Molecular Biology Reports (2021) 48:4879–4895
https://doi.org/10.1007/s11033-021-06479-7

REVIEW

CRISPR/Cas based gene editing: marking a new era in medical science

Kirti Prasad[1,3] · Anila George[1,2] · Nithin Sam Ravi[1,2] · Kumarasamypet M. Mohankumar[1]

Received: 22 February 2021 / Accepted: 5 June 2021 / Published online: 18 June 2021

Abstract

CRISPR/Cas9 system, a bacterial adaptive immune system developed into a genome editing technology, has emerged as a powerful tool revolutionising genome engineering in all branches of biological science including agriculture, research and medicine. Rapid evolution of CRISPR/Cas9 system from the generation of double strand breaks to more advanced applications on gene regulation has made the wide-spread use of this technology possible. Medical science has benefited greatly from CRISPR/Cas9; being both a versatile and economical tool, it has brought gene therapy closer to reality. In this review, the development of CRISPR/Cas9 system, variants thereof and its application in different walks of medical science- research, diagnostics and therapy, will be discussed.

Keywords Programmable nucleases · CRISPR/Cas9 · Gene therapy · Gene editing

Abbreviations

CRISPR	Clustered regularly interspaced short palindromic repeats
Cas	CRISPR associated proteins
DNA	Deoxyribo nucleic acid
HDR	Homology directed repair
NHEJ	Non homologous end joining
ZFN	Zinc finger nucleases
TALEN	Transcription activator like effector nucleases
RNA	Ribo nucleic acid
crRNA	Clustered regularly interspaced short palindromic repeats ribo nucleic acid
tracrRNA	Trans- activating clustered regularly interspaced short palindromic repeats ribo nucleic acid
PAM	Protospacer adjacent motifs
InDel	Insertions and deletions
sgRNA	Single guide RNA
CTD	C terminal domain
RNP	Ribo nucleo protein
GOI	Gene of interest
NLS	Nuclear localisation signal
AAV	Adeno associated virus
mRNA	Messenger Ribo nucleic acid
spCas	Streptococcus pyogenes CRISPR associated proteins
ssDNA	Single stranded deoxyribo nucleic acid
ssoDNA	Single stranded oligo deoxyribo nucleic acid
saCAS	Staphylococcus aureus CRISPR associated proteins
APOBEC1	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide 1
UGI	Uracil glycosylase inhibitor
dCAS	Dead CRISPR associated proteins
ABE	Adenine base editor
CBE	Cytosine base editor
CGBE	C-to-G base editor
PEG RNA	Prime editing guide RNA
ACE	Adenine and cytosine base editor
CRISPRi	Clustered regularly interspaced short palindromic repeats interference
CRISPRa	Clustered regularly interspaced short palindromic repeats activation
KRAB	Krüppel associated box
enCHIP	Engineered DNA-binding molecule-mediated chromatin immune precipitation

✉ Kumarasamypet M. Mohankumar
Mohankumarkm@cmcvellore.ac.in

[1] Centre for Stem Cell Research (a Unit of inStem, Bengaluru), Christian Medical College Campus, Bagayam, Vellore 632002, Tamil Nadu, India

[2] Sree Chitra Tirunal Institute for Medical Sciences and Technology, Thiruvananthapuram 695 011, Kerala, India

[3] Manipal Academy of Higher Education, Manipal 576104, Karnataka, India

CAPTURE	CRISPR affinity purification in situ of regulatory elements
CRISPR GO	CRISPR-genome organization
CLOuD9	Chromatin loop re-organization using CRISPR-dCas9
HSC	Hematopoietic stem cell
CCR5	C-C chemokine receptor type 5
CAR	Chimeric antigen receptor
IPSC	Induced pluripotent stem cells
BCL11A	B-cell lymphoma 11A
SHERLOCK	Specific high-sensitivity enzymatic reporter unlocking
DETECTER	DNA endonuclease targeted CRISPR Trans reporter
COVID	Corona virus disease

Introduction

Precision targeting and creation of double strand break for gene manipulation are the two most vital requirements in genome editing. Early research on double strand break, repair and recombination were conducted in the 1980s which paved the way for the development of gene editing tools [1, 2]. Maria Jasin in 1994 used Meganucleases to create double strand breaks in DNA and showed that it can be corrected by cellular repair machinery either using error free Homology directed repair (HDR) or error prone Non homologous end joining (NHEJ) pathways, revealing the scope for gene editing [3]. Although the targeting scope of Meganucleases were very limited, this discovery furthered the research for programmable nucleases that can be targeted to precise genomic loci. Two important gene editing tools thus developed in the 2000s were Zinc finger nucleases (ZFNs) and Transcription activator like effector nucleases (TALENs). In both the tools DNA binding proteins (Zinc fingers in ZFNs and TAL effectors in TALENs) are fused with an endonuclease (FokI) to obtain precisely targetable nucleases [4–7]. Once the cut is made in the DNA, the resulting double strand breaks will be corrected by cellular break repair pathways forming small insertions or deletions causing DNA editing. The major limitation of the protein based nucleases were the difficulty in synthesizing new DNA binding proteins each time to target a new location and hence scientists were on the lookout for easily programmable nucleases [8].

The research into CRISPR Cas system had begun long back. Francisco Mojica et al. noticed that the bacterial genome had certain sequences which repeated several times with regular spaces in between them [9]. Further research identified these as fragments of DNA from bacteriophages which attack the bacteria, but how and why these fragments were integrated into the bacterial genome remained unknown [10]. It was in 2012 that Jennifer Doudna and Emmanuelle Charpentier elucidated the mechanism of CRISPR/Cas system and developed it into a genome editing tool [11]. From the discovery of CRISPR/Cas system in bacteria it took almost 20 years to understand the mechanism and develop it into a gene editing tool; but the following decade witnessed a rapid surge in research, improvement and application of this system highlighting its potential in both research and clinical applications (Fig. 1). In this review, we will be discussing about the CRISPR/Cas system and its development as a gene editing tool, potential applications in genome editing and the current limitations.

CRISPR in bacterial system

CRISPR/Cas system is the RNA guided adaptive immune system in bacteria analogous to the adaptive immune system in humans. The bacteria which escapes the primary attack by bacteriophages or other mobile genetic elements stores memories of the invasion in the form of short DNA fragments in its chromosome. Upon reinvasion by the same phage, bacteria uses this stored information for silencing the invaders [12].

The clustered regularly interspaced short palindromic repeats–CRISPR-associated protein (CRISPR/Cas) system as the name suggests consists of two parts- the nucleic acid and the protein components. On escaping a viral attack, small fragments (~ 20 bp) of viral DNA are cut and integrated into specific genomic loci in bacteria, known as the CRISPR array. This locus also codes for the components of Cas (CRISPR associated) protein which is an endonuclease. The integrated DNA fragments in the CRISPR array are transcribed and processed subsequently to produce short pre crRNAs, with a unique spacer sequence at the 5' end and a conserved repeat sequence at the 3' end. Pre crRNA forms a RNA duplex with a long non-coding RNA termed tracrRNA forming mature crRNA and recruits the Cas protein to form a Ribonucleoprotein (RNP) complex which will be surveilling the bacterial cell and upon attack can specifically target and cleave the invader nucleic acid having sequence similarity to the spacer sequence. The target nucleic acid is distinguished from the host DNA by the presence of a unique short Protospacer Adjacent Motif (PAM) sequence, adjacent to the spacer sequence which is present only in the invader genome [13–15]. Different species of bacteria possess unique Cas proteins which differ in structure, PAM requirements and functionality [16]. This review will be confined to discussions on SpCas 9, which was identified in Streptococcus *pyogenes* and is the most commonly used type of Cas protein (Fig. 2a).

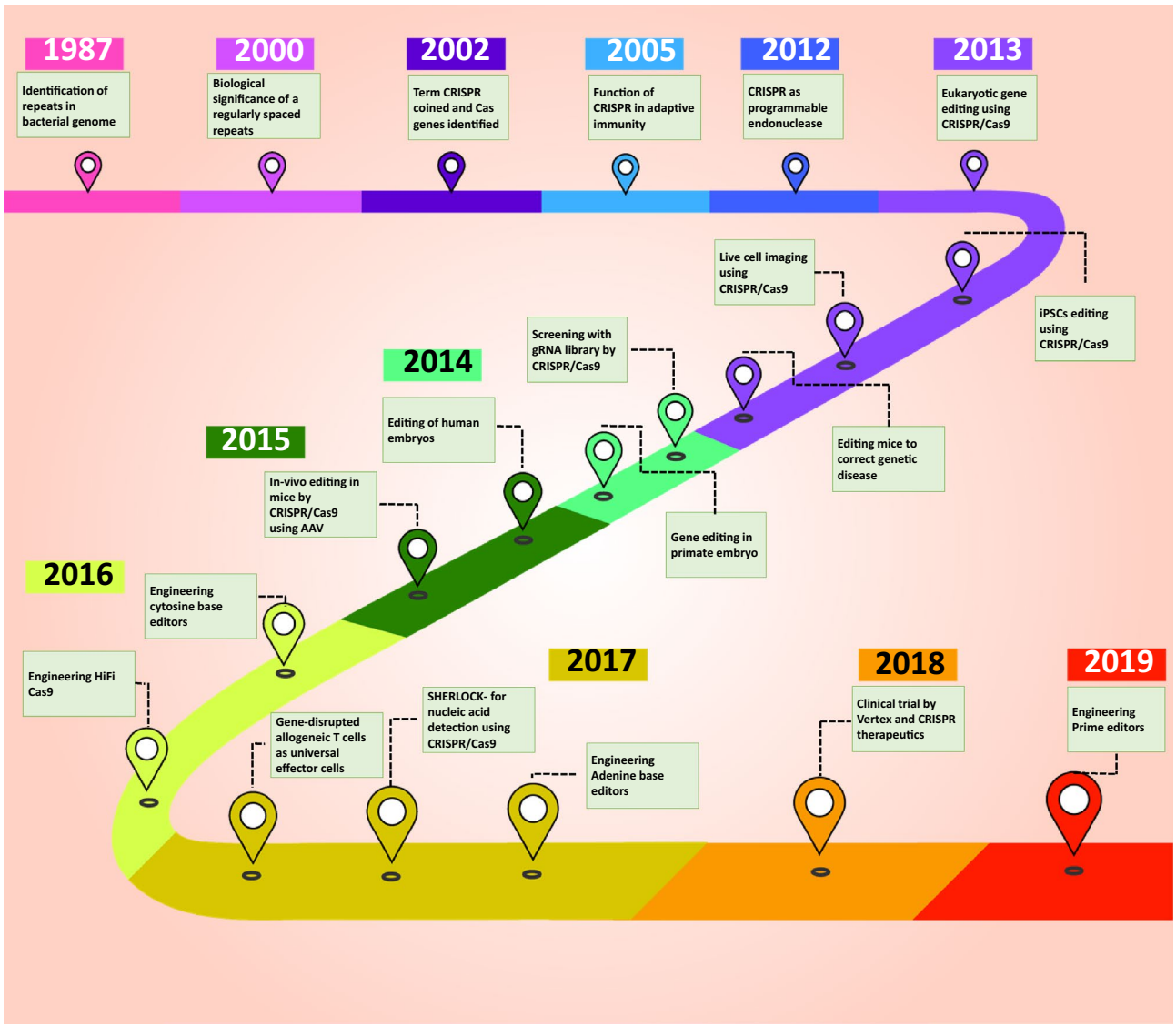


Fig. 1 The Brief history on evolution of CRISPR/Cas system for genome editing: Discovery of CRISPR/Cas system for genome editing (1987–2012); Various approaches for eukaryotic genome editing (2013–2017); Evolution of CRISPR/Cas9 based tools (2016–2019); First clinical trial based on CRISPR/Cas9 system (2018) are represented in the roadmap

Development of CRISPR as a gene editing tool

The early research and development of gene editing tools had established that the prerequisite for any genome editing tool is the ability to specifically bind and introduce double strand breaks in the target region, which will further be corrected by the cellular repair machinery causing small insertions and deletions (InDels). The CRISPR/Cas system fulfilled both the requirements: specific target recognition by the spacer sequence and double strand break by the Cas9 nuclease. Target recognition by CRISPR is mediated by the 20 nucleotides spacer sequence in the crRNA and not by the protein as in the forerunner targetable endonucleases such as TALENs and ZFNs [17]. Designing and synthesising varying nucleic acid sequences binding to target regions is easier than designing proteins for the same; thus CRISPR/Cas9 provided a simple, economical and versatile tool which can be made to target any desired sequence by altering only the ~ 20 bp spacer sequences (Fig. 2b).

The engineered CRISPR Cas system for gene editing consist of two essential components: The Cas protein which is the nuclease and a single guide RNA (sgRNA) consisting of fused crRNA and tracrRNA, which recruits Cas9 protein and recognizes the target site with appropriate PAM. Once the double strand break is made, the DNA damage is repaired by either NHEJ or HDR pathways in the cell which

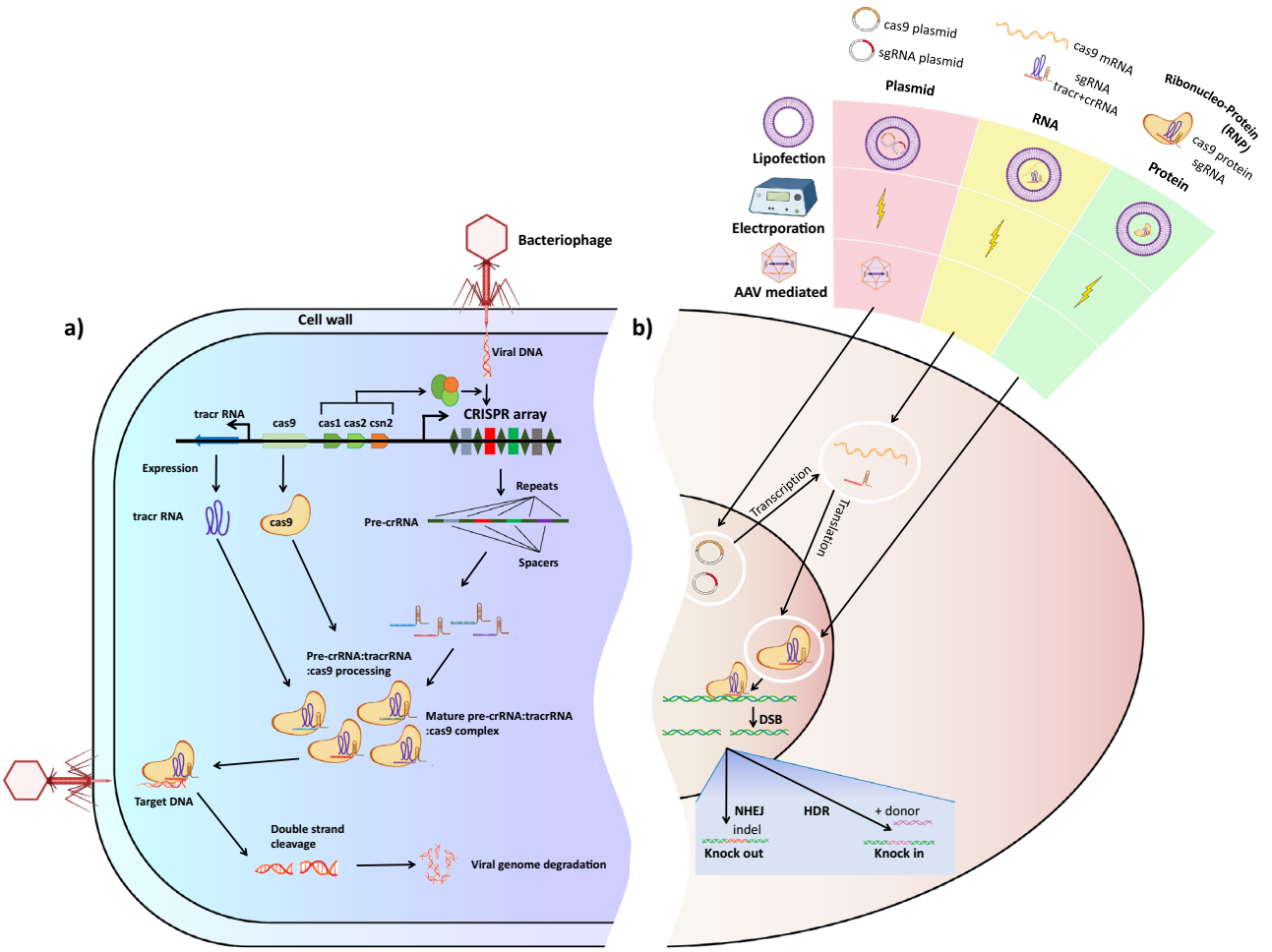


Fig. 2 Schematic representation of CRISPR-Cas9 activity in prokaryotes and eukaryotes **a** Adaptive immunity in bacteria by CRISPR/Cas9; After bacteriophage infection, cas1, cas2 and csn2 forms complex with viral DNA fragments and integrates into the host DNA as CRISPR array. For CRISPR/Cas9 system to be active against the invading bacteriophages, tracer RNA, pre-crRNA and Cas9 are individually expressed and mature RNA–protein complex is formed. Once the complex is formed the sgRNA (tracr + crRNA) helps guiding the Cas9 protein to the target foreign DNA molecule for double strand cleavage, thus silencing the invading genetic material. **b** CRISPR-Cas9 delivery strategies in eukaryotes; Different formats of delivering Cas9 in the form of plasmid, mRNA and protein along with sgRNA via lipofection, electroporation and viral mediated delivery is represented in the figure

leads to small insertions or deletions at the target site causing genome editing [18].

Components of the engineered CRISPR/Cas9 system

Cas9 protein

SpCas9 is a large, multidomain, single turnover endonuclease that cuts the target DNA 3-5 bp upstream of the PAM sequence. It has a distinct bilobed structure: a recognition (REC) lobe and a nuclease (NUC) lobe, which are connected by linker sequences. The REC lobe of Cas9 includes the bridge helix motif and REC1, REC2 and REC3 domains [19, 20]. The NUC lobe consists of two distinct nuclease domains, RuvC and HNH, along with a c-terminal domain (CTD) consisting of PAM interacting sites. The REC lobe and the NUC lobe of Cas9 fold to present a positively charged groove at their interface which accommodates the negatively charged sgRNA: target DNA heteroduplex. The nuclease domains are highly conserved whereas the PAM interacting domain is quite variable among the different Cas proteins. The RuvC domain is split in the primary structure and comes together after folding to form the nuclease domain that cleaves the non-target strand through

a two-metal ion catalytic mechanism. The HNH domain on the other hand uses a one metal ion catalytic system to cleave the target strand [21]. The PAM interacting site in c-terminal domain of NUC lobe is responsible for PAM interrogation and is kept in an inactive conformation as long as the sgRNA is not bound to the protein thereby preventing unwanted nucleic acid binding of the protein.

sgRNA

The sgRNA (Single guide RNA) for gene editing has been engineered by fusing crRNA and tracrRNA. The 20 nucleotides user defined spacer is at the 5' end followed by the repeat sequence of crRNA which forms an RNA duplex with the anti-repeat sequence of the tracrRNA. The tracrRNA has a unique stem-loop structure and the repeat-anti-repeat RNA duplex along with the stem-loop1 is required for the interaction with Cas9 protein while stem-loop 2 and 3 at the 3' end of tracrRNA provides stability to the RNP complex (Fig. 3). The target specificity of sgRNA is provided by the 10–12 nucleotides at the 3' end of the spacer sequence known as the seed sequence which is essential for R-loop formation and any mismatch in this region leads to loss of specificity of the sgRNA [21, 22].

Mechanism of CRISPR based gene editing

Target recognition, unwinding of dsDNA and cutting occurs in a sequential and coordinated manner by forming ribonucleoprotein (RNP) complexes which consist of Cas9 and sgRNA (Fig. 3).

sgRNA-Cas interaction

RNP complex formation is the first and most essential step in CRISPR mediated gene editing. The Cas9 protein forms hydrogen bonds with the stem loop structure of tracrRNA. It has been shown that sgRNA competes with other cellular RNAs for binding to Cas9 protein and the stem loops 2 and 3 increases the binding efficiency of sgRNA. The interaction of sgRNA with Cas9 activates the latter for PAM searching and recognition by repositioning the HNH and RuvC domains to a DNA binding conformation [22, 23]. This sequential interaction of sgRNA with Cas9 prevents the unnecessary binding of Cas9 to DNA targets without activation.

PAM recognition

Presence of PAM sequence is essential for target recognition and cleavage by the CRISPR/Cas system. spCas9 requires a 5'NGG3' PAM sequence on the non-target strand for its activity. The Cas9-sgRNA complex starts the search for target sequence by first searching for the presence of 3' PAM sequence. Cas9 stays associated for a longer time with the DNA containing PAM sequence which facilitates the unwinding of adjacent DNA sequence and formation of DNA-sgRNA duplex. The PAM sequence in the non-target strand interacts and forms hydrogen bond with crucial arginine residues in the PAM interacting domain of Cas9; at the same time upstream (+ 1 of PAM) phosphate group in the target strand is stabilized by critical lysine and serine residues of C-terminal domain which creates a kink in the strand. These interactions facilitate the local DNA melting and RNA–DNA hybridization [24].

RNA – DNA duplex (R – loop) formation

After PAM recognition the sgRNA initiates complementary base pairing with the target DNA in an unidirectional manner starting at the PAM proximal nucleotide. The base pairing will occur only if there is sufficient homology between the target region and the spacer sequence and mismatches in the target strand can lead to displacement of RNP complex from the DNA. Once separated, the target strand forms a RNA–DNA hybrid with the spacer region of sgRNA and will be placed in a channel between the two lobes of Cas9, while the non-target strand will be positioned within a tunnel in the NUC lobe. Similar to the target strand, the non-target strand also kinks at + 1 position with an additional kink at + 4 position. The kinks make both strands susceptible to cleavage by the nucleases [25].

Target cleavage

RNA–DNA duplex formation activates Cas9 to make a blunt double stranded cut in the target DNA 3–5 bp upstream of the PAM sequence. The HNH domain cleaves the target strand while the RuvC domain cleaves the non-target strand by hydrolysing the phosphodiester bonds in the respective strands in a metal ion dependent manner. The completion of complementary base pairing at the PAM distal end induces conformational activation of HNH domain which in turn allosterically regulates the positioning and activity of RuvC domain via the two linkers thus affecting a concerted double stranded cut. Once the cut is made the enzyme remains bound to the DNA until certain cellular factors displace it [26].

DNA repair and editing

The double strand breaks induced by the CRISPR/Cas system will be corrected by Non Homologous End Joining (NHEJ) or homology directed repair (HDR). NHEJ creates small random insertions or deletions (indels) in the target

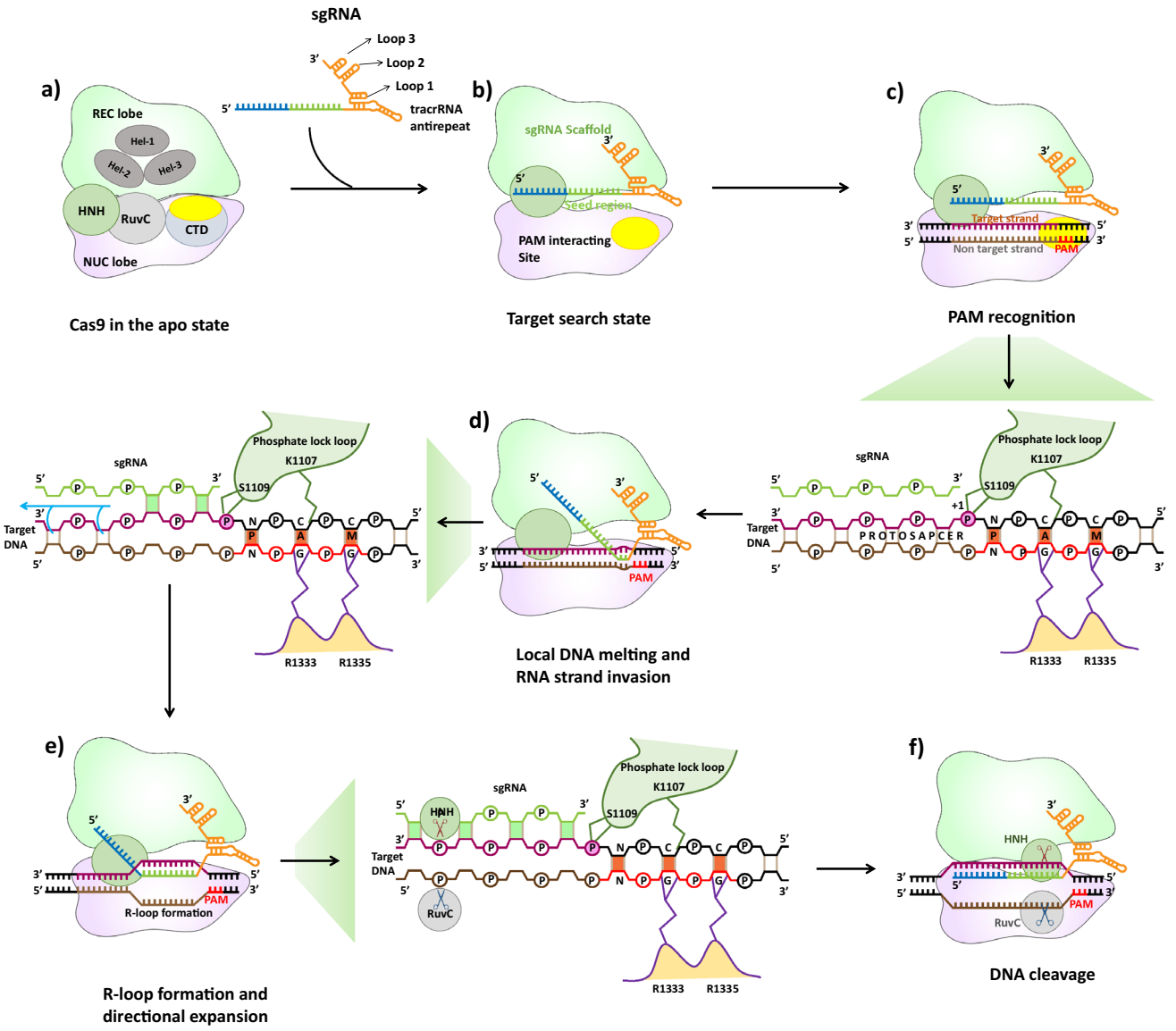


Fig. 3 Schematic representation of Cas9 mediated target recognition and cleavage: **a** Structure of Cas9 in inactive state; sgRNA consist of crRNA and tracrRNA connected by an artificial tetra loop; the 20 nucleotide spacer region is represented in blue and green colour where the green color represents the seed region. **b** Target Search; Upon sgRNA binding the REC lobe of Cas9 undergoes a conformational change so as to position the 5' end of sgRNA inside a cavity formed between the two-nuclease domains (HNH &RuVC) thus preventing its degradation. **c** PAM Recognition; the major and minor grooves of the conserved dinucleotide guanine bases in the PAM(NGG) forms hydrogen bond with critical residues (arginine (R1333 and R1335) and lysine (K1107) respectively) of the CTD. This facilitates the interaction of S1109 in the phosphate lock loop with phosphate at position + 1 of the PAM via hydrogen bonding, thus stabilizing the DNA in such a way that the first base of the target sequence rotates upwards towards the sgRNA **d** Local DNA melting and RNA strand invasion; PAM recognition leads to melting of DNA adjacent to the PAM site and invasion of RNA strand to the unwound region. **e** R-loop formation and directional expansion; sufficient complementarity will lead to the unwinding of DNA and flipping of the target strand to form an RNA–DNA duplex with the sgRNA while the non target strand will remain free. **f** Cleavage; Upon R-loop formation each nuclease domain of Cas9 cleaves the respective strands of the target DNA in between 3rd and 5th bp from the PAM sequence and produces double strand break

site creating a variety of mutations. The resulting amino acid change can lead to a permanent knock-out of the gene of interest although the efficiency may vary from site to site.

HDR on the other hand relies on a donor template, which carries a gene of interest (GOI) or smaller mutation to be inserted at the double strand break site with right and left homology arms, which can be delivered along with Cas9

and sgRNA. Although less efficient and more cumbersome than NHEJ mediated repair, HDR has the advantage of giving seamless editing. The homologous DNA strand can be delivered as single stranded oligonucleotide (ssODN) or as double stranded plasmid DNA through transfection or transduction. HDR can be utilized for targeted knocking-in of the desired gene, epitope tagging of genes etc. There have been considerable efforts at increasing HDR efficiency by using small molecules which stalls the cell division phase at S and G2-M phase, inhibiting enzymes involved in NHEJ pathway, making staggered cuts and fusing HDR repair proteins or donor template to Cas9 [5, 27].

The development of CRISPR/Cas9 system was a major boost to scientific research as well as gene therapy application. Ease of generating knockout led to the development of various cellular, plant and animal models of diseases. It also made possible to create screening libraries targeting various regions, revealing the functional roles of various regulatory elements in the genome. Unlike the preceding programmable nucleases, CRISPR/Cas9 made genome editing of desired genes easier.

CRISPR based genome editing in mammalian cells

One of the major hurdles in using CRISPR/Cas9 for editing eukaryotic system was the presence of nucleus which separates the DNA from cytoplasm, which was not in case of prokaryotic system and the challenge of repurposing a prokaryotic system for eukaryotic application was addressed by various groups. A system which was designed to act on nucleic acids in the cytoplasm now had to enter the nucleus for editing. George M. Church's and Feng Zhang's group in 2013 synthesised human codon optimised SpCas9 with C-terminal SV40 nuclear localization signal (NLS) and cloned into a mammalian expression system [28, 29]. They performed experiment in human cell lines and delivered both sgRNA (crRNA-tracrRNA fusion transcripts) and Cas9 in the form of plasmids via nucleofection for suspension culture and lipofection for adherent culture. Adding a NLS facilitated the efficient transfer of Cas9 to the nucleus, while codon optimization of the Cas9 sequence facilitated efficient protein translation in the eukaryotic system [29].

However, plasmid mediated delivery of Cas9 pose problems like decreased editing efficiency due to slow onset of translation, cell toxicity and lack of efficient methods to deliver them in-vitro [30]. To overcome these issues ribonucleoprotein (RNP) complex was introduced, a system in which the Cas9 in the form of protein and sgRNA in the form of RNA are complexed together and then delivered into the cell by liposome mediated or nucleofection mediated delivery [31, 32]. The delivery of Cas9 as RNP complex gives higher editing efficiency since it is delivered in functional form and it is one of the widely used strategies for genome editing in research and therapy. It also has less toxicity and the delivery is fairly easy when compared to plasmid DNA and mRNA. Though it can be used effectively for in-vitro genome editing this system cannot be used for *in-vivo* genome editing. For *in-vivo* genome editing a new strategy like AAV based cas9 delivery system was developed [33, 34], which acts as a stand-alone system without being coalesced with an additional delivery system. Though it is one of the best strategies available for *in-vivo* genome editing AAV delivery system has few drawbacks like capsid induced immune response, difficulty in achieving empty capsid free virus particles and high production cost. To address these problems, chemically modified Cas9 mRNA was developed and is delivered using lipid nano particles primed towards the tissue of interest. These modified mRNA are more stable, cost effective, and do not elicit an immune response when compared to the viral mediated delivery system [35] (Fig. 2b).

A large repository of computational tools and databases have been developed which can aid in all steps of genome engineering with higher specificity. Software's and web based tools are now available for efficient sgRNA designing (CHOP-CHOP, CRISPOR), analysis of editing outcomes (ICE, TIDE,EDITR), prediction and detection of off- target effects and has been extensively reviewed by Sledzinski et al. [36].

Variants of CRISPR/Cas

As the gene editing applications with SpCas-9 system gained momentum the search for variants of Cas proteins from other bacterial and archaeal species continued and there has been a steady increase in the number of variants reported with altered PAM requirements, nature of cutting, target nucleic acid (DNA/RNA), size etc. The most recent classification of the Cas proteins describes two classes with 6 types and 33 subtypes. The Cas variants are broadly classified into two based on the effector protein; class I consists of multiple proteins while Class II consists of a single effector protein [36]. Although many variants have been reported only a handful has potential gene editing applications. A few notable ones include Cas-12 which makes a staggered double strand cut, Cas-13 which can target RNA, Cas-14 which targets ssDNA, SaCas-9 which has a smaller size etc. [37–39]. Advances in protein engineering and evolution also facilitated the development of Cas variants with desired properties.

Cas9 variants based on PAM requirement

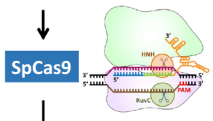
PAM flanking the target sequence is one of the essential requirements of Cas mediated gene editing. Longer PAM sequence increases the specificity while shorter PAM increases the targeting scope and hence there is considerable amount of research going on to increase the range of targetable PAM sequences. By exploring the natural diversity of Cas proteins a number of orthologs have been identified with unique PAM requirements although only a few have been reported for use in human cells (Table 1).

With the advancements in protein engineering, a variety of Cas proteins have been developed either by targeted mutagenesis or by directed evolution to have altered PAM requirements and even near PAM-less targeting (Table 1). Engineering of PAM requirement often relies on logically altering the critical amino acids in the PAM interacting site, thus affecting the binding specificities of the protein. This has considerably

Table 1 CRISPR-Cas9 derived tools for genome editing

Common Cas9 Orthologs

Cas9 Orthologs	Organism Isolated from	PAM (5′ to 3′)	Target
SpCas9	*Streptococcus pyogenes*	NGG	DNA
SaCas9	*Staphylococcus aureus*	(NNGRRT or NNGRRN)	DNA
NmeCas9	*Neisseria meningitidis*	(NNNNGATT)	RNA
StCas9	*Streptococcus thermophilus*	NNAGAAW	DNA
FnCas9	*Francisella novicida*	NG	DNA
ScCas9	*Streptococcus canis*	NNG	DNA
NcCas9	*Neisseria cinerea*	NNNNGTA	DNA



Based on nuclease domain activity

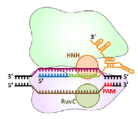
Modified Cas9		Nuclease domain	Genome editing tool	Target
dCas9		Mutated RuvC & mutated HNH	Epigenetic modifiers, Base editing	DNA
nCas9	D10A	Mutated RuvC domain	Base editing	DNA
	H840A	Mutated HNH domain	Prime editing	DNA

Based on improved efficiency and specificity

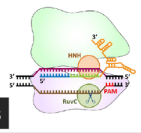
Engineered Cas9	Species	PAM	Target
eSpcas9	*Staphylococcus aureus*	NGG	DNA
SpCas9-HF1	*Staphylococcus aureus*	NGG	DNA
HypaCas9	*Staphylococcus aureus*	NGG	DNA
evoCas9	*Staphylococcus aureus*	NGG	DNA
HiFi Cas9	*Staphylococcus aureus*	NGG	DNA
Sniper Cas9	*Staphylococcus aureus*	NGG	DNA

Based on PAM variability

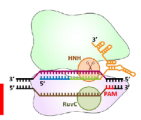
Cas9	Description	PAM (5′ to 3′)
VER Spcas9	*Staphylococcus aureus*	NGCG
VQR Spcas9	*Staphylococcus aureus*	NGAN OR NGNG
EQR Spcas9	*Staphylococcus aureus*	NGAG
QQR1 SpyCas9	*Staphylococcus aureus*	NAAG
VRER SpyCas9	*Staphylococcus aureus*	NGCG
xCas9	*Staphylococcus aureus*	NG, GAA, GAT
SpCas9-NG	*Staphylococcus aureus*	NG
SpRY	*Staphylococcus aureus*	NRN, NYN



S.No.	dCas9	Fusion Proteins	Function	Target
1.	Transcriptional activators	Direct fusion: VP64, VP48, VP120, p65, p65-HSF1, VP64-p65-Rta	VP64: strong transcriptional activation domain that recruits a variety of transcription factors and chromatin remodelling factors	DNA
2.	Transcriptional repressors	KRAB, SID4x, Mxi1, DNA Methyltransferase MQ1	KRAB: It associates with KP1, which forms a scaffold to recruit several proteins involved in inducing and spreading of heterochromatin over large distances	DNA
3.	Histone acetylation	p300 acetyltransferase, CBP (CREB-binding protein)	Increase levels of H3K27 ac histone modification	DNA
4.	DNA demethylation	Tet1 demethylase, TET1	Decrease methylation	DNA
5.	Histone demethylation	LSD1 (Lysine specific Demethylase 1)	Demethylates histone 3 at lysine residues 4 and 9, leading to silencing of enhancers	DNA
6.	DNA methylation	DNMT3A (DNA Methyltransferase 3 alpha)	Methylated CpGs in promoters are associated with silenced transcription	DNA



nCas9	Fusion Proteins	Function	Target
Prime editor Mutated HNH domain (H840A)	Murine leukemia virus (M-MLV) reverse transcriptase (RT)	Targeted small insertions, deletions and base swapping in precise way	DNA



nCas9	Fusion Proteins	Function	Target
Adenine Base Editor (ABE) Mutated RuvC domain (D10A)	tRNA-specific adenosine deaminase (TadA)	Converts Adenine to guanine	DNA
Cytosine Base Editor (CBE) Mutated RuvC domain (D10A)	Cytidine deaminase, uracil DNA glycosylase inhibitor (UGI)	Converts cytosine to thymine	DNA

enhanced the targeting scope of Cas protein with previously inaccessible regions now available for targeting.

Cas9 variants with improved efficiency and specificity

The efficiency and specificity of gene editing tools are very critical especially when it comes to gene therapy applications. Efficiency and specificity often bear an inverse relation with each other, higher efficiency leading to lesser specificity and vice versa.

There have been two main approaches in increasing the targeting specificity of Cas mediated editing. The first approach is to reduce the exposure time of DNA to Cas by delivering as RNP complex or as light or drug inducible Cas which can be spatially controlled. The other involves protein engineering which has produced Cas variants such as eSpCas9 (ala substitution that weaken the interaction b/w RuvC and non-target strand), spCas9hf1 (disrupted interaction b/w cas9 and phosphate backbone of DNA), hypaCas9 (increased cas9 proofreading activity) with improved specificity. Latest addition to this group was Sniper-Cas9 with an improved sensitivity and target specificity compared to the existing variants [40–46] (Table 1).

Cas9 variants with modified nucleases

dCas9

The ability of Cas-gRNA system to bind to specific sequences prompted the scientists to think about engineering it in such a way that the protein can bind to but not cleave the nucleic acid. This was achieved by inactivating the nuclease activity of Cas9 by creating two point mutations in the HNH and RuvC domains of Cas9. The H840A mutation in HNH and D10A mutation in RuvC created dCas9 which lost its nuclease activity while retaining the full potential to bind to the specific nucleic acid when guided by sgRNA (Table 1).

Nickases

Cas9 nickases were created by mutating only one of the two nucleases thus having the potential for cutting a single strand rather than causing a double strand break. The D10A mutant with inactive RuvC nuclease cleaves only the target strand while H840A mutant cleaves the non-target strand as HNH nuclease is mutated [47, 48] (Table 1).

Cas9 derived genome editing tools

The possibility to fuse Cas9 with other proteins opened up a variety of potential applications for the CRISPR/Cas9 system beyond the creation of double strand breaks at the target site.

dCas derived tools

dCas9 being able to bind to a specific region without causing strand breaks served as an excellent vehicle to take any proteins to the gene of interest. The modulator proteins can either be fused directly Cas protein or to the sgRNA carrying a suitable aptamer in the stem loop structure. It was now possible to do CRISPR mediated activation and repression of target genes, epigenetic modifications, study of chromatin interactions and live cell imaging with different dCas9 fusion proteins [49–51] (Table 1).

Nickase derived tools

Base editors

The possibility to fuse other proteins with Cas9 prompted the idea of base editors which can convert a single nucleotide to another without creating a double strand breaks. The first base editor was developed by David Liu's group in 2016. They fused APOBEC1, an enzyme in the cytosine deaminase family, to Cas9 nickase thus creating a cytosine base editor which can convert C.G to T.A. Cas9 will position the cytosine deaminase in correct orientation on the target DNA strand allowing it to convert cytosine base to uracil by deamination. The U-G mismatch will be corrected to U-A which in turn will be converted to T-A thus resulting in a C to T conversion at the target position. A uracil glycosylase inhibitor (UGI) fused to Cas9 prevents the conversion of U-G back to C-G by base excision repair. The use of nickase rather than dCas9 makes a single strand cut in the non-edited strand aiding in the preferential conversion of U-A to U-G during the repair thus increasing the efficiency. At about the same time another base editor termed Target AID was created by fusing activation induced cytosine deaminase to dCas9 [52, 53].

As there were no naturally occurring adenosine deaminases which can convert A to G, bacterial enzyme TadA which converts A to G in RNA was evolved to use DNA as substrate and create the desired edit. Multiple rounds of evolution created ABE which can convert adenosine to inosine which will subsequently be converted to guanosine. ABEs showed better efficiency and less off target effects compared to CBEs [54].

Development of Base Editors expanded the application of CRISPR/Cas9 system. It is now possible to create mutations without causing any double strand break or significant amount of Indels, thus making it possible to edit the coding regions without the risk of knock-out. Overtime several groups worked on improving the efficiency of both base editors. Codon optimisation, addition of NLS and other protein engineering techniques created a vast set of base editors with increased editing efficiency, altered PAM

requirement, modified editing window, reduced off target effects and smaller size (Table 1) [55–57]. Recently scientists also developed CGBE which is able to induce C to G transversion, thus expanding the number of genetic diseases that can be corrected by using base editor [58].

Adenosine and Cytosine Base editor (ACE), with the ability to simultaneously convert adenosine and cytosine bases within the window region, was created by fusing TadA from ABE and PmCDA1 from target AID to Cas9 nickase (D10A). Although less efficient than ABE and CBE, it offers the possibility for co-editing thus increasing the codon conversion potential in the target region [59].

EvolvR

Fusion of an error prone, nick translating DNA polymerase to the Cas9 nickase created this tool which can be utilized for targeted mutagenesis in the region of interest. The sgRNA guides PolI3M-nCas9 complex to a particular region of interest and creates a nick in the ssDNA via nCas9 endonuclease and dissociates from PolI3M. PolI3M is then recruited to the nick site and introduces errors by polymerizing new DNA strands and displacing the old DNA strand. EvolvR can be used to engineer microbes with different phenotypes, to study protein-protein and protein-DNA interactions, to investigate the functional role of DNA segments as well as for cellular barcoding [60].

Prime editor

Inability to create transversion mutations as well as non-specific editing within the window region of base editors (bystander effect) fuelled the research for systems that can create specific edits with single base precision. In 2019, Liu's group developed prime editing which can mediate single base conversion, insertions or deletions at the target site with reduced PAM constraints. They fused a reverse transcriptase enzyme to Cas9 nickase which can transcribe the sequence contained in the RNA template to the end of nicked DNA. The RNA template with the desired edit is given as a modified sgRNA termed pegRNA, which contains the spacer sequence at the 5' end followed by the RNA scaffold sequence and the RNA template at the 3' end. More flexibility compared to base editors and higher efficiency compared to HDR makes prime editing a very promising gene editing tool [61].

Applications of CRISPR/Cas9 system

The applications of CRISPR/Cas9 system will be described under three broad categories: research, therapeutics and diagnostics (Fig. 4).

Research applications

CRISPR/Cas9 system can be used for genetic studies as well as generation of cellular and animal models.

Gene repression, activation by epigenetic modifiers

Repression of transcription by dCas9 termed CRISPRi was first demonstrated in bacterial cells; it utilises the property of dCas9 to bind to a specific segment of DNA so that it can prevent the transcription elongation by blocking RNA polymerase movement or by blocking the attachment of specific transcription factors in the promoter region of target gene. The efficiency of CRISPRi was low in mammalian cells initially but has been enhanced by binding transcription repressors such as KRAB or SID4X to dCas9. This technique is specific, reversible and multiplexible although there is a possibility of regulating other non-targeted genes within the same operon [44, 62–64].

Similar to CRISPRi, transcription activation (CRISPRa) is also achieved by fusing transcription activation effectors such as p65 or VP64 to dCas9. These proteins are targeted to the promoter regions of genes to enhance the transcription of the gene without causing any mutations. CRISPR SAM is a protein complex engineered for activation of endogenous genes and can be used in conjunction with sgRNA libraries. It consist of a nucleolytically inactive Cas9-VP64 fusion, an sgRNA incorporating two MS2 RNA aptamers and MS2-P65-HSF1 activation helper protein and can be used for transcriptional activation of coding regions. The fusion of modifier enzymes such as histone acetyl transferase, methylcytosine dioxygenase, DNA methyl transferase to dCas9 can cause epigenetic modifications such as acetylation and methylation of histones and DNA [65–68].

Genome wide functional screens

The possibility for genome wide screening was perhaps one of the important advantages that CRISPR/Cas9 system had over its forerunners. Using a library of gRNAs, it is possible to conduct genome wide screening for establishing genotype-phenotype correlation. Numerous studies have been published where loss of function screens have been done using Cas9 system especially in cancer genetics. Use of dCas9 on the other hand allows both gain of function and loss of function studies [69–71].

Pooled sgRNA libraries, where hundreds to millions of different sgRNAs are cloned in the same backbone plasmid and mixed together in a single tube, has increased the efficiency of CRISPR based genome wide screens. CRISPRKo library (e.g., GeCKO, Avana and Brunello) contain multiple sgRNAs to create indel at target sites across the genome and makes it non-functional allele. Similarly, CRISPR repression

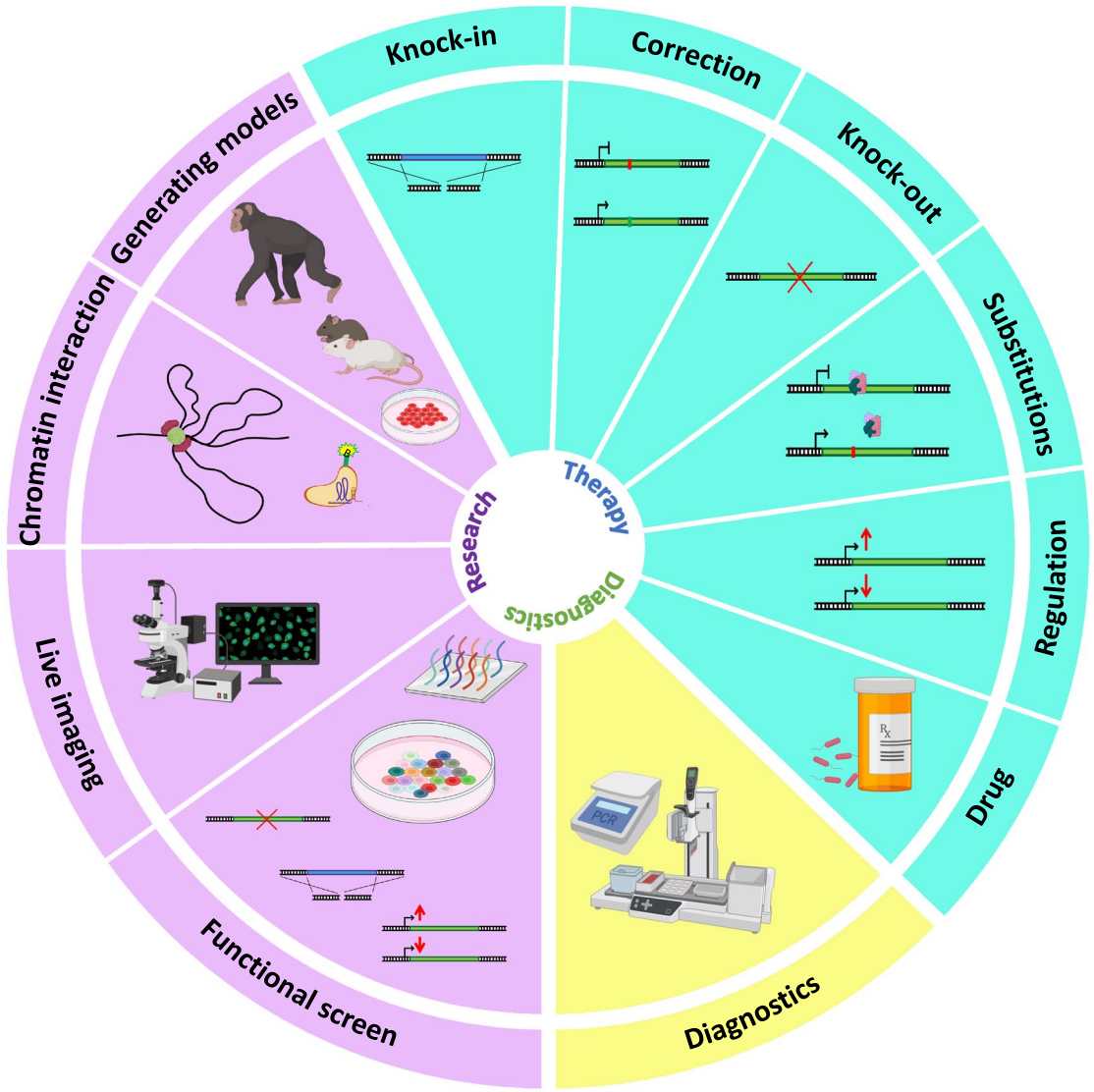


Fig. 4 Applications of CRISPR-Cas in medical science: CRISPR-Cas based technologies have varied applications in the field of medicine and can be broadly classified into research, therapeutic and diagnostic applications. The introduction of CRISPR in research has increased our understanding of biological system and has also facilitated the creation of cellular and animal models. Recent improvements, although in the early phase holds promise for CRISPR based gene therapy. CRISPR based diagnostics has also enabled rapid and easy detection of microbial as well as other diseases

library or CRISPRi library (e.g., Dolcetto, CRiNCL, Sub-pooled CRISPRi-v2 human library) and CRISPRa library (e.g., Calabrese, Human SAM genome-wide library, Human SAM lncRNA activation library) uses different types of sgRNAs to target the dCas9 bound to a transcriptional repressor and transcriptional activator respectively to regulate gene expression [72, 73].

Live cell imaging

dCas-9 fusion with fluorescent markers enables visualisation of chromosome regions in live cells. The region of interest can be targeted by using a sgRNA that will direct the binding of fluorescent tagged dCas9 which can be used to visualise the movement and location of the particular loci in the live cell. Repetitive or non-repetitive sequences in the chromosome can be imaged simultaneously in this way using single or multi-colour fluorescent tagging. This method can also aid in the rough estimation of distance between different loci. CRISPR-Sirius is an improved version which allows the detection of even low copy genomic loci [74–76].

Studying chromatin interactions and gene expression

dCas-9 platform can also be used for studying gene regulation by chromatin interactions. Termed enCHIP (engineered DNA-binding molecule-mediated chromatin immunoprecipitation), this method works on the basis of immunoprecipitation of the genomic locus of interest and the associated protein by using an antibody against a tagged dCas9 which is bound to the DNA with the help of sgRNA [77]. CRISPR affinity purification in situ of regulatory elements (CAPTURE) is one such method which utilizes dCas9 tagged with biotin to identify proteins interacting with the target DNA site. The sgRNA directs biotin tagged dCas-9 to the target site and once the protein is bound, protein-DNA interaction is fixed using formaldehyde. The chromatin is then sheared into small pieces and the pieces with biotin tagged dCas-9 are precipitated using streptavidin affinity. Mass spectrometry (for Trans Regulatory Elements) or deep sequencing (for Cis Regulatory Elements) is then used for identifying the regulatory elements. Recently an improved version of CAPTURE capable of multiplexed analysis of chromatin interactions was also described by which multiple enhancers and promoters can be studied in a single experiment [78, 79].

CRISPR-genome organization (CRISPR-GO) is another approach where repositioning of particular genomic loci to different nuclear locations is achieved using CRISPR and can be used to study and regulate gene expression based on gene position [80]. Chromatin loop re-organization using CRISPR-dCas9 (CLOuD9) is another method to reversibly induce the formation of chromatin loops to study their role in gene regulation. This method depends on the dimerisation of dCas9 fused with appropriate dimerising proteins targeted to two different genomic loci for inducing looping [81].

Generation of cellular and animal models

The advent of gene editing tools enabled rapid and efficient generation of cellular and animal models of genetic diseases. CRISPR/Cas-9 is a versatile and cost effective platform for generation of disease models owing to its ease of design and delivery. It can be used to create various types of mutations including small insertions/deletions, large deletions, or point mutations resulting in generation of precise models. The development of prime editing has further widened the scope of CRISPR based disease modelling. CRISPR/Cas-9 system can be delivered into the cells or embryo via electroporation, microinjection, adenoviral transduction etc. Ability to specifically control genes using CRISPR has helped in reprogramming cell fate and differentiation [82, 83].

Therapeutic applications

Therapeutic gene editing and gene therapy can be achieved either through ex vivo manipulation of cells or through in vivo delivery of gene editing tools. Although with limitations, both approaches are being harnessed for treatment of a variety of diseases including cancer, cardiovascular disorders, neurological disorders and haematological diseases.

Delivery of a copy of functional gene

In diseases involving defective genes, a functional copy of the gene can be delivered to be expressed by an endogenous promoter, utilising the HDR pathway in the cell after a double strand break. The donor gene is delivered either as a double or single stranded DNA template, which gets inserted into the target region based on homology to the region flanking the double strand break. One example is the correction of haemophilia by supplying a functional copy of factor IX. Although the efficiency is low, HDR mediated gene insertion is a feasible approach for correction of genetic diseases and can be expected to reach clinical trials soon [84, 85].

Correction of causative mutations

Many disorders like cystic fibrosis and sickle cell disease is caused by mutations that result in the loss of function. CRISPR/Cas9 has enabled the precise rectification of such mutations using ex vivo ssODN based gene correction. ssODN mediated correction of sickle cell disease has already been demonstrated in human HSCs. The latest addition to the CRISPR tool box, prime editing also holds promise for therapeutic correction of genetic mutations without causing any undesired edits [86–88].

Introduction of protective mutations

Creating a beneficial mutation can be useful in diseases that have a genetic and non-genetic aetiology. One such example is the introduction of CCR5 gene mutation in lymphocytes for protection against HIV. Another example is the introduction of anti-sickling mutations in the beta globin gene for the rescue of sickle cell disease [89].

Engineering of therapeutic cells

CAR T- cells hold great promise for the treatment of different types of cancer. CRISPR Cas system can be utilised for the efficient generation of CAR T- cells by knocking in of functional genes, knock out of genes for MHCs and receptors, inserting engineered CAR cassette into specific locus etc. Another example of therapeutic cells is engineered IPSCs which has then been differentiated to pancreatic beta

cells for the treatment of diabetes. IPSCs in theory can be differentiated into any cell type and can be used for cell based therapy for human diseases [90, 91].

Repression or activation of gene expression

CRISPR/Cas9 system allows for both regulated and unregulated activation or repression of gene expression in the human system. Disruption of an activator binding site can cause gene repression while disruption of repressor binding site can activate gene expression. Gene therapy for beta hemoglobinopathies by activating fetal haemoglobin expression through disruption of BCL11A binding site is already in clinical trials. Gene knockout of transcription repressors or activators using CRISPR is also being sought for regulation of gene expression [92, 93].

Use of appropriate sgRNAs that target two sites simultaneously can also lead to deletion or inversion of the sequence in between the two targets. This can be used for gene knockout as well as for therapy of certain disorders especially neurological disorders where repeat sequences are pathogenic. Recently precise genomic deletions using dual prime editing gRNAs has also been successfully demonstrated.

The outcome of gene editing can be analysed at all molecular levels such as DNA (sequencing, PCR, T7 endonuclease, surveyor assay etc.), RNA (qRT PCR, RNA seq) or at the protein expression level (western blot, FACS etc.). The choice of method will depend on the expected outcome as well as the level of validation required for the experiment in question.

Development of anti-microbial agents: CRISPR/Cas9 system targeting bacterial virulence gene delivered via bacteriophages can efficiently kill the bacteria. This system can be engineered to target only the virulent strains while allowing the survival of non-virulent ones. Resistant bacteria can also be sensitized to antibiotics by targeting the sgRNA to antibiotic resistance gene either in the bacterial chromosome or in intracellular plasmids [94].

Diagnostic applications

CRISPR/Cas system has gained popularity as a diagnostic tool for both microbial and non microbial diseases. The first diagnostic system using CRISPR was developed by Pardee et al. in 2016 for the detection of Zika virus [95]. Later SHERLOCK was developed as a diagnostic platform for nucleic acid detection relying on Cas13 [96]. Numerous other researchers also developed CRISPR based nucleic acid detection methods latest addition being its use in detecting COVID-19 [97–99]. DETECTR which is a CRISPR based DNA detection tool, in addition to being accurate also requires less turnaround time [37]. The major disadvantage of CRISPR based diagnostics is the off target effect which can give false positive results and need for high nucleic acid level (viral load), thereby increasing the chance of false negatives. In addition to detecting viral, bacterial and fungal pathogens CRISPR can also be utilised for cancer detection. Nevertheless CRISPR bases diagnosis appears to be a promising tool for easy, rapid and cost effective diagnosis of infectious and non-infectious diseases [100] (Fig. 4).

Challenges in CRISPR based gene editing

Ever since its discovery there has been a considerable amount of scientific effort in improving CRISPR/Cas9 technology for gene editing resulting in highly efficient and better versions. Even so there are still some hurdles along the way for this system, the most important being it's off target effect. Off-target binding of the sgRNA is a problem in both research as well as therapy, sometimes causing unintended effects and sometimes masking the desired effect [101]. Although careful sgRNA designing can reduce the chances for off target editing, it is important to rule out any undesired editing outcomes. Possible off target sites can be predicted using in silico tools based on different algorithms and can be subjected to high throughput sequencing to check for off target editing. Unbiased in vivo and in vitro genome wide assays (SITE-seq, CIRCLE-seq, Digenome-seq, BLESS, IDLVS, Guide-seq) can also be used for off target analysis [102–104]. Although there are a considerable number of off target prediction and detection tools, no technique can detect off targets with complete efficiency and the best methods often require expensive whole genome sequencing [101]. The fact that CRISPR/Cas9 system behave differently in each individual obviates the need for off target analysis in each patient even though the sgRNA used is the same. Another challenge is finding the optimal delivery strategy; each one possessing its own merits and demerits. Lenti-viral delivery, although efficient, possesses the risk of random integration and sustained expression. Delivery as Cas9 RNP complex or AAV mediated delivery are the most preferred methods to date. The concerns regarding the delivery using viral vectors has also prompted a drift towards non-viral delivery methods like nanoparticles and liposomes for in vivo gene editing [30]. Safety concerns regarding the immunogenicity and potential oncogenesis by Cas9 mediated gene editing is also not minimal. Recently, it has been shown that double strand breaks lead to p53 activation and cell death. Therefore, the cells which survive during gene editing might have a less active p53 pathway, posing a risk of oncogenesis in a later stage. It has also been demonstrated that a large proportion of the population possess antibodies against Cas9 protein. Although not a concern in ex vivo gene editing, immunogenicity is a serious obstacle for in vivo gene editing and is thus the main reason for research on Cas

orthologues [105, 106]. Other factors like cost, ethical concerns and regulatory challenges also needs to be considered before CRSIPR based gene editing can enter main stream medical application.

Conclusion

The story of how a bacterial immune system has been repurposed for gene editing is an inspiring one. It highlights the importance of basic science research in advancing medical research and care. The ease, versatility and cost effectiveness of CRISPR has made it a very popular gene editing tool accessible to even the smallest, less funded laboratories around the world and has in turn contributed to the diversified research applications ranging from bacterial and plant research to gene editing for human diseases. Many laboratories across the world has worked towards improving gene editing using CRISPR and has contributed to improved versions with better efficiency, specificity and safety. Research still continues to identify new Cas orthologues and also to improve the existing versions and the scope of application. CRISPR has proved to be an indispensable tool in both research and therapy within this short period of its discovery and can be expected to benefit medical science in an unprecedented manner.

Acknowledgements Schematic representations were generated in Adobe Illustrator with the support of Biorender (©BioRender-biorender.com).

Funding This work was supported by EMR Grant: EMR/2017/004363 (Science and Engineering Research Board (SERB), New Delhi, India), ICMR Grant: 2019-0916/SCR/ADHOC_BMS (Indian Council of Medical Research, New Delhi, India). Anila George is supported by Research Fellowship from the Council of Scientific and Industrial Research (CSIR), India.

Declarations

Conflict of interest The authors declare no conflicts of interest.

References

1. Haber JE (2016) A life investigating pathways that repair broken chromosomes. Annu Rev Genet 23(50):1–28
2. Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW (1983) The double-strand-break repair model for recombination. Cell 33(1):25–35
3. Rouet P, Smih F, Jasin M (1994) Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. Mol Cell Biol 14(12):8096–8106
4. Kim YG, Cha J, Chandrasegaran S (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci U S A 93(3):1156–1160
5. Bibikova M, Carroll D, Segal DJ, Trautman JK, Smith J, Kim YG et al (2001) Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. Mol Cell Biol 21(1):289–297
6. Li T, Huang S, Zhao X, Wright DA, Carpenter S, Spalding MH et al (2011) Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. Nucleic Acids Res 39(14):6315–6325
7. Moscou MJ, Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. Science 326(5959):1501
8. Joung JK, Sander JD (2013) TALENs: a widely applicable technology for targeted genome editing. Nat Rev Mol Cell Biol 14(1):49–55
9. Mojica FJ, Juez G, Rodríguez-Valera F (1993) Transcription at different salinities of Haloferax mediterranei sequences adjacent to partially modified PstI sites. Mol Microbiol 9(3):613–621
10. Mojica FJM, Díez-Villaseñor C, García-Martínez J, Soria E (2005) Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J Mol Evol 60(2):174–182
11. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. Science 337(6096):816–821
12. Gasiunas G, Barrangou R, Horvath P, Siksnys V (2012) Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc Natl Acad Sci 109(39):E2579–E2586
13. Hille F, Richter H, Wong SP, Bratovič M, Ressel S, Charpentier E (2018) The biology of CRISPR-Cas: backward and forward. Cell 172(6):1239–1259
14. Mojica FJM, Rodriguez-Valera F (2016) The discovery of CRISPR in archaea and bacteria. FEBS J 283(17):3162–3169
15. Karginov FV, Hannon GJ (2010) The CRISPR system: small RNA-guided defense in bacteria and archaea. Mol Cell 37(1):7
16. Gleditzsch D, Pausch P, Müller-Esparza H, Özcan A, Guo X, Bange G et al (2018) PAM identification by CRISPR-Cas effector complexes: diversified mechanisms and structures. RNA Biol 16(4):504–517
17. Carroll D (2014) Genome engineering with targetable nucleases. Annu Rev Biochem 83(1):409–439
18. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F (2013) Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8(11):2281–2308
19. Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N et al (2014) Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell 156(5):935–949
20. Zhu Y, Huang Z (2019) Recent advances in structural studies of the CRISPR-Cas-mediated genome editing tools. Natl Sci Rev 6(3):438–451
21. Jiang F, Doudna JA (2017) CRISPR–Cas9 structures and mechanisms. Annu Rev Biophys 46(1):505–529
22. Mekler V, Minakhin L, Semenova E, Kuznedelov K, Severinov K (2016) Kinetics of the CRISPR-Cas9 effector complex assembly and the role of 3′-terminal segment of guide RNA. Nucleic Acids Res 44(6):2837–2845
23. Lim Y, Bak SY, Sung K, Jeong E, Lee SH, Kim J-S et al (2016) Structural roles of guide RNAs in the nuclease activity of Cas9 endonuclease. Nat Commun 7(1):13350
24. Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature 507(7490):62–67

25. Szczelkun MD, Tikhomirova MS, Sinkunas T, Gasiunas G, Karvelis T, Pschera P et al (2014) Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. Proc Natl Acad Sci 111(27):9798–9803
26. Gong S, Yu HH, Johnson KA, Taylor DW (2018) DNA unwinding is the primary determinant of CRISPR-Cas9 activity. Cell Rep 22(2):359–371
27. Liu M, Rehman S, Tang X, Gu K, Fan Q, Chen D et al (2018) Methodologies for improving HDR efficiency. Front Genet 9:691
28. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE et al (2013) RNA-guided human genome engineering via Cas9. Science 339(6121):823–826
29. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N et al (2013) Multiplex genome engineering using CRISPR/Cas systems. Science 339(6121):819–823
30. Lino CA, Harper JC, Carney JP, Timlin JA (2018) Delivering CRISPR: a review of the challenges and approaches. Drug Deliv 25(1):1234–1257
31. Kim S, Kim D, Cho SW, Kim J, Kim J-S (2014) Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Res 24(6):1012–1019
32. Liang X, Potter J, Kumar S, Zou Y, Quintanilla R, Sridharan M et al (2015) Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. J Biotechnol 20(208):44–53
33. Yang Y, Wang L, Bell P, McMenamin D, He Z, White J et al (2016) A dual AAV system enables the Cas9-mediated correction of a metabolic liver disease in newborn mice. Nat Biotechnol 34(3):334–338
34. Carvalho LS, Turunen HT, Wassmer SJ, Luna-Velez MV, Xiao R, Bennett J et al (2017) Evaluating efficiencies of dual AAV approaches for retinal targeting. Front Neurosci 11:503
35. Yin H, Song C-Q, Suresh S, Wu Q, Walsh S, Rhym LH et al (2017) Structure-guided chemical modification of guide RNA enables potent non-viral in vivo genome editing. Nat Biotechnol 35(12):1179–1187
36. Makarova KS, Wolf YI, Iranzo J, Shmakov SA, Alkhnbashi OS, Brouns SJJ et al (2020) Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. Nat Rev Microbiol 18(2):67–83
37. Chen JS, Ma E, Harrington LB, Da Costa M, Tian X, Palefsky JM et al (2018) CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. Science 360(6387):436–439
38. Abudayyeh OO, Gootenberg JS, Essletzbichler P, Han S, Joung J, Belanto JJ et al (2017) RNA targeting with CRISPR–Cas13. Nature 550(7675):280–284
39. Friedland AE, Baral R, Singhal P, Loveluck K, Shen S, Sanchez M et al (2015) Characterization of Staphylococcus aureus Cas9: a smaller Cas9 for all-in-one adeno-associated virus delivery and paired nickase applications. Genome Biol 16(1):257
40. Lee JK, Jeong E, Lee J, Jung M, Shin E, Kim Y et al (2018) Directed evolution of CRISPR-Cas9 to increase its specificity. Nat Commun 9(1):3048
41. Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F (2016) Rationally engineered Cas9 nucleases with improved specificity. Science 351(6268):84–88
42. Chen JS, Dagdas YS, Kleinstiver BP, Welch MM, Sousa AA, Harrington LB et al (2017) Enhanced proofreading governs CRISPR–Cas9 targeting accuracy. Nature 550(7676):407–410
43. Kulcsár PI, Tálas A, Huszár K, Ligeti Z, Tóth E, Weinhardt N et al (2017) Crossing enhanced and high fidelity SpCas9 nucleases to optimize specificity and cleavage. Genome Biol 18(1):190
44. Casini A, Olivieri M, Petris G, Montagna C, Reginato G, Maule G et al (2018) A highly specific SpCas9 variant is identified by in vivo screening in yeast. Nat Biotechnol 36(3):265–271
45. Ikeda A, Fujii W, Sugiura K, Naito K (2019) High-fidelity endonuclease variant HypaCas9 facilitates accurate allele-specific gene modification in mouse zygotes. Commun Biol 2(1):1–7
46. Kleinstiver BP, Pattanayak V, Prew MS, Tsai SQ, Nguyen NT, Zheng Z et al (2016) High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. Nature 529(7587):490–495
47. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK et al (2013) High frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol 31(9):822–826
48. Kim H, Kim J-S (2014) A guide to genome engineering with programmable nucleases. Nat Rev Genet 15(5):321–334
49. Tian P, Wang J, Shen X, Rey JF, Yuan Q, Yan Y (2017) Fundamental CRISPR-Cas9 tools and current applications in microbial systems. Synth Syst Biotechnol 2(3):219–225
50. Brocken DJW, Tark-Dame M, Dame RT (2018) dCas9: a versatile tool for epigenome editing. Curr Isuues Mol Biol 26:18
51. Duke CG, Bach SV, Revanna JS, Sultan FA, Southern NT, Davis MN et al (2020) An improved CRISPR/dCas9 interference tool for neuronal gene suppression. Front Genome Ed. https://doi.org/10.3389/fgeed.2020.00009/full
52. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature 533(7603):420–424
53. Nishida K, Arazoe T, Yachie N, Banno S, Kakimoto M, Tabata M et al (2016) Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. Science 353(6305)
54. Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI et al (2017) Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. Nature 551(7681):464–471
55. Kantor A, McClements ME, MacLaren RE (2020) CRISPR-Cas9 DNA base-editing and prime-editing. Int J Mol Sci 21(17):6240
56. Anzalone AV, Koblan LW, Liu DR (2020) Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. Nat Biotechnol 38(7):824–844
57. Rees HA, Liu DR (2018) Base editing: precision chemistry on the genome and transcriptome of living cells. Nat Rev Genet 19(12):770–788
58. Kurt IC, Zhou R, Iyer S, Garcia SP, Miller BR, Langner LM et al (2020) CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. Nat Biotechnol 20:1–6
59. Sakata RC, Ishiguro S, Mori H, Tanaka M, Tatsuno K, Ueda H et al (2020) Base editors for simultaneous introduction of C-to-T and A-to-G mutations. Nat Biotechnol 38(7):865–869
60. Halperin SO, Tou CJ, Wong EB, Modavi C, Schaffer DV, Dueber JE (2018) CRISPR-guided DNA polymerases enable diversification of all nucleotides in a tunable window. Nature 560(7717):248–252
61. Anzalone AV, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM et al (2019) Search-and-replace genome editing without double-strand breaks or donor DNA. Nature 576(7785):149–157
62. Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffini LA (2013) Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic Acids Res 41(15):7429–7437
63. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP et al (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152(5):1173–1183
64. Larson MH, Gilbert LA, Wang X, Lim WA, Weissman JS, Qi LS (2013) CRISPR interference (CRISPRi) for sequence-specific control of gene expression. Nat Protoc 8(11):2180–2196
65. Dahlman JE, Abudayyeh OO, Joung J, Gootenberg JS, Zhang F, Konermann S (2015) Orthogonal gene knockout and

activation with a catalytically active Cas9 nuclease. Nat Biotechnol 33(11):1159–1161
66. Ding X, Seebeck T, Feng Y, Jiang Y, Davis GD, Chen F (2019) Improving CRISPR-Cas9 genome editing efficiency by fusion with chromatin-modulating peptides. CRISPR J 2(1):51–63
67. Xie N, Zhou Y, Sun Q, Tang B (2018) Novel epigenetic techniques provided by the CRISPR/Cas9 system. Stem Cells Int 2018:7834175
68. Pulecio J, Verma N, Mejía-Ramírez E, Huangfu D, Raya A (2017) CRISPR/Cas9-based engineering of the epigenome. Cell Stem Cell 21(4):431–447
69. Adelmann CH, Wang T, Sabatini DM, Lander ES (2019) Genome-wide CRISPR/Cas9 screening for identification of cancer genes in cell lines. Methods Mol Biol 1907:125–136
70. Yu JSL, Yusa K (2019) Genome-wide CRISPR-Cas9 screening in mammalian cells. Methods 15(164–165):29–35
71. Joung J, Konermann S, Gootenberg JS, Abudayyeh OO, Platt RJ, Brigham MD et al (2017) Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. Nat Protoc 12(4):828–863
72. Sanson KR, Hanna RE, Hegde M, Donovan KF, Strand C, Sullender ME et al (2018) Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. Nat Commun 9(1):5416
73. Momen-Roknabadi A, Oikonomou P, Zegans M, Tavazoie S (2020) An inducible CRISPR interference library for genetic interrogation of saccharomyces cerevisiae biology. Commun Biol 3(1):1–12
74. Wu X, Mao S, Ying Y, Krueger CJ, Chen AK (2019) Progress and challenges for live-cell imaging of genomic loci using CRISPR-based platforms. Genomics Proteomics Bioinformatics 17(2):119–128
75. Zhou Y, Wang P, Tian F, Gao G, Huang L, Wei W et al (2017) Painting a specific chromosome with CRISPR/Cas9 for live-cell imaging. Cell Res 27(2):298–301
76. Qin P, Parlak M, Kuscu C, Bandaria J, Mir M, Szlachta K et al (2017) Live cell imaging of low- and non-repetitive chromosome loci using CRISPR-Cas9. Nat Commun 8(1):14725
77. Fujita T, Fujii H (2013) Efficient isolation of specific genomic regions and identification of associated proteins by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) using CRISPR. Biochem Biophys Res Commun 439(1):132–136
78. Liu X, Zhang Y, Chen Y, Li M, Zhou F, Li K et al (2017) In situ capture of chromatin interactions by Biotinylated dCas9. Cell 170(5):1028-1043.e19
79. Liu X, Chen Y, Zhang Y, Liu Y, Liu N, Botten GA et al (2020) Multiplexed capture of spatial configuration and temporal dynamics of locus-specific 3D chromatin by biotinylated dCas9. Genome Biol 21(1):59
80. Wang H, Xu X, Nguyen CM, Liu Y, Gao Y, Lin X et al (2018) CRISPR-mediated programmable 3D genome positioning and nuclear organization. Cell 175(5):1405-1417.e14
81. Morgan SL, Mariano NC, Bermudez A, Arruda NL, Wu F, Luo Y et al (2017) Manipulation of nuclear architecture through CRISPR-mediated chromosomal looping. Nat Commun 13(8):15993
82. Zarei A, Razban V, Hosseini SE, Tabei SMB (2019) Creating cell and animal models of human disease by genome editing using CRISPR/Cas9. J Gene Med 21(4):e3082
83. Antao AM, Karapurkar JK, Lee DR, Kim K-S, Ramakrishna S (2020) Disease modeling and stem cell immunoengineering in regenerative medicine using CRISPR/Cas9 systems. Comput Struct Biotechnol J 1(18):3649–3665
84. Park C-Y, Lee DR, Sung JJ, Kim D-W (2016) Genome-editing technologies for gene correction of hemophilia. Hum Genet 135(9):977–981
85. Wang Q, Zhong X, Li Q, Su J, Liu Y, Mo L et al (2020) CRISPR-Cas9-mediated in vivo gene integration at the albumin locus recovers hemostasis in neonatal and adult hemophilia B mice. Mol Ther Methods Clin Dev 30(18):520–531
86. DeWitt MA, Magis W, Bray NL, Wang T, Berman JR, Urbinati F et al (2016) Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells. Sci Transl Med 8(360):360ra134
87. Pattabhi S, Lotti SN, Berger MP, Singh S, Lux CT, Jacoby K et al (2019) In vivo outcome of homology-directed repair at the HBB gene in HSC using alternative donor template delivery methods. Mol Ther Nucleic Acids 7(17):277–288
88. Park SH, Lee CM, Dever DP, Davis TH, Camarena J, Srifa W et al (2019) Highly efficient editing of the β-globin gene in patient-derived hematopoietic stem and progenitor cells to treat sickle cell disease. Nucleic Acids Res 47(15):7955–7972
89. Xu L, Wang J, Liu Y, Xie L, Su B, Mou D et al (2019) CRISPR-edited stem cells in a patient with HIV and acute lymphocytic leukemia. N Engl J Med 381(13):1240–1247
90. Li C, Mei H, Hu Y (2020) Applications and explorations of CRISPR/Cas9 in CAR T-cell therapy. Brief Funct Genomics 19(3):175–182
91. Lim D, Sreekanth V, Cox KJ, Law BK, Wagner BK, Karp JM et al (2020) Engineering designer beta cells with a CRISPR-Cas9 conjugation platform. Nat Commun 11(1):4043
92. Psatha N, Reik A, Phelps S, Zhou Y, Dalas D, Yannaki E et al (2018) Disruption of the BCL11A erythroid enhancer reactivates fetal hemoglobin in erythroid cells of patients with b-thalassemia major. Mol Ther 10:313–326
93. Weber L, Frati G, Felix T, Hardouin G, Casini A, Wollenschlaeger C et al (2020) Editing a γ-globin repressor binding site restores fetal hemoglobin synthesis and corrects the sickle cell disease phenotype. Sci Adv 6(7):eaay9392
94. Li Y, Peng N (2019) Endogenous CRISPR-Cas system-based genome editing and antimicrobials: review and prospects. Front Microbiol. https://doi.org/10.3389/fmicb.2019.02471/full
95. Pardee K, Green AA, Takahashi MK, Braff D, Lambert G, Lee JW et al (2016) Rapid, low-cost detection of zika virus using programmable biomolecular components. Cell 165(5):1255–1266
96. Kellner MJ, Koob JG, Gootenberg JS, Abudayyeh OO, Zhang F (2019) SHERLOCK: nucleic acid detection with CRISPR nucleases. Nat Protoc 14(10):2986–3012
97. Chiu C (2018) Cutting-edge infectious disease diagnostics with CRISPR. Cell Host Microbe 23(6):702–704
98. Wang M, Zhang R, Li J (2020) CRISPR/cas systems redefine nucleic acid detection: principles and methods. Biosens Bioelectron 165:112430
99. Broughton JP, Deng X, Yu G, Fasching CL, Servellita V, Singh J et al (2020) CRISPR–Cas12-based detection of SARS-CoV-2. Nat Biotechnol 38(7):870–874
100. Jolany Vangah S, Katalani C, Boone HA, Hajizade A, Sijercic A, Ahmadian G (2020) CRISPR-based diagnosis of infectious and noninfectious diseases. Biol Proced Online 22(1):22
101. Kim D, Luk K, Wolfe SA, Kim J-S (2019) Evaluating and enhancing target specificity of gene-editing nucleases and deaminases. Annu Rev Biochem 88(1):191–220
102. Naeem M, Majeed S, Hoque MZ, Ahmad I (2020) Latest developed strategies to minimize the off-target effects in CRISPR-Cas-mediated genome editing. Cells 9(7):1068
103. Kang S-H, Lee W, An J-H, Lee J-H, Kim Y-H, Kim H et al (2020) Prediction-based highly sensitive CRISPR off-target validation using target-specific DNA enrichment. Nat Commun 11(1):3596
104. Cui Y, Liao X, Peng S, Tang T, Huang C, Yang C (2020) OffScan: a universal and fast CRISPR off-target sites detection tool. BMC Genomics 21(1):872

105. Enache OM, Rendo V, Abdusamad M, Lam D, Davison D, Pal S et al (2020) Cas9 activates the p53 pathway and selects for p53-inactivating mutations. Nat Genet 52(7):662–668
106. Charlesworth CT, Deshpande PS, Dever DP, Camarena J, Lemgart VT, Cromer MK et al (2019) Identification of preexisting adaptive immunity to Cas9 proteins in humans. Nat Med 25(2):249–254