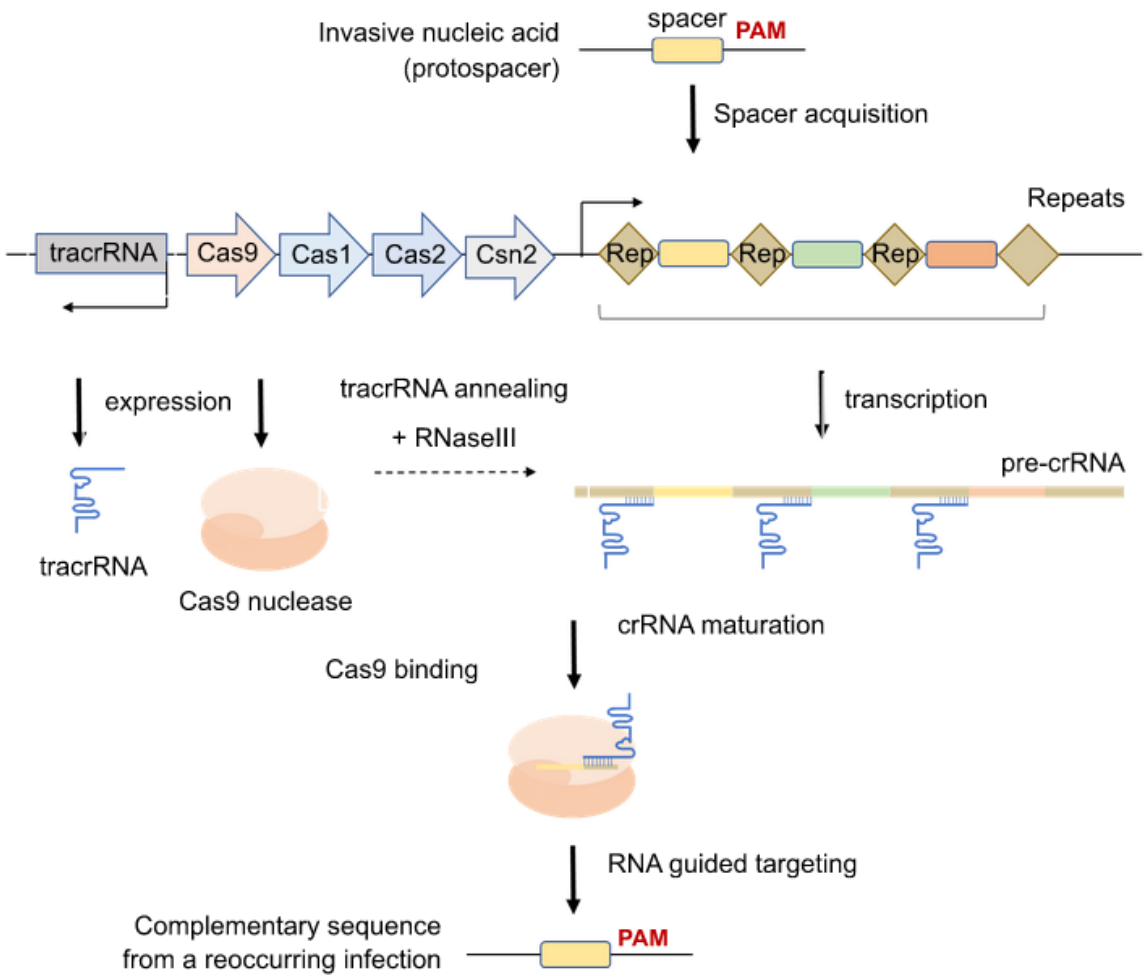

Figure 2. CRISPR-Cas immunity system in type II of *S. pyogenes*. Adaptive immunity is acquired through spacer acquisition facilitated by all the Cas effector proteins. New spacer sequences are integrated into the CRISPR array downstream of the leader sequences. Expression of the CRISPR loci produces pre-crRNA which will bind to tracrRNA to initiate crRNA maturation. Pre-crRNA:tracrRNA complex is cleaved by RNaseIII and binds to Cas9. Further processing involves an unknown nuclease which trims the 5 ends of the crRNA yielding a mature crRNA for Cas9 activation. crRNA-driven Cas9 targets sequences with complementary spacer sequences upstream of a PAM motif, protecting the bacteria from reoccurring infections.

Spacer acquisition and integration to the CRISPR array

The CRISPR adaptive immunity emanates from the ability of the CRISPR system to store memory of previous infections by acquiring DNA sequences, known as spacer sequences (Marraffini & Sontheimer, 2008; Mojica et al., 2009; Garneau et al., 2010; Deltcheva et al., 2011; Makarova et al., 2015). The major machinery involved in this spacer adaptation are Cas1 and Cas2 protein effectors (Makarova et al., 2015). Depending on the subtype, additional auxiliary acquisition proteins might be needed as well. For example, Csn2 is required for spacer acquisition in type II-A, as well Cas4 for type II-B (Chylinski et al., 2014; Wright & Doudna, 2016).

Cas1 is the most conserved cas effector in the CRISPR system (Koonin et al., 2017). The catalytic function of Cas1 is induced by binding to Cas2, which acts as a structural protein that stabilizes the complex (Arslan et al., 2014; Chylinski et al., 2014). In type II-A, Csn2 binds to the complex to form an active conformation. It is thought that with the high affinity to DNA ends, Csn2's role in the complex is to protect the DSB ends from further degradation (Heler et al., 2015; Wei et al., 2015; Wilkinson et al., 2019).

Spacer acquisition is initiated by PAM-proximal sampling. The Cas1:Cas2 complex requires a DNA-bound Cas9 to initiate spacer acquisition (Heler et al., 2015; Wei et al., 2015). Electron imaging of crystalized structures reveals that Cas1, Cas2, Csn2, and Cas9 form a complex and undergoes changes in their conformation, suggesting that spacer acquisition is done through multiple steps (Wilkinson et al., 2019). Previous studies further suggested that Cas9 used in this complex does not have a catalytic activity, which supports previous findings (Heler et al., 2015; Wei et al., 2015; Wilkinson et al., 2019). It is thought that an unknown non-Cas nuclease might be recruited to the complex to cleave the DNA-bound complex (Wilkinson et al., 2019).

Upon obtaining the new spacer sequence, the Cas9 dissociates from the complex and the Cas1-Cas2-csn2 complex locates the CRISPR array. New spacers are predominantly added at the leader sequence of the CRISPR array (Barrangou et al., 2007; Datsenko et al., 2012; Wei et al., 2015). Cas1 contains a leader-recognizing structure, which makes a sequence-specific contact with the leader sequence through the leader anchoring sequence (LAS) (Rollie et al., 2015; Wright and Doudna, 2016). Unlike type I which requires an additional integrase host factor (IHF), type II Cas1-Cas2 sufficiently adds new spacers into the CRISPR array without any additional host factors (Wright & Doudna, 2016; Xiao et al., 2017).

Spacer integration into the leader sequence is catalyzed by a nucleophilic attack by the terminal 3' OH new spacer on the leader and the first repeat end of the CRISPR array (Wright &Doudna, 2016; Xiao et al., 2017). The 5' end of the spacer is ligated to the leader proximal end, while the 3' end is ligated to the repeat end which is directed by DNA bending. (Xiao, et al., 2017). Upon ligation, this creates an intermediate product of dsDNA-ssDNA (Mcginn and Marraffini, 2019). Bacterial DNA polymerase is recruited to the site to fill in the gaps and complete spacer integration and repeat duplication (Wright & Doudna, 2016; Xiao et al., 2017; Wilkinson et al., 2019).

crRNA biogenesis

Type II crRNA maturation requires the use and interaction of tracrRNA (Deltcheva et al., 2011; Gasiunas et al., 2012; Jinek et al., 2012). TracrRNAs are typically 75 to 100 nucleotides long, containing conserved 'anti-repeat' sequences to facilitate crRNA base pairing (Figure 3) (Chylinski et al., 2013; Karvelis et al., 2013). Bioinformatic and crystal structure studies have revealed the secondary structures of tracrRNA, which include stem-bulge, nexus, and hairpin 1 and 2 formations (Anders et al., 2014; Nishimasu et al., 2014; Jiang et al., 2015). Briner and colleagues conducted a study to assess the function of these structures and found that the stem and nexus interact directly with the arginine-rich bridge, promoting proper positioning of the guide RNA into the Cas9 protein (Briner et al., 2014). The hairpin structures, together with the nexus, dictate the orthogonality of the guide RNA and the Cas9 protein (Briner et al., 2014).

Formation of a crRNA:tracrRNA double-stranded RNA (dsRNA) duplex is stabilized by Cas9 binding to the CRISPR array (Deltcheva et al., 2011; Chylinski et al., 2013). A RNaseIII ribonuclease is then recruited to cut the 3' end of the pre-crRNA releasing individual guide RNA:Cas9 complex (Deltcheva et al., 2011). Further 5' end crRNA trimming was done by an unknown exonuclease to generate mature guide RNA (Deltcheva et al., 2011).

CRISPR-Cas9 interference

The *Streptococcus pyogenes* Cas9 (SpCas9) nuclease consists of two large domains, the alpha-helical recognition (REC) lobe and the nuclease (NUC) lobe, which are linked together by an arginine-rich bridge and a non-conserved linker. The REC lobe is composed of three alpha-helical domains which interact with the gRNA and are responsible for gRNA:Cas9 complex stabilization (Jinek et al., 2014; Sternberg et al., 2014). The NUC lobe contains two nuclease domains, the HNH and RuvC, as well as a C-terminal domain (CTD) (Gasiunas

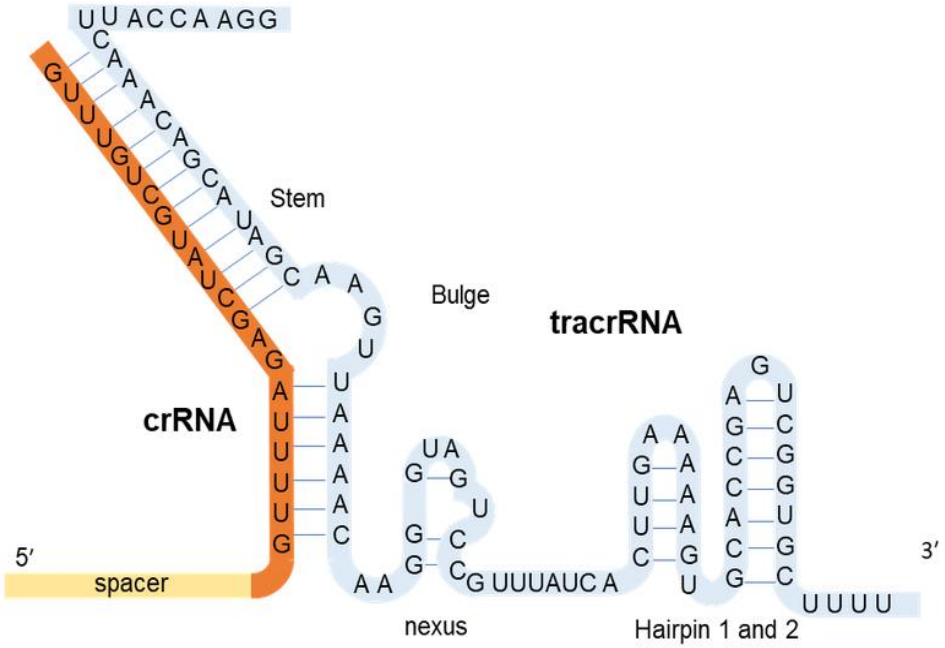

Figure 3. The crRNA and tracrRNA duplex. The crRNA (orange line), which carries the target spacer sequence (yellow line) at the 5′ end, hybridized with the tracrRNA (blue line), forming the crRNA:tracrRNA duplex through the formation of hydrogen bonds (blue strips). The 3′ end of the tracrRNA sequence folds into nexus, hairpin 1, and 2, acting as a functional structure for binding to the Cas9 protein.

et al., 2012; Jinek et al., 2012; Jinek et al., 2014). PAM recognition site is within the CTD, which is disordered in the absence of a guide RNA (Sternberg et al., 2014).

Guide RNA recruitment initiates major conformational changes of inactive (apoprotein) Cas9 conformation to a DNA recognition-competent conformation (Figure 4) (Jinek et al., 2014; Jiang et al., 2015). This supported previous studies that established the prerequisite guide RNA to Cas9 binding for DNA-targeted cleavage (Gasiunas et al., 2012; Jinek et al., 2012). Crystal structure studies of sgRNA-bound Cas9 further reveal that stabilization was achieved by protein-to-RNA interaction through extensive hydrogen bonds within the stem and hairpin structures (Jiang et al., 2015). On the other end, the bound gRNA is reordered to expose 10 nucleotides upstream of the PAM site in an A-form helix conformation (Künne et al., 2014). PAM-distal sequences of the spacer are placed between the HNH and RuvC domains, protecting the gRNA from RNA degradation (Jiang et al., 2016). Conformational changes were also seen at the CTD, in which PAM interacting arginine sites are prepositioned to an active DNA-sampling position to initiate DNA sampling for PAM sequences (Jiang et al., 2015).

Cas9 will then probe dsDNA for PAM sequences and dissociate quickly at non-PAM sites (Marraffini & Sontheimer, 2010; Jinek et al., 2012; Sternberg et al., 2014). The SpCas9 recognizes NGG PAM motif with N being any other base (Mojica et al., 2009). The interactions between the Cas9 complex and the PAM was revealed by PAM duplex-bound structure studies, in which the first nucleotide of the PAM sequence, N, remains base-paired and does not interact directly with the protein. On the contrary, the -GG dinucleotides are flipped out to interact with the two arginine residues, R1333 and R1335 at the PAM recognition sites (Anders et al., 2014; Jiang et al., 2016).

Upon appropriate PAM recognition, the Cas9 complex induces local DNA denaturation at the PAM-adjacent site in a unidirectional orientation (Sternberg et al., 2014; Szczelkun et al., 2014). On the non-target strand, stabilization was achieved through van der Waals interaction between PAM proximal sequences and the RuvC domain. The first nucleotide upstream of the PAM (-1 position) is stacked into the PAM duplex to

create an intra-strand base stacking, which further stabilizes the complex. Nucleotide positions -2 and -3 are flipped out and interact with Cas9 protein. Another kink is observed at the -4 position, placing the disordered PAM distal non-target strand into the NUC lobe (Jiang et al., 2016).

On the target strand, a sharp kink at the phosphodiester group upstream of the PAM (referred to as +1 phosphate) was observed, stabilized by a phosphate lock loop by residues K1107-S1109 (Anders et al., 2014). DNA sequences upstream of the PAM were rotated and flipped up to initiate guide RNA-DNA hybridization via Watson-Crick base pairing. The first 10 nucleotides of the PAM-proximal guide sequence, known as the seed sequence, is crucial for DNA cleavage activity (Jinek et al., 2012; Cong et al., 2013; Pattanayak et al., 2013; Sternberg et al., 2014). Mismatches of the seed region greatly decrease or possibly abolish Cas9 activity, although multiple reports have demonstrated that Cas9 activity could be observed at off-target sites with close homologous sequences (Fu et al., 2013; Cong et al., 2013; Pattanayak et al., 2013; Zheng et al., 2017).

RNA:DNA hybridization in the seed region will continue gRNA invasion and further anneal complementary base pairing at the PAM-distal sites (Jiang et al., 2016). It should be noted that mismatches of the non-seed region are more tolerable by the Cas9 complex and contribute the majority of off-target activity (Wu et al., 2014). Completion of base pairing resulted in a stabilized R-loop formation with a guide RNA:DNA heteroduplex on the target strand and a displaced non-target strand (Sternberg et al., 2014; Szczelkun et al., 2014). Interestingly, the dwell time of the Cas9 complex on the dsDNA is dependent on the complementary sequence and proper PAM sequence at the target DNA (Knight et al., 2015; Ma et al., 2016).

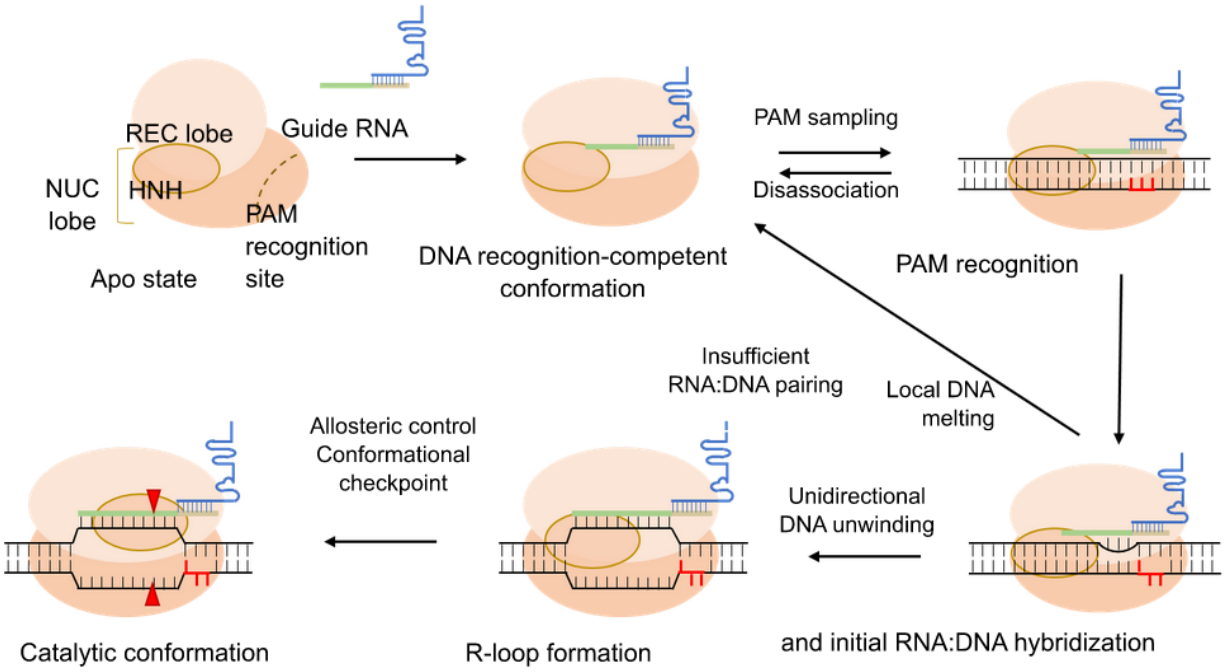

Figure 4. CRISPR interference processing through multiple conformational changes to the Cas9 complex. Guide RNA binding to the Cas9 initiates major conformational changes from an inactive state to a DNA-recognition competent conformation. This complex then performs PAM sampling through the DNA, and dissociates quickly from PAM-absent DNA sequences. PAM recognition will initiate local DNA unwinding, exposing the target DNA to the seeding region sequences of the guide RNA. Insufficient RNA:DNA pairing will result in the dissociation of the Cas9 complex of the DNA sequence. Perfect complementary annealing will further RNA:DNA heteroduplex formation while keeping non-target separated by the RuvC domain, forming the stabilized R-loop formation. Complete annealing of the guide sequence will trigger HNH transitioning to the active site, initiating allosteric control to the RuvC nuclease domain which will trigger DNA nicks on both strands to produce a double-strand break.

Early Chromatin immunoprecipitation (ChIP) studies evaluating off-target activity suggested that Cas9 binding to dsDNA fragment alone is insufficient to initiate Cas9-directed cleavage (Kuscu et al., 2014; Wu et al., 2014). It was later revealed that the Cas9 complex underwent further conformational changes before acquiring the ability to cleave DNA fragments. (Sternberg et al., 2015). Through bulk FRET experiments, Sternberg and colleagues were able to visualize HNH transitioning to a 'docking' active site after the completion of an RNA:DNA heteroduplex formation. PAM distal complementary pairing and divalent cations highly influence this HNH transition (Dagdas et al., 2017). Interestingly, crystal structure studies revealed that the HNH domain does not interact directly with the RNA:DNA heteroduplex (Nishimasu et al., 2014; Jiang et al., 2015), suggesting allosteric interactions between the protein domains.

Other structural studies indicate the role of the REC3 domain in triggering global conformational changes by directly interacting with the RNA:DNA heteroduplex (Nishimasu et al., 2014; Jiang et al., 2015). Additionally, REC2 and REC3 domains were also shown to direct HNH transitioning to the active site (Chen et al., 2017). HNH docking then triggers the folding-unfolding rearrangement of two linkers, which in turn activates the RuvC nuclease domain (Sternberg et al., 2014; Jiang et al., 2016). This globally allosteric cross-talk between REC lobe, HNH, and RuvC domains is critical for the conformational change of Cas9 to a catalytic state (Nishimasu et al., 2014; Sternberg et al., 2015; Jiang et al., 2016).

On Cas9 cleavage activity, both nuclease domains are dependent on the recruitment of ions to stimulate nucleophilic attack by an activated water molecule (Jinek et al., 2012; Nishimasu et al., 2014). The HNH nuclease domain consists of a conserved ββα-fold containing three active site residues: Asp839, His840, and Asn863 (Anders et al., 2014; Jinek et al., 2014; Nishimasu et al., 2014; Jiang et al., 2015). Asp839 and Asn863 recruit oxygen atoms from the scissile phosphate, which in turn coordinate Mg^{2+} recruitment to the HNH domain. This activates the His840 residue, which stimulates a nucleophilic attack by water molecules at the third nucleotide (+3 phosphate) of the target strand (Jinek et al., 2012; Nishimasu et al., 2014).

The RuvC nuclease domain consists of four active sites: Asp10, Glu762, His983, and Asp986 (Jinek et al., 2012; Nishimasu et al., 2014). Asp10, Glu762, and Asp986 recruit oxygen atoms from the scissile phosphate, which then coordinates two ion recruitments to the RuvC nuclease domain (Jinek et al., 2012; Nishimasu et al., 2014). Activated His983 residue activates water molecules to nucleophilically attack the non-target strand. Interestingly, in the initial report by Jinek et al. (2012), the non-target strand is cleaved at multiple sites within three to eight bases upstream of the PAM site. This observation was further elaborated by a molecular dynamic simulation, which suggested preferential binding of the two Mg^{2+} ions at the -4 phosphate instead of the -3 phosphate (Zuo & Liu, 2016). Indeed, further studies using next-generation sequencing platforms revealed that SpCas9 produces multiple end products with predominant blunt-end DNA DSBs (Kim et al., 2016; Gisler et al., 2019).

THE DEVELOPMENT: AS A PROGRAMMABLE GENOME EDITING TOOL

The remarkable precision of the CRISPR system was able to be demonstrated in multiple organisms and transferred into other organisms. The proof-of-concept was first demonstrated in bacteria, where the team of Sapranauskas (2011) and colleagues managed to transfect the CRISPR locus from *S. thermophilus* into *E. coli*. The transformed CRISPR array was able to prevent the transformation of *E. coli* by known transforming plasmid specific to *S. thermophilus* (pSP1 and pSP2). Similar results were also achieved upon challenging the bacteria with bacteriophage infections. These findings implied that the CRISPR system is transferable between bacterial organisms.

In 2012, two major *in vitro* studies managed to redefine the mechanism of CRISPR-Cas9 cleavage, solidifying its possible use as a programmable RNA-guided endonuclease for specific genetic targeting (Gasiunas et al., 2012; Jinek et al., 2012). Three major revelations were achieved from both studies. First,

Cas9 could be directed to target single-stranded DNA (ssDNA) or dsDNA by providing short twenty-nucleotide sequences into the crRNA. This programmable RNA-guided targeting was demonstrated by both groups using two different CRISPR systems from *S. thermophilus* and *S. pyogenes*. They managed to demonstrate that Cas9 cleavage is sequence-specific and cleaves exactly three bases upstream of the PAM. Further characterization indicates that Cas9 activity is sensitive to mutations, as PAM-proximal mutations reduce Cas9 activity while mutations at the PAM sequence abolish the activity entirely. This finding was in agreement with previous *in vivo* studies in *S. thermophilus*, highlighting the importance of PAM recognition for CRISPR-Cas9 interference.

Second, the creation of dsDNA double-strand break was accomplished by the HNH and RuvC nucleases domains, each nicking the target and non-target strRuvaand, respectively (Gasiunas et al., 2012; Jinek et al., 2012). Both groups carried out amino acid mutations in each nuclease domain, reverting the inactivity of the nuclease. Using radio-labeled DNA substrate and PAGE analysis, they managed to demonstrate each nuclease activity. Interestingly, while HNH cleavage is precise at three nucleotides upstream of the PAM, the RuvC creates multiple products, indicating possible exonuclease activity. These nuclease mutations, D10A for HNH mutation and H840A for RuvC mutation, are used widely to create Cas9 nickase and deactivated Cas9 (dCas9) variants (Chen et al., 2013).

Third, the tracrRNA is indispensable for DNA recognition and Cas9 cleavage. Previous studies examining various CRISPR classes revealed the role of tracrRNA in assisting crRNA maturation (Chylinski et al., 2013; Karvelis et al., 2013). Jinek and colleagues managed to notice that the absence of tracrRNA greatly reduces Cas9 binding to DNA through a mobility shift assay. They further proceed with their study by carefully examining the secondary structures of the tracRNA:crRNA hybrid. Moreover, they managed to design a chimeric single guide RNA (sgRNA) which provided equal activity to native tracrRNA:crRNA guided nuclease activity. Altogether, these findings provided evidence of a simple and programmable tool that could potentially be used for universal genome editing.

In light of these revelations, it did not take long for CRISPR to be applied as a genome-editing tool in mammalian cells. Three major publications from three leading laboratories followed in the early months of 2013. The first paper was published in the late-January edition of eLife journal by the Doudna lab. Following their previously *in vitro* studies, Jinek and colleagues managed to demonstrate CRISPR-Cas9 targeted DNA cleavage activity in HEK293T cell lines. Human-codon-optimized Cas9 and sgRNA were cloned into plasmid-expressing systems, where the guide RNA is driven by a human RNA Polymerase III promoter. Cas9 cleavage activity was detected by the occurrence of insertion and deletion (indels) mutations at the target site due to error-prone DNA repair of DNA double-strand breaks (Jinek et al., 2013).

In the mid-February issue of Science, two back-to-back reports from the Church and Zhang labs strengthen this discovery. Using two nuclear locating signals (NLS) fused into the C-terminus of human-codon-optimized Cas9, Cong et al. (2013) reported efficient editing in mouse and human cell lines. Targeting the *EMX1* gene, they managed to report comparable targeting efficiency to transcription activator-like effector nucleases (TALENs). Moreover, efficient multiplex genome targeting was also demonstrated by simultaneous targeting of *EMX1* and *PVLB* genes in the human genome (Cong et al., 2013).

Additionally, the group explored the use of the Cas9 D10A nickase variant for genome editing and found no detectable indels at the target site. This finding is in agreement with previous studies that claimed faithful repair of DNA nicks. Targeted donor integration was also observed by providing a homology repair (HR) template to the transfection reaction. Integration of the donor template to the target site was observed with both Cas9 and Cas9 nickase.

Similar to Zhang's group's findings, Mali et al. (2013) managed to demonstrate individual and multiplexed Cas9 targeting in multiple human cell lines. Furthermore, through a green fluorescent protein (GFP) reporter assay, which requires HR donor exchange to rescue GFP expression, the group revealed that Cas9-mediated DSB could induce site-specific HR to integrate donor DNA sequences. They confirmed their

findings by integrating a dsDNA donor construct at an endogenous locus, the AAVS1 locus, unveiling the feasibility of integrating exogenous sequences into specific sites in the genome (Mali et al., 2013).

These proof-of-concept studies provided clear evidence that the CRISPR-Cas9 system could indeed be applied for targeted genome editing in mammalian cells. Further advancements to the systems were later developed to address current limitations and to expand the genome editing toolbox.

THE ADVANCEMENTS: EXPANDING THE GENOME EDITING TOOL BOX

Since its development, CRISPR-Cas9 technology has been improved greatly. One major contributing factor is the wealth of knowledge in DNA repair mechanisms and the growing understanding of DNA repair responses to CRISPR-Cas9 mediated DSBs (van Overbeek et al., 2016; Brinkman et al., 2018). Here, we will go through multiple strategies and advancements that were made to the CRISPR-Cas9 system, which includes the use of Cas9 orthologs, rationally engineered Cas9 proteins, and Cas9 fusion with other functional proteins.

Expanding Cas9 targeting sites

The SpCas9 is the most characterized and used native Cas9 in research and development. Several SpCas9 orthologs have been identified in several other species (Table 1). Notably, each Cas9 ortholog from each species recognized its own specific PAM sequences. This in turn expands the possible targeting sites in the genome. Furthermore, a few native Cas9 orthologs are smaller in size compared to SpCas9, giving them an advantage for vector packaging in in vivo applications.

Table 1. Cas9 orthologs utilized as a genetic editing tool.

Name	Origin	PAM recognition	Size (amino acids)	Reference
SpCas9	*Streptococcus pyogenes*	NGG	1368	(Jinek et al., 2012)
ScCas9	*Streptococcus cunis*	NNG	1375	(Chatterjee et al., 2018)
StCas9	*Streptococcus thermophilus*	NNAGAAW	1121	(Deveau et al., 2008; Gasiunas et al., 2012)
NmCas9	*Neisseria meningitidis*	NNNNGATT	1082	(Hou et al., 2013)
SaCas9	*Staphylococcus aureus*	NNGRRT	1053	(Ran et al., 2015)
CjCas9	*Campylobacter jejuni*	NNNVRYM	984	(Yamada et al., 2017)

Despite all the options Cas9 orthologs have to offer, due to its targeting efficiency in mammalian cells, SpCas9 remains the most used in research. However, SpCas9 possesses major limitations, such as the following: (i) PAM recognition site limits Cas9 targeting sites and (ii) off-target activity of Cas9 which restricts it from crucial in vivo genetic editing use. A different approach is made by rationally engineering Cas9 proteins to relax the PAM recognition requirement and increase its specificity. These improvements were made possible with the advancement of crystal structure and modeling studies.

Relaxing the PAM recognition requirement would increase possible targeting sites, as it would not be restricted to an -NGG motif. This offers possible genome editing at locations lacking G nucleotides, predictable template-free end joining, and base editing applications (Rees & Liu, 2018; Iyer et al., 2019).

Several engineered Cas9 variants with increased PAM scope have been developed to expand the potential of Cas9 for precise genome editing (Table 2).

Alanine substitution of the PAM recognition residues, R1333 and R1335, abolished the catalytic activity of Cas9 in *in vitro* and *in vivo* studies (Anders et al., 2014; Kleinstiver et al., 2015). This suggested that mutations at the PAM interacting site could be done to re-engineer PAM requirements. Using directed evolution through a bacterial selection system that selects for NGA PAM sites, Kleinstiver et al. (2015) managed to obtain and characterize two Cas9 variants, which are referred to as VQR and VRER variants.

A similar directed evolution system was used by David Liu's group. Instead of using bacteria, they used a phage-assisted continuous evolution (PACE) system to rapidly generate Cas9 variants that recognize multiple PAM sequences. Through this system, they obtained the xCas9 variant, which recognizes broad PAM sequences such as NG, GAA, and GAT (Hu et al., 2018). Additionally, the authors demonstrated that xCas9 has less off-target activity on NGG off-target sites compared to native Cas9.

A different approach was used to generate the SpCas9-NG. Nishimasu et al. (2018) rationally engineered Cas9 by amino acid substitution of the PAM recognition R1335 residue and other residues surrounding the PAM duplex. After characterizing several mutational variants, they identified a seven amino acid substitute variant, referred to as SpCas9-NG. Compared to the previous xCas9, SpCas9-NG produced higher targeting efficiency at NG PAM sites in *in vitro* and *in vivo* experiments.

Table 2. Engineered SpCas9 with broad PAM recognition.

Cas9 Variant	Mutation	PAM recognition	Reference
Cas9-VQR	D1135V, R1335Q, T1337R	NGRN	(Kleinstiver et al., 2015)
Cas9-VRER	D1135V, G1218R, R1335E, T1337R	NGGN	
xCas9	A262T, R324L, S409I, E480K, E543D, M694I, E1219V	NG, GAA, GAT	(Hu et al., 2018)
SpCas9-NG	R1335V, L1111R, D1135V, G1218R, E1219F, A1322R, T1337R	NGNN and NANG	(Nishimasu et al., 2018)
Cas9-NRRH	R1114G, D1135N, V1139A, D1180G, E1219V, Q1221H, A1320V and R1333K	NRRH	(Miller et al., 2020)
Cas9-NRTH	R1114G, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, E1253K, P1321S, D1332G and R1335L	NRTH	
Cas9-NRCH	R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N, S1338T and H1349R	NRCH	
SpG	D1135L, S1136W, G1218K, E1219Q, R1335Q, T1337R	NGN	(Walton et al., 2020)
SpRY	SpG mutations with L1111R, A1322R, A61R, N1317R, R1333P	NHN	

In 2020, David Liu's group, using a phage-assisted non-continuous evolution (PANCE) selection strategy, identified three new Cas9 variants that recognize non-G PAM sequences: NRRH, NRTH, and NRCH sequences (N: any base; R: A or G; H: A or C or T). Mutagenesis analysis at endogenous genes showed comparable targeting efficiencies to native SpCas9. The authors further claimed that with the addition of these new variants, 95% of known pathogenic single nucleotide polymorphisms (SNP) compiled in the ClinVar database can now be targeted and corrected (Miller et al., 2020).

In April 2020, Kleinstiver's lab took the SpCas9-VRER version from 2015 and profiled additional alanine amino acid exchange around the PAM interacting residues to identify a near-PAMless SpCas9 version termed as SpG. This SpG version recognizes NGN PAM sequences and showed comparable activity to native SpCas9. Further rational substitution to the SpG produced a SpRY variant which is capable of targeting NRN, and to a lesser extent, NYN PAM sequences (Walton et al., 2020). Altogether, these engineered variants have expanded the targeting possibilities for human genomic editing.

Rationally engineered Cas9 variants to increase specificity

Early Cas9 nuclease characterization has indicated that CRISPR-Cas complex tolerates one to three mismatches to the gRNA spacer sequence (Jinek et al., 2012; Sternberg et al., 2014). Cas9 binding assays and deep genome sequencing analysis reveal that Cas9 complex binds to off-target sites, often resulting in off-target activity (Pattanayak et al., 2013; Kim et al., 2015). This presents a major issue for therapeutic applications. To resolve this, enhanced specificity Cas9 variants have been engineered (Table 3).

Table 3. Engineered Cas9 variants to enhance targeting specificity.

Variant	Mutation	Engineering method	Reference
dCas9	D10A and H840A	Rationally engineered	(Chen et al., 2013; Gilbert et al., 2012)
Cas9 nickase	D10A (RuvC) or H840A (HNH)	Rationally engineered	(Mali et al., 2013; Ran et al., 2013; Shen et al., 2014)
Cas9-HF	N497A, R661A, Q695A, Q926A	Rationally engineered	
eSpCas9(1.1)	K848A, K1003A, R1060A	Rationally engineered	(Slaymaker et al., 2016)
HypaCas9	N692A, M694A, Q695A, H698A	Rationally engineered	(Chen et al., 2017)
HiFi Cas9	R691A	Random mutational and high-throughput bacterial screening	(Vakulskas et al., 2018)
evoCas9	M495V, Y515N, K562E, R661Q	Random mutational and high-throughput bacterial screening	(Casini et al., 2018)
SniperCas9	A203D, H277N, G366S, M763I, F1038Y, T1102P, D1127E	Random mutational and high-throughput bacterial screening	(Lee et al., 2018)
LZ3-Cas9	n.a.	Rationally engineered	(Schmid-Burgk et al., 2020)

Cas9 nickases bear an alanine substitution at one of its nuclease residues in each domain. A D10A or H840A mutation will render inactivation of the HNH and RuvC domain, respectively. This mutation will result in the cleavage of one DNA strand, generating a nicked DNA product. In mammalian cells, this nicked DNA is faithfully repaired by the base excision repair pathway (Dianov & Hu, 2013). From the perspective of genome editing, the use of Cas9 nickases will reduce off-target indel formations, a hypothesis supported by the findings from several reports (Cong et al., 2013; Mali et al., 2013; Ran et al., 2013). However, the application of single-nicked DNA is limited for gene editing purposes due to the low editing efficiencies, as shown in previous reports (Cong et al., 2013; Mali et al., 2013; Ran et al., 2013).

Pairing two guided RNA Cas9 nickase, done to offset each other (Figure 5), mediates on-target editing while reducing off-target activity from individual nicking guides (Mali et al., 2013; Ran et al., 2013). Orientation and the length of the offset between the gRNA pairs greatly affect DSB formation. The PAM-out orientation was shown to produce higher indel formation, indicating the requirements of tail-to-tail Cas9 formation for proper strand separation of nicked DNA (Mali et al., 2013; Ran et al., 2013; Bothmer et al., 2017). With this orientation, a D10A pair would result in a 5’ overhang and H840A in a 3’ overhang DSBs. Experimental studies indicate that the optimal offset length to produce indel formation or microdeletion is up to 100 bp.

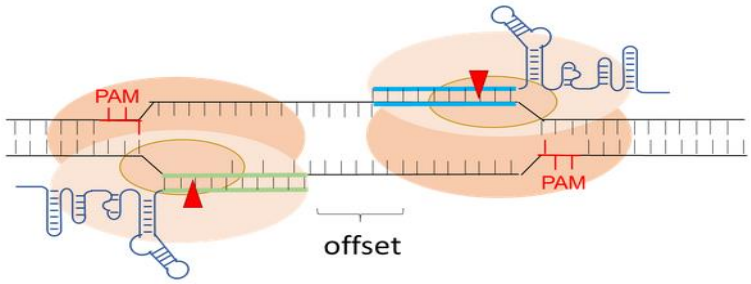

Figure 5. Paired Cas9 D10A nickases with a PAM-out orientation. Two Cas9 D10A nickase-guideRNA complexes are paired together to create nicks at opposing strands. Efficient editing with paired Cas9 nickases is achieved through a PAM out orientation. The offset is the space between the two ends of the guide RNA.

Advances in the characterization of the Cas9 protein have enabled scientists to rationally engineer Cas9 proteins that alter the binding properties to DNA. Improved specificity of Cas9 variants through this rationale was first reported in 2016. Here, two reports demonstrated that reducing Cas9 binding affinity towards guide RNA:DNA mismatch sequences would result in a reduction of DNA cleavage activity at off-target sites.

Slaymaker et al. (2016) generated espCas9(1.1), which has three mutations position at K848A, K1003A, and R1060A. These alanine amino acid changes neutralize positive residues on the non-target strand groove, weakening the interaction between the DNA sequence and the Cas9 complex (Figure 6). This would result in a Cas9 protein that is less tolerant to mismatched sequences. Analysis using break labeling, enrichment on streptavidin and next-generation sequencing (BLESS) demonstrated the reduction of mismatch and endogenous off-target activity of espCas9(1.1) against wild-type Cas9 (Slaymaker et al., 2016).

Conversely, Kleinstiver and colleagues hypothesized that disrupting the Cas9 contacts with the target DNA strand would require a more stringent complementary pairing of the guide RNA to the DNA target. Cas9-HF, which has four mutations at N497A, R661A, Q695A, and Q926A, was generated and assessed against wild-type Cas9 using genome-wide unbiased identification of double-stranded breaks enabled by sequencing (GUIDE-Seq). Sequence assessment on multiple endogenous sites revealed that even though on-target performance was slightly decreased, global off-target activity was significantly reduced (Kleinstiver et al., 2016).

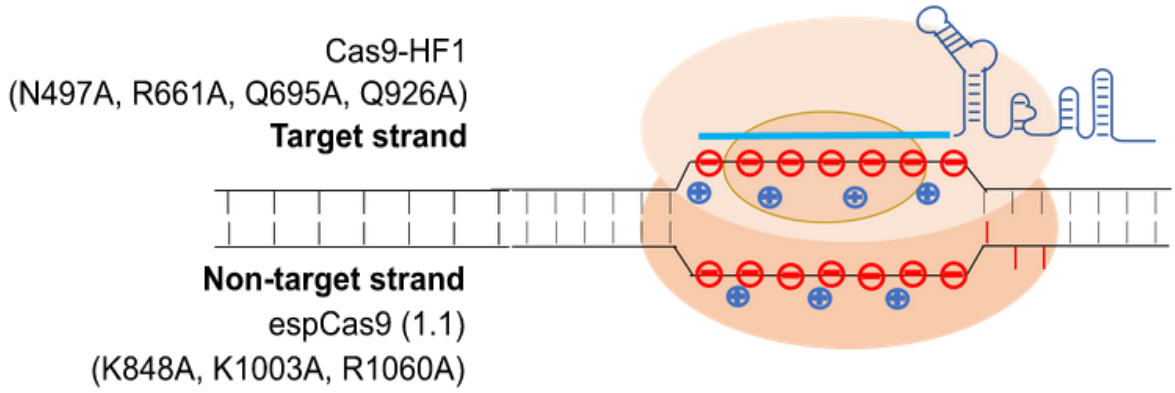

Figure 6. The enhanced specificity Cas9 with reduced binding affinity to mismatched target sequences. Two rationally engineered Cas9 variants introduce mutations at the DNA interacting domains to reduce the binding affinity of Cas9 to mismatch sequences. Cas9-HF1 introduces four mutations that affect the interaction between the Cas9 and the target strand, as the eSpCsa9(1.1) introduces three mutations that weaken the interaction between Cas9 and the non-target strand.

Since then, multiple engineered Cas9 variants have been developed, with each variant providing an increase in their targeting specificity or Cas9 activity (Table 3). However, it must be noted that each report employed a different assay system in evaluating their Cas9 variant performances.

Cas9-protein fusions for targeted molecular reactions

The first Cas9 fusion with other protein factors was first demonstrated in 2013. Gilbert and colleagues repurposed the CRISPR-Cas9 system into a transcription regulator by fusing chromatin modifier domains onto dCas9 (Gilbert et al., 2012). Thereafter, multiple proteins of interest have been fused to Cas9 protein for various purposes, such as genomic imaging and epigenetic modification applications (Hong et al., 2018; Stepper et al., 2017; Yeo et al., 2018).

Regarding genome editing, activator proteins fused to Cas9 nucleases are generally aimed to increase HDR-mediated repair (Table 4). This could be achieved through: (i) the formation of a drug-inducible system to control Cas9 activity, (ii) cell-cycle regulated expression of Cas9, and (iii) fusion of Cas9 with NHEJ inhibitors or HDR-promoting factors.

Drug-inducible systems have been widely used to fine-control the expression of a protein of interest. Multiple CRISPR-Cas9 inducible systems with various mechanisms of action have been developed, with one notable example being the ER-driven nuclear translocation system. The estrogen receptor (ER) is a transcription factor that is regulated by t estrogen. In its absence, ER binds to heat shock protein 90 (Hsp90), preventing its translocation to the nucleus. Advances in protein engineering have produced ER mutants that could effectively bind to a synthetic drug, 4-Hydroxytamoxifen (4-OHT) (Zhang et al., 2019). Cas9 fusion to this ER mutant, known as ERT2, would form an inducible Cas9 activity system that requires the presence of 4-OHT to translocate to the nucleus and initiate targeted genomic cleavage (Zhao et al., 2018).

An alternative way to control Cas9 activity is to fuse a destructive domain (DD) ligand to the Cas9. The DD is an engineered ligand domain that initiates the binding of the protein to the proteasome in the absence of a specific small molecule ligand (Zhang et al., 2019). Several ligand-DD pairs have been developed, in which CMP8/4-OHT-estrogen receptor destabilized domain (ER50 DD) and trimethoprim (TMP)-dihydrofolate reductase (DFHR) have been demonstrated to control Cas9 activity in cell cultures and mice models (Maji et al., 2017; Senturk et al., 2017).

Regulation of Cas9 activity dependent on the cell cycle is one other interesting approach. Conversion of Cas9 protein into a cell-cycle-dependent manner could be attained through the fusion of geminin (Gutschner et al., 2016; Howden et al., 2016). This system would not require any drug additions and would also limit Cas9 activity to the S-phase of the cell cycle. Geminin is a replication licensing factor that is ubiquitinylated by the Apc/Cdh1 E3 ligase complex during M and G1 cell cycle phase. Ubiquitinoylation of Geminin at the destruction box motif would lead to proteosome-dependent degradation (Ballabeni et al., 2013). Indeed, the fusion of this minimal geminin destruction box motif to Cas9 was shown to confine Cas9 activity to the S-phase of the cell cycle and increase HDR events in in vitro and in vivo studies (Gutschner et al., 2016; Howden et al., 2016).

Fusion of Cas9 with DNA repair regulator proteins was also developed with two notable Cas9 fusion variants, including the fusion of CtIP domain and DN18 (Charpentier et al., 2018; Jayavaradhan et al., 2019). A phosphorylated CtIP is a co-factor of the MRN complex, which is the initiator for DNA end resection. Fusion of a minimal CtIP domain to the Cas9, termed Cas9-HE, was demonstrated to stimulate HDR-mediated integrations (Charpentier et al., 2018). Moreover, the use of Cas9-HE was shown to induce higher MMEJ-associated and less NHEJ-associated indel formations at the target site, further supporting the concept of Cas9-HE increasing DNA end resection events.

A similar concept of increasing DNA end resection was demonstrated through the inhibition of 53BP1. The 53BP1 acts as a DNA repair regulator by inhibiting DNA end resection and promoting the NHEJ pathway (Jayavaradhan et al., 2019). Jayavaradhan and colleagues managed to suppress the accumulation of 53BP1 to the Cas9-mediated DSB site by fusing a dominant negative version of 53BP1 (DN1S) to the Cas9 protein. This local NHEJ inhibition would then reduce the formation of NHEJ-associated indels and increase HDR-transgene insertion, as shown in a proof-of-concept leukocyte editing of leukocyte adhesion deficiency (LAD) patient-derived B lymphocytes (Jayavaradhan et al., 2019).

Table 4. Cas9 fusion proteins for CRISPR-Cas9 genome editing.

Fusion protein	Effect	Result	Reference
DFHR-DD Cas9	Inducible ligand-DD stabilisation	Reduce off-target activity	(Maji et al., 2017)
FKBP-DD-Cas9	Inducible ligand-DD stabilisation	Reduced off-target activity	(Senturk et al., 2017)
Cas9-Geminin	Cell-cycle dependent Cas9 activity (induce Cas9 proteolysis in G1)	Modest increase in HR integration, lower indel formation, and reduce off-target activity	(Gutschner et al., 2016; Howden et al., 2016)
Cas9- ER^{T2} (Cas9-HIT)	Drug inducible Cas9 expression	Reduced off-target activity	(Zhao et al., 2018)
Cas9-CtIP (Cas9-HE)	Inducing HR by CtIP activation	Modest increase in HR integration and lower on-target indel formation	(Charpentier et al., 2018)
Cas9-DN1S	Reduces NHEJ events by blocking 53BP1 activation	Reduction of NHEJ indel formation and increase biallelic HR repair	(Jayavaradhan et al., 2019)

Base editing and prime editing

The introduction of DNA DSBs by CRISPR-Cas9 complexes may cause unwanted off-targeting events, which is a major limitation for faithful genome editing (van Overbeek et al., 2016; Brinkman et al., 2018). Advances in the field of Cas9 engineering have provided new strategies for precise editing without the generation of DNA DSBs. Base editing enables directed A-to-T or C-to-G base changes with the aid of base editors. Base editors are Cas9 nickases which have an adenine deaminase or cytosine deaminase fused to the C-terminus. Base changes are mediated by these deaminases, which will then trigger a mismatch repair pathway to establish the directed base change (Komor et al., 2016; Gaudelli et al., 2017).

Application of base editing includes any base alteration of DNA, which will result in permanent editing, as well as RNA, resulting in temporary genetic editing (Cox et al., 2017; Surun et al., 2020). Single nucleotide polymorphism or single point mutations that result in the pathogenesis of diseases, such as sickle cell anemia, are mostly the target diseases for this technology. Single base corrections of SNP or mutations are able to be done by applying this technology. Among these genetic diseases, the base editing strategy has been applied to reverse single base mutation of sickle cell disease and to lower blood cholesterol (Newby, 2021; Musunuru et al., 2021). These studies are currently being studied for their safety in human clinical trials.

Similar to base editing, prime editing allows precise genome editing without the generation of DNA DSBs. Prime editors consist of Cas9 nickases fused with an engineered reverse transcriptase (RT) domain and an extended guide RNA, termed prime editing guide RNA (pegRNA) (Anzalone et al., 2019). Prime editing can perform targeted small editing such as insertions, deletions, or sequence substitutions at the targeted site. The pegRNA provides the sequence template, which will then be integrated into the nicked genome by the RT domain. With this capacity, prime editing has the potential to be used to correct the majority of known genetic variants associated with diseases (Chen & Liu, 2023).

CONCLUSION: GENOME EDITING REVOLUTION

The discovery, adaptation, and proof-of-concept of the CRISPR-Cas9 system to specifically target and cleave DNA in eukaryote cells has marked a new era in genome editing. Previously, two promising site-specific genome editing tools have been developed, the zinc finger nucleases (ZFN) and transcription activator-like effector-nucleases (TALEN). Furthermore, the advancements in expanding the genomic toolbox, such as the use of Cas9 orthologs, has rationally engineered Cas9 to relax its PAM motives and increase its specificity. Moreover, the fusion of molecular regulatory proteins to Cas9 has had a substantial impact on the potential use of these technologies to solve and cure known genetic diseases

REFERENCES

Anders, C., Niewoehner, O., Duerst, A., & Jinek, M. (2014). Structural basis of pam-dependent target DNA recognition by the cas9 endonuclease. *Nature*, *513*(7519), 569–573. https://doi.org/10.1038/nature13579

Anzalone, A. V., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., Chen, P. J., Wilson, C., Newby, G. A., Raguram, A., & Liu, D. R. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*, *576*(576). https://doi.org/10.1038/s41586-019-1711-4

Arslan, Z., Hermanns, V., Wurm, R., Wagner, R., & Pul, Ü. (2014). Detection and characterization of spacer integration intermediates in type I-E CRISPR–CAS system. *Nucleic Acids Research*, *42*(12), 7884–7893. https://doi.org/10.1093/nar/gku510

Ballabeni, A. et al. (2013) 'Geminin deploys multiple mechanisms to regulate Cdt1 before cell division thus ensuring the proper execution of DNA replication', Proceedings of the National Academy of Sciences of the United States of America, 110(30), pp. 8–13. doi: 10.1073/pnas.1310677110

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 (New York, N.Y.)*, *315*(5819), 1709–1712. https://doi.org/10.1126/science.1138140

Bothmer, A., Phadke, T., Barrera, L. A., Margulies, C. M., Lee, C. S., Buquicchio, F., Moss, S., Abdulkerim, H. S., Selleck, W., Jayaram, H., Myer, V. E., & Cotta-Ramusino, C. (2017). Characterization of the interplay between DNA repair and CRISPR/Cas9-induced DNA lesions at an endogenous locus. *Nature Communications*, *8*(1). https://doi.org/10.1038/ncomms13905

Briner, Alexandra E., Donohoue, Paul D., Gomaa, Ahmed A., Selle, K., Slorach, Euan M., Nye, Christopher H., Haurwitz, Rachel E., Beisel, Chase L., May, Andrew P., & Barrangou, R. (2014). Guide RNA functional modules direct cas9 activity and orthogonality. *Molecular Cell*, *56*(2), 333–339. https://doi.org/10.1016/j.molcel.2014.09.019

Brinkman, E. K., Chen, T., de Haas, M., Holland, H. A., Akhtar, W., & van Steensel, B. (2018). Kinetics and fidelity of the repair of cas9-induced double-strand DNA breaks. *Molecular Cell*, *70*(5), 801-813.e6. https://doi.org/10.1016/j.molcel.2018.04.016

Casini, A., Olivieri, M., Petris, G., Montagna, C., Reginato, G., Maule, G., Lorenzin, F., Prandi, D., Romanel, A., Demichelis, F., Inga, A., & Cereseto, A. (2018). A highly specific SpCas9 variant is identified by in vivo screening in yeast. *Nature Biotechnology*, *36*(3), 265–271. https://doi.org/10.1038/nbt.4066

Charpentier, M., Khedher, A. H. Y., Menoret, S., Brion, A., Lamribet, K., Dardillac, E., Boix, C., Perrouault, L., Tesson, L., Geny, S., De Cian, A., Itier, J. M., Anegon, I., Lopez, B., Giovannangeli, C., & Concordet, J. P. (2018). CtIP fusion to Cas9 enhances transgene integration by homology-dependent repair. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-03475-7

Chatterjee, P., Jakimo, N., & Jacobson, J. M. (2018). Minimal PAM specificity of a highly similar SpCas9 ortholog. *Science Advances*, *4*(10), eaau0766. https://doi.org/10.1126/sciadv.aau0766

Chen, B., Gilbert, Luke A., Cimini, Beth A., Schnitzbauer, J., Zhang, W., Li, G.-W., Park, J., Blackburn, Elizabeth H., Weissman, Jonathan S., Qi, Lei S., & Huang, B. (2013). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell*, *155*(7), 1479–1491. https://doi.org/10.1016/j.cell.2013.12.001

Chen, J. S., Dagdas, Y. S., Kleinstiver, B. P., Welch, M. M., Sousa, A. A., Harrington, L. B., Sternberg, S. H., Joung, J. K., Yildiz, A., & Doudna, J. A. (2017). Enhanced proofreading governs CRISPR–Cas9 targeting accuracy. *Nature*, *550*(7676), 407–410. https://doi.org/10.1038/nature24268

Chen, P. J., & Liu, D. R. (2022). Prime editing for precise and highly versatile genome manipulation. *Nature Reviews Genetics*, *24*, 1–17. https://doi.org/10.1038/s41576-022-00541-1

Chylinski, K., Le Rhun, A., & Charpentier, E. (2013). The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA Biology*, *10*(5), 726–737. https://doi.org/10.4161/rna.24321

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. https://doi.org/10.1126/science.1231143

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*, *358*(6366), 1019–1027. https://doi.org/10.1126/science.aaq0180

Dagdas, Y. S., Chen, J. S., Sternberg, S. H., Doudna, J. A., & Yildiz, A. (2017). A conformational checkpoint between DNA binding and cleavage by CRISPR-Cas9. *Science Advances*, *3*(8). https://doi.org/10.1126/sciadv.aao0027

Datsenko, K. A., Pougach, K., Tikhonov, A., Wanner, B. L., Severinov, K., & Semenova, E. (2012). Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nature Communications*, *3*(1). https://doi.org/10.1038/ncomms1937

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., Eckert, M. R., Vogel, J., & Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, *471*(7340), 602–607. https://doi.org/10.1038/nature09886

Deveau, H., Barrangou, R., Garneau, J. E., Labonte, J., Fremaux, C., Boyaval, P., Romero, D. A., Horvath, P., & Moineau, S. (2007). Phage response to CRISPR-Encoded resistance in Streptococcus thermophilus. *Journal of Bacteriology*, *190*(4), 1390–1400. https://doi.org/10.1128/jb.01412-07

Dianov, G. L., & Hübscher, U. (2013). Mammalian base excision repair: The forgotten archangel. *Nucleic Acids Research*, *41*(6), 3483–3490. https://doi.org/10.1093/nar/gkt076

Fu, Y. et al. (2013) 'High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells', Nature Biotechnology. Nature Publishing Group, 31(9), pp. 822–826. doi: 10.1038/nbt.2623.

Garneau, J. E., Dupuis, M.-È., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A. H., & Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, *468*(7320), 67–71. https://doi.org/10.1038/nature09523

Gasiunas, G., Barrangou, R., Horvath, P., & Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences*, *109*(39), E2579–E2586. https://doi.org/10.1073/pnas.1208507109

Gaudelli, N. M., Komor, A. C., Rees, H. A., Packer, M. S., Badran, A. H., Bryson, D. I., & Liu, D. R. (2017). Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature*, *551*(7681), 464–471. https://doi.org/10.1038/nature24644

Gilbert, Luke A., Larson, Matthew H., Morsut, L., Liu, Z., Brar, Gloria A., Torres, Sandra E., Stern-Ginossar, N., Brandman, O., Whitehead, Evan H., Doudna, Jennifer A., Lim, Wendell A., Weissman, Jonathan S., & Qi, Lei S. (2013). CRISPR-Mediated modular rna-guided regulation of transcription in eukaryotes. *Cell*, *154*(2), 442–451. https://doi.org/10.1016/j.cell.2013.06.044

Gisler, S., Gonçalves, J. P., Akhtar, W., de Jong, J., Pindyurin, A. V., Wessels, L. F. A., & van Lohuizen, M. (2019). Multiplexed Cas9 targeting reveals genomic location effects and gRNA-based staggered breaks influencing mutation efficiency. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-09551-w

González Castro, N., Bjelic, J., Malhotra, G., Huang, C., & Alsaffar, S. H. (2021). Comparison of the feasibility, efficiency, and safety of genome editing technologies. *International Journal of Molecular Sciences*, *22*(19), 10355. https://doi.org/10.3390/ijms221910355

Gopalappa, R., Suresh, B., Ramakrishna, S., & Kim, H. (2018). Paired D10A Cas9 nickases are sometimes more efficient than individual nucleases for gene disruption. *Nucleic Acids Research*, *46*(12), e71–e71. https://doi.org/10.1093/nar/gky222

Gutschner, T., Haemmerle, M., Genovese, G., Draetta, Giulio F., & Chin, L. (2016). Post-translational regulation of Cas9 during G1 enhances homology-directed repair. *Cell Reports*, *14*(6), 1555–1566. https://doi.org/10.1016/j.celrep.2016.01.019

Heler, R., Samai, P., Modell, J. W., Weiner, C., Goldberg, G. W., Bikard, D., & Marraffini, L. A. (2015). Cas9 specifies functional viral targets during CRISPR–Cas adaptation. *Nature*, *519*(7542), 199–202. https://doi.org/10.1038/nature14245

Hong, Y., Lu, G., Duan, J., Liu, W., & Zhang, Y. (2018). Comparison and optimization of CRISPR/dCas9/gRNA genome-labeling systems for live cell imaging. *Genome Biology*, *19*(1). https://doi.org/10.1186/s13059-018-1413-5

Hou, Z., Zhang, Y., Propson, N. E., Howden, S. E., Chu, L.-F. ., Sontheimer, E. J., & Thomson, J. A. (2013). Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. *Proceedings of the National Academy of Sciences*, *110*(39), 15644–15649. https://doi.org/10.1073/pnas.1313587110

Howden, Sara E., McColl, B., Glaser, A., Vadolas, J., Petrou, S., Little, Melissa H., Elefanty, Andrew G., & Stanley, Edouard G. (2016). A cas9 variant for efficient generation of indel-free knockin or gene-corrected human pluripotent stem cells. *Stem Cell Reports*, *7*(3), 508–517. https://doi.org/10.1016/j.stemcr.2016.07.001

Hu, J. H., Miller, S. M., Geurts, M. H., Tang, W., Chen, L., Sun, N., Zeina, C. M., Gao, X., Rees, H. A., Lin, Z., & Liu, D. R. (2018). Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature*, *556*(7699), 57–63. https://doi.org/10.1038/nature26155

Huang, J., Zhou, Y., Li, J., Lu, A., & Liang, C. (2022). CRISPR/Cas systems: Delivery and application in gene therapy. *Frontiers in Bioengineering and Biotechnology*, *10*. https://doi.org/10.3389/fbioe.2022.942325

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., & Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *Journal of Bacteriology*, *169*(12), 5429–5433. https://doi.org/10.1128/jb.169.12.5429-5433.1987

Iyer, S., Suresh, S., Guo, D.-S., Daman, K., Chen, J. K., Liu, P., Zieger, M., Luk, K., Roscoe, B. P., Mueller, C., King, O. D., Emerson, C. P., & Wolfe, S. M. (2019). Precise therapeutic gene correction by a simple nuclease-induced double-stranded break. *Nature*, *568*(7753), 561–565. https://doi.org/10.1038/s41586-019-1076-8

Jansen, Ruud., Embden, Jan. D. A. van, Gaastra, Wim., & Schouls, Leo. M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology*, *43*(6), 1565–1575. https://doi.org/10.1046/j.1365-2958.2002.02839.x

Jayavaradhan, R., Pillis, D. M., Goodman, M., Zhang, F., Zhang, Y., Andreassen, P. R., & Malik, P. (2019). CRISPR-Cas9 fusion to dominant-negative 53BP1 enhances HDR and inhibits NHEJ specifically at Cas9 target sites. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-10735-7

Jiang, F., Taylor, D. W., Chen, J. S., Kornfeld, J. E., Zhou, K., Thompson, A. J., Nogales, E., & Doudna, J. A. (2016). Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science*, *351*(6275), 867–871. https://doi.org/10.1126/science.aad8282

Jiang, F., Zhou, K., Ma, L., Gressel, S., & Doudna, J. A. (2015). A Cas9-guide RNA complex preorganized for target DNA recognition. *Science*, *348*(6242), 1477–1481. https://doi.org/10.1126/science.aab1452

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. https://doi.org/10.1126/science.1225829

Jinek, M., Jiang, F., Taylor, D. W., Sternberg, S. H., Kaya, E., Ma, E., Anders, C., Hauer, M., Zhou, K., Lin, S., Kaplan, M., Iavarone, A. T., Charpentier, E., Nogales, E., & Doudna, J. A. (2014). Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science*, *343*(6176), 1247997–1247997. https://doi.org/10.1126/science.1247997

Karvelis, T., Gasiunas, G., Miksys, A., Barrangou, R., Horvath, P., & Siksnys, V. (2013). CrRNA and tracrRNA guide Cas9-mediated DNA interference inStreptococcus thermophilus. *RNA Biology*, *10*(5), 841–851. https://doi.org/10.4161/rna.24203

Kim, D., Kim, J., Hur, J. K., Been, K. W., Yoon, S., & Kim, J.-S. (2016). Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nature Biotechnology*, *34*(8), 863–868. https://doi.org/10.1038/nbt.3609

Kim, N., Kim, H. K., Lee, S., Seo, J. H., Choi, J. W., Park, J., Min, S., Yoon, S., Cho, S.-R., & Kim, H. H. (2020). Prediction of the sequence-specific cleavage activity of Cas9 variants. *Nature Biotechnology*, 1–9. https://doi.org/10.1038/s41587-020-0537-9

Kim, S., Kim, D., Cho, S. W., Kim, J., & Kim, J.-S. . (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Research*, *24*(6), 1012–1019. https://doi.org/10.1101/gr.171322.113

Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z., & Joung, J. K. (2016). High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature*, *529*(7587), 490–495. https://doi.org/10.1038/nature16526

Kleinstiver, B. P., Prew, M. S., Tsai, S. Q., Topkar, V. V., Nguyen, N. T., Zheng, Z., Gonzales, A. P. W., Li, Z., Peterson, R. T., Yeh, J.-R. J., Aryee, M. J., & Joung, J. K. (2015). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature*, *523*(7561), 481–485. https://doi.org/10.1038/nature14592

Knight, S. C., Xie, L., Deng, W., Guglielmi, B., Witkowsky, L. B., Bosanac, L., Zhang, E. T., El Beheiry, M., Masson, J.-B., Dahan, M., Liu, Z., Doudna, J. A., & Tjian, R. (2015). Dynamics of CRISPR-Cas9 genome interrogation in living cells. *Science*, *350*(6262), 823–826. https://doi.org/10.1126/science.aac6572

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. https://doi.org/10.1038/nature17946

Koonin, E. V., Makarova, K. S., & Zhang, F. (2017). Diversity, classification and evolution of CRISPR-Cas systems. *Current Opinion in Microbiology*, *37*, 67–78. https://doi.org/10.1016/j.mib.2017.05.008

Künne, T., Swarts, D. C., & Brouns, S. J. J. (2014). Planting the seed: Target recognition of short guide RNAs. *Trends in Microbiology*, *22*(2), 74–83. https://doi.org/10.1016/j.tim.2013.12.003

Kuscu, C., Arslan, S., Singh, R., Thorpe, J., & Adli, M. (2014). Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nature Biotechnology*, *32*(7), 677–683. https://doi.org/10.1038/nbt.2916

Lee, J. K., Jeong, E., Lee, J., Jung, M., Shin, E., Kim, Y., Lee, K., Jung, I., Kim, D., Kim, S., & Kim, J.-S. (2018). Directed evolution of CRISPR-Cas9 to increase its specificity. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-05477-x

Li, T., Yang, Y., Qi, H., Cui, W., Zhang, L., Fu, X., He, X., Liu, M., Li, P., & Yu, T. (2023). CRISPR/Cas9 therapeutics: Progress and prospects. *Signal Transduction and Targeted Therapy*, *8*(1), 1–23. https://doi.org/10.1038/s41392-023-01309-7

Lin, Y., Jun, L., Li, C., Tu, Z., Li, S., Shengbo Eben Li, & Yan, S. (2022). Application of CRISPR/Cas9 system in establishing large animal models. *Frontiers in Cell and Developmental Biology*, *10*. https://doi.org/10.3389/fcell.2022.919155

Ma, H., Tu, L.-C., Naseri, A., Huisman, M., Zhang, S., Grunwald, D., & Pederson, T. (2016). CRISPR-Cas9 nuclear dynamics and target recognition in living cells. *Journal of Cell Biology*, *214*(5), 529–537. https://doi.org/10.1083/jcb.201604115

Maji, B., Moore, C. L., Zetsche, B., Volz, S. E., Zhang, F., Shoulders, M. D., & Choudhary, A. (2016). Multidimensional chemical control of CRISPR–Cas9. *Nature Chemical Biology*, *13*(1), 9–11. https://doi.org/10.1038/nchembio.2224

Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., Barrangou, R., Brouns, S. J. J., Charpentier, E., Haft, D. H., Horvath, P., Moineau, S., Mojica, F. J. M., Terns, R. M., Terns, M. P., White, M. F., Yakunin, A. F., Garrett, R. A., van der Oost, J., & Backofen, R. (2015). An updated evolutionary classification of CRISPR–Cas systems. *Nature Reviews Microbiology*, *13*(11), 722–736. https://doi.org/10.1038/nrmicro3569

Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M., Kosuri, S., Yang, L., & Church, G. M. (2013). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature Biotechnology*, *31*(9), 833–838. https://doi.org/10.1038/nbt.2675

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E., & Church, G. M. (2013). RNA-Guided human genome engineering via Cas9. *Science*, *339*(6121), 823–826. https://doi.org/10.1126/science.1232033

Marraffini, L. A., & Sontheimer, E. J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science*, *322*(5909), 1843–1845. https://doi.org/10.1126/science.1165771

McGinn, J., & Marraffini, L. A. (2018). Molecular mechanisms of CRISPR–Cas spacer acquisition. *Nature Reviews Microbiology*, *17*(1), 7–12. https://doi.org/10.1038/s41579-018-0071-7

Miller, S. M., Wang, T., Randolph, P. B., Arbab, M., Shen, M. W., Huang, T. P., Matuszek, Z., Newby, G. A., Rees, H. A., & Liu, D. R. (2020). Continuous evolution of SpCas9 variants compatible with non-G PAMs. *Nature Biotechnology*, *38*(4), 471–481. https://doi.org/10.1038/s41587-020-0412-8

Mojica, F. J. M., Dı, C. and Garcı, J. (2009) 'Short motif sequences determine the targets of the prokaryotic CRISPR defence system', Microbiology, pp. 733–740. doi: 10.1099/mic.0.023960-0.

Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J., & Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of Molecular Evolution*, *60*(2), 174–182. https://doi.org/10.1007/s00239-004-0046-3

Mojica, F. J. M., Diez-Villasenor, C., Soria, E., & Juez, G. (2000). Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Molecular Microbiology*, *36*(1), 244–246. https://doi.org/10.1046/j.1365-2958.2000.01838.x

Musunuru, K., et al. (2021). In vivo CRISPR base editing of PCSK9 durably lowers cholesterol in primates. Nature 593, 429–434. https://doi.org/10.1038/s41586-021-03534-y

Newby, G. A., Yen, J. S., Woodard, K. J., Mayuranathan, T., Lazzarotto, C. R., Li, Y., Sheppard-Tillman, H., Porter, S. N., Yao, Y., Mayberry, K., Everette, K. A., Jang, Y., Podracky, C. J., Thaman, E., Lechauve, C., Sharma, A., Henderson, J. M., Richter, M. F., Zhao, K. T., & Miller, S. M. (2021). Base editing of haematopoietic stem cells rescues sickle cell disease in mice. *Nature*, *595*(7866), 295–302. https://doi.org/10.1038/s41586-021-03609-w

Nishimasu, H., Ran, F. Ann, Hsu, Patrick D., Konermann, S., Shehata, Soraya I., Dohmae, N., Ishitani, R., Zhang, F., & Nureki, O. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell*, *156*(5), 935–949. https://doi.org/10.1016/j.cell.2014.02.001

Nishimasu, H., Shi, X., Ishiguro, S., Gao, L., Hirano, S., Okazaki, S., Noda, T., Abudayyeh, O. O., Gootenberg, J. S., Mori, H., Oura, S., Holmes, B., Tanaka, M., Seki, M., Hirano, H., Aburatani, H., Ishitani, R., Ikawa, M., Yachie, N., & Zhang, F. (2018). Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science*, *361*(6408), 1259–1262. https://doi.org/10.1126/science.aas9129

Pattanayak, V., Lin, S., Guilinger, J. P., Ma, E., Doudna, J. A., & Liu, D. R. (2013). High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nature Biotechnology*, *31*(9), 839–843. https://doi.org/10.1038/nbt.2673

Pourcel, C., Salvignol, G., & Vergnaud, G. (2005). CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology*, *151*(3), 653–663. https://doi.org/10.1099/mic.0.27437-0

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. https://doi.org/10.1038/nature14299

Rees, H. A., & Liu, D. R. (2018). Base editing: Precision chemistry on the genome and transcriptome of living cells. *Nature Reviews Genetics*, *19*(12), 770–788. https://doi.org/10.1038/s41576-018-0059-1

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

Sapranauskas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., & Siksnys, V. (2011). The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. *Nucleic Acids Research*, *39*(21), 9275–9282. https://doi.org/10.1093/nar/gkr606

Schmid-Burgk, J. L., Gao, L., Li, D., Gardner, Z., Strecker, J., Lash, B., & Zhang, F. (2020). Highly parallel profiling of Cas9 variant specificity. *Molecular Cell*, *78*(4), 794-800.e8. https://doi.org/10.1016/j.molcel.2020.02.023

Senturk, S., Shirole, N. H., Nowak, D. G., Corbo, V., Pal, D., Vaughan, A., Tuveson, D. A., Trotman, L. C., Kinney, J. B., & Sordella, R. (2017). Rapid and tunable method to temporally control gene editing based on conditional Cas9 stabilization. *Nature Communications*, *8*(1). https://doi.org/10.1038/ncomms14370

Shen, M. W., Arbab, M., Hsu, J. Y., Worstell, D., Culbertson, S. J., Krabbe, O., Cassa, C. A., Liu, D. R., Gifford, D. K., & Sherwood, R. I. (2018). Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature*, *563*(7733), 646–651. https://doi.org/10.1038/s41586-018-0686-x

Sheridan, C. (2023). The world's first CRISPR therapy is approved: Who will receive it? *Nature Biotechnology*, *42*(1). https://www.nature.com/articles/d41587-023-00016-6

Slaymaker, I. M., Gao, L., Zetsche, B., Scott, D. A., Yan, W. X., & Zhang, F. (2015). Rationally engineered Cas9 nucleases with improved specificity. *Science*, *351*(6268), 84–88. https://doi.org/10.1126/science.aad5227

Stepper, P., Kungulovski, G., Jurkowska, R. Z., Chandra, T., Krueger, F., Reinhardt, R., Reik, W., Jeltsch, A., & Jurkowski, T. P. (2016). Efficient targeted DNA methylation with chimeric dCas9–Dnmt3a–Dnmt3L methyltransferase. *Nucleic Acids Research*, *45*(4), 1703–1713. https://doi.org/10.1093/nar/gkw1112

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. https://doi.org/10.1038/nature13011

Sürün, D., Schneider, A., Mircetic, J., Neumann, K., Lansing, F., Paszkowski-Rogacz, M., Hänchen, V., Lee-Kirsch, M. A., & Buchholz, F. (2020). Efficient generation and correction of mutations in human iPS cells utilizing mRNAs of CRISPR base editors and prime editors. *Genes*, *11*(5), 511. https://doi.org/10.3390/genes11050511

Szczelkun, M. D., Tikhomirova, M. S., Sinkunas, T., Gasiunas, G., Karvelis, T., Pschera, P., Siksnys, V., & Seidel, R. (2014). Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proceedings of the National Academy of Sciences*, *111*(27), 9798–9803. https://doi.org/10.1073/pnas.1402597111

Vakulskas, C. A., Dever, D. P., Rettig, G. R., Turk, R., Jacobi, A. M., Collingwood, M. A., Bode, N. M., McNeill, M. S., Yan, S., Camarena, J., Lee, C. M., Park, S. H., Wiebking, V., Bak, R. O., Gomez-Ospina, N., Pavel-Dinu, M., Sun, W., Bao, G., Porteus, M. H., & Behlke, M. A. (2018). A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nature Medicine*, *24*(8), 1216–1224. https://doi.org/10.1038/s41591-018-0137-0

van Overbeek, M., Capurso, D., Carter, Matthew M., Thompson, Matthew S., Frias, E., Russ, C., Reece-Hoyes, John S., Nye, C., Gradia, S., Vidal, B., Zheng, J., Hoffman, Gregory R., Fuller, Christopher K., & May, Andrew P. (2016). DNA repair profiling reveals nonrandom outcomes at Cas9-mediated breaks. *Molecular Cell*, *63*(4), 633–646. https://doi.org/10.1016/j.molcel.2016.06.037

Walton, R. T., Christie, K. A., Whittaker, M. N., & Kleinstiver, B. P. (2020). Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science*, *368*(6488). https://doi.org/10.1126/science.aba8853

Wei, Y., Terns, R. M., & Terns, M. P. (2015). Cas9 function and host genome sampling in type II-A CRISPR–Cas adaptation. *Genes & Development*, *29*(4), 356–361. https://doi.org/10.1101/gad.257550.114

Wilkinson, M., Drabavicius, G., Silanskas, A., Gasiunas, G., Siksnys, V., & Wigley, D. B. (2019). Structure of the dna-bound spacer capture complex of a type II CRISPR-Cas system. *Molecular Cell*, *75*(1), 90-101.e5. https://doi.org/10.1016/j.molcel.2019.04.020

Wright, A. V., & Doudna, J. A. (2016). Protecting genome integrity during CRISPR immune adaptation. *Nature Structural & Molecular Biology*, *23*(10), 876–883. https://doi.org/10.1038/nsmb.3289

Wright, Addison V., Nuñez, James K., & Doudna, Jennifer A. (2016). Biology and applications of CRISPR systems: Harnessing nature's toolbox for genome engineering. *Cell*, *164*(1-2), 29–44. https://doi.org/10.1016/j.cell.2015.12.035

Wu, X., Scott, D. A., Kriz, A. J., Chiu, A. C., Hsu, P. D., Dadon, D. B., Cheng, A. W., Trevino, A. E., Konermann, S., Chen, S., Jaenisch, R., Zhang, F., & Sharp, P. A. (2014). Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nature Biotechnology*, *32*(7), 670–676. https://doi.org/10.1038/nbt.2889

Xiao, Y., Luo, M., Hayes, R. P., Kim, J., Ng, S., Ding, F., Liao, M., & Ke, A. (2017). Structure basis for directional R-loop formation and substrate handover mechanisms in type I CRISPR-Cas system. *Cell*, *170*(1), 48-60.e11. https://doi.org/10.1016/j.cell.2017.06.012

Xiao, Y., Ng, S., Nam, K. H., & Ke, A. (2017). How type II CRISPR–Cas establish immunity through Cas1–Cas2-mediated spacer integration. *Nature*, *550*(7674), 137–141. https://doi.org/10.1038/nature24020

Yamada, M., Watanabe, Y., Gootenberg, J. S., Hirano, H., F. Ann Ran, Takanori Nakane, Ryuichiro Ishitani, Zhang, F., Hiroshi Nishimasu, & Osamu Nureki. (2017). Crystal structure of the minimal Cas9 from campylobacter jejuni reveals the molecular diversity in the CRISPR-Cas9 systems. *Molecular Cell*, *65*(6), 1109-1121.e3. https://doi.org/10.1016/j.molcel.2017.02.007

Yao, R., Liu, D., Jia, X., Zheng, Y., Liu, W., & Xiao, Y. (2018). CRISPR-Cas9/Cas12a biotechnology and application in bacteria. *Synthetic and Systems Biotechnology*, *3*(3), 135–149. https://doi.org/10.1016/j.synbio.2018.09.004

Yeo, N. C., Chavez, A., Lance-Byrne, A., Chan, Y., Menn, D., Milanova, D., Kuo, C.-C., Guo, X., Sharma, S., Tung, A., Cecchi, R. J., Tuttle, M., Pradhan, S., Lim, E. T., Davidsohn, N., Ebrahimkhani, M. R., Collins, J. J., Lewis, N. E., Kiani, S., & Church, G. M. (2018). An enhanced CRISPR repressor for targeted mammalian gene regulation. *Nature Methods*, *15*(8), 611–616. https://doi.org/10.1038/s41592-018-0048-5

Zhang, J., Chen, L., Zhang, J., & Wang, Y. (2019). Drug inducible CRISPR/Cas systems. *Computational and Structural Biotechnology Journal*, *17*, 1171–1177. https://doi.org/10.1016/j.csbj.2019.07.015

Zhao, C., Zhao, Y., Zhang, J., Lu, J., Chen, L., Zhang, Y., Ying, Y., Xu, J., Wei, S., & Wang, Y. (2018). HIT-Cas9: A CRISPR/Cas9 genome-editing device under tight and effective drug control. *Molecular Therapy - Nucleic Acids*, *13*, 208–219. https://doi.org/10.1016/j.omtn.2018.08.022

Zheng, T., Hou, Y., Zhang, P., Zhang, Z., Xu, Y., Zhang, L., Niu, L., Yang, Y., Liang, D., Yi, F., Peng, W., Feng, W., Yang, Y., Chen, J., Zhu, Y. Y., Zhang, L.-H., & Du, Q. (2017). Profiling single-guide RNA specificity reveals a mismatch sensitive core sequence. *Scientific Reports*, *7*(1). https://doi.org/10.1038/srep40638

Zhu, H., Li, C., & Gao, C. (2020). Applications of CRISPR–Cas in agriculture and plant biotechnology. *Nature Reviews Molecular Cell Biology*, *21*(11), 661–677. https://doi.org/10.1038/s41580-020-00288-9

Zuo, Z., & Liu, J. (2016). Cas9-catalyzed DNA cleavage generates staggered ends: Evidence from molecular dynamics simulations. *Scientific Reports*, *6*(1). https://doi.org/10.1038/srep37584