CRISPR-Cas9: Revolution in Gene Editing

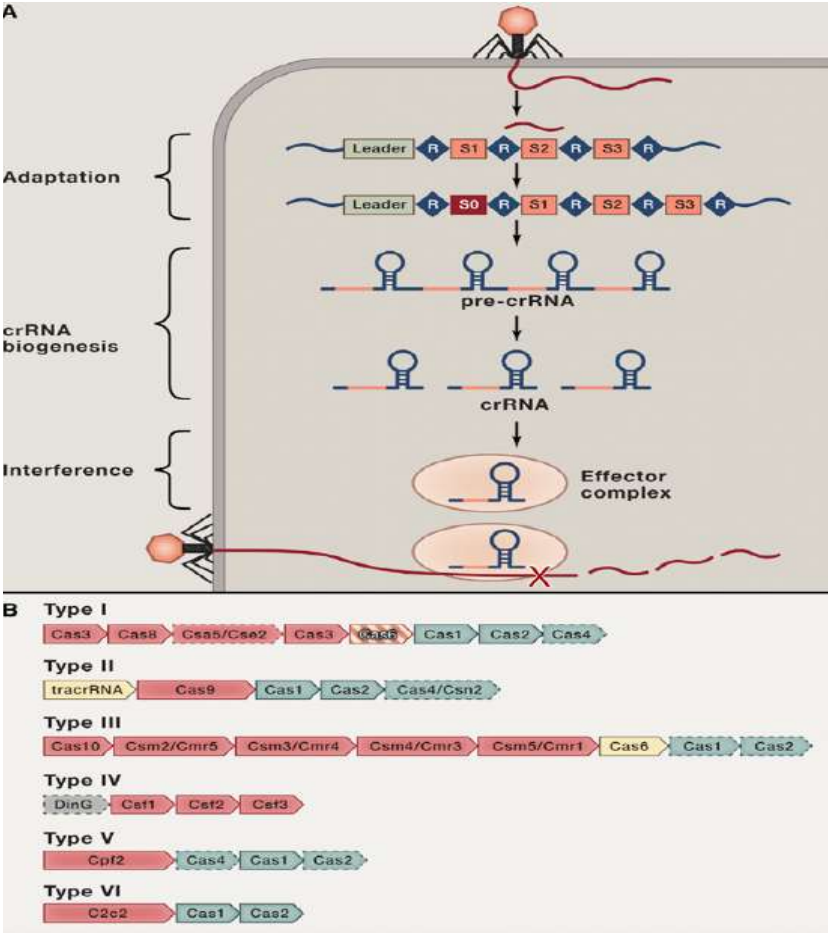
Mou Saha[1]; Sreayosi Sinha Roy[2]; Piali Chakraborty[3]

[1]Assistant Professor ; Department of Microbiology, R.K.V.M. Sarada Ma Girls' College ; Barasat, Kolkata-700126; Email- saha.mou136@gmail.com

[2]Post Graduate Student of Biomedical Genetics; Vellore Institute of Technology, Vellore; Email- sreayosis@gmail.com

[3]Undergraduate student of Microbiology, R.K.V.M. Sarada Ma Girls' College; Barasat, Kolkata-700126; Email- cpiali023@gmail.com

Gene editing and CRISPR -Cas9



Gene editing enables scientists to engineer DNA of organisms and thus encompasses an innovative line-up of technologies. These innovative techniques empower researchers to modify, insert or remove genetic material at exact locations within an organism's genome. Among these advancements, one notably prominent technique is CRISPR-Cas9, implying CRISPR-linked protein-9. Basically they are arranged in DNA as short palindromic repeated sequences in clustered form in regularly interspaced manner. It is one of the most prominent techniques within the scientific community than other gene editing methods. CRISPR-Cas9 has incited significant enthusiasm due to its remarkable attributes, including its swiftness, affordability, precision, and efficiency.

Fig: Organization of CRISPR-Cas9 system

Using CRISPR-Cas9 as a form of immune defence, a naturally occurring genome editing mechanism in which bacteria capture viral DNA fragments, therefore, integrating the uptaken fragments with their inherent DNA in a precise manner, resulting in CRISPR arrangement.

These arrangements enable bacteria to recognize previously infected viruses or correlated organisms. If a subsequent viral invasion occurs within the same bacterial cell, RNA segments are produced within the bacteria from the CRISPR arrangements that integrate to certain segments of the viral genetic material. In this regard, bacteria utilize Cas9 enzymes to cleave the viral DNA, thereby neutralizing the threat posed by the virus.

○ History

- **The invention of clustered DNA repeats** – Clustered DNA repeats were discovered separately in three different corners of the world. In 1987, at Osaka University in Japan; molecular biologist Yoshizumi Ishino et al. first described this technology that would later be known as CRISPR. The researchers unintentionally cloned iap gene (functions as alkaline phosphatase isozyme conversion), along with a segment of CRISPR sequence. Frequently repeated sequences are arranged in a straight lined order where no other sequences are present in between them. The scientists were unacquainted of the functional role of this clustered repeated parts.

  Later in 1993 Dutch researchers studied the interrupted direct repeats (DR) lying in clustered form in *Mycobacterium tuberculosis* genome and found diverse sequences positioned between the direct repeats in several strains of *Mycobacterium tuberculosis*. They identified the diversity of the sequences that stood between the direct repeats in various *M. tuberculosis* strains. Now a days this feature is used to develop spoligotyping procedure.

  In Spain at the University of Alicante, Francisco Mojica studied purpose of repeats in archaeal organisms like *Haloarcula* and *Haloferax*. He believed these clustered repeats function for segregation of duplicated DNA and producing daughter cells during cell division. It was the first detailed characterization of CRISPR which was identified through the transcription of interrupted sequence repetitions. They determined that 20 species of microorganisms with interrupted repeats belonged to the same family. Mojica initially referred to these sequences as "short regularly spaced repeats" (SRSR) because of the interspaced nature of those sequences.

  The abbreviation of **CRISPR** (Clustered Regularly Interspaced Short Palindromic Repeats) was introduced by **Mojica** and **Ruud Jansen** in 2001, who were looking for further interrupted repeats in order to reduce the confusion caused by the multiple abbreviations used to depict the sequences in the scientific descriptions.

- **Cas9 and PAM region Identification**: Alexander Bolotin , a researcher of National Institute for Agricultural Research, France, was researching the recently sequenced bacteria Streptococcus thermophilus, which featured a unique CRISPR locus. The CRISPR array resembled previously described systems, but some of the recognized Cas genes were missing. This studied arrangement contained some unique Cas genes,one of which codes for Cas 9 ,a large protein which have nuclease activity . Researchers studied about the spacers which is analogous to the viral genes both of which have common sequence in their one terminal end. He classified these sequences as **The protospacer adjacent motif (PAM).**
- **Hypothetical approach to adaptive immunity:** Eugene Koonin, a researcher of the US National Centre for Biotechnology Information**,** dismissed the previously studied theory abot the Cas protein which might form a novel DNA repair set-up. While researching clusters of orthologous protein groups using computational analysis. He then hypothetically proposed CRISPR cascades as a part of bacterial immune system depending on insertion of homologous DNA to phage DNA in the portion of spacer array.

- **Adaptive immunity demonstrated experimentally:** *Streptococcus thermophilus* is a popular strain used in the dairy sector to produce yoghurt and cheese. Researchers at Danisco wanted to learn more about how S. thermophilus handles phage attack, a typical issue in the production of industrial yoghurt. According to an experimental demonstration by Horvath and colleagues, as they incorporate newly synthesized phage DNA into their arrangement CrispR system actually functions as adaptive immune system. Cas9 protein mostly functions as the only protein necessary for interference but the exact mechanism of neutralising the attacking phage is still unknown.
- **Transcription of spacer sequences into guide RNAs:** Scientist of University of Wageningen, Netherlands John van der Oost proposed the specifics about the mechanism by which CRISPR-Cas system neutralise the attacking phage . The first important piece of information was provided by John van der Oost et al. who demonstrated that in *Escherichia coli*, spacer sequences obtained from phage are translated into short RNAs which are known as CRISPR RNAs (crRNAs), which direct Cas proteins for targeting the DNA.

- **Discovery of the target molecule**: Marraffini and Sontheimer's invention was the next crucial step in understanding the mechanism of interference (Marraffini and Sontheimer, 2008) that not RNA but DNA is the target molecule. This invention was surprizing since mostly everyone believed that CRISPR had similarities with eukaryotic RNA interference silencing mechanisms, targeting the RNA.Sontheimer and Marraffini stated bu their experiment about the application of this system. According to their study,f this method get applied to non-bacterial system it will be more beneficial tool .
- **Cleaving mechanism of Cas9**: In 2010, Sylvain Moineau et al showed that CRISPR-Cas9 causes double-stranded breaks in the target DNA at University of Laval, Canada. Additionally, they verified that the CRISPR-Cas9 system is the only system which

requires Cas9 protein as the necessary protein for cleavage. The interference is carried out by a single large protein (Cas9 in this case) and this is the significance of type II CRISPR-Cas9 system.

- **tracrRNA discovery**: In 2011, Emmanuelle Charpentier et al. proposed the ultimate mechanism of natural CRISPR-Cas9-guided interference. They implemented small RNA sequencing on the CRISPR-Cas9-equipped Streptococcus pyogenes. As a result , they discovered a second short RNA in addition to the crRNA, which was then named trans-activating CRISPR RNA (tracrRNA). In conclusion, they stated that Cas9 is guided to its targets by the duplex created by tracrRNA and crRNA.

- **Biochemical characterization of Cas9 mediated cleavage**: Researcher Virginijus Siksnys and his fellows examined E. coli strain to purify Cas9 in complex with crRNA. They modified it with CRISPR locus from the *S. thermophiles* and performed different biochemical experiments to understand the mode of action. This was done to take advantage of their different system. Scientists confirmed the location of cleavage and the need for the PAM . They used point mutations they proved that the HNH domain present in Cas9 plays the role in cleaving the complementary site. On the other hand, non-complementary strands are ckeaved by the RuvC domain only.

    - **Introducing CRISPR Cas9 into genetic engineering**: In 2013, American researcher Feng Zhang, used two orthologs of Cas9 derived from *S. pyogenes* and *S. thermophilus*. Zhang et al. performed the cleavage of the targeted genome in mice and human cells. Their experiment led to the first successful adaptation of CRISPR-Cas9 for genome editing in eukaryotic cells. Additionally, they illustrated how the first system was assumed to be set up to target numerous genomic locations, and the second system could activate Homology-directed repair (hdr). Researchers from the lab of George Church's Harvard University published the analogous discoveries in the same Science publication issue.

○ Structural configurations of CRISPR Cas9

Two main components of CRISPR Cas9 are there in the prokaryotic system:

→ CRISPR arrays and

→ CRISPR-associated (Cas) proteins.

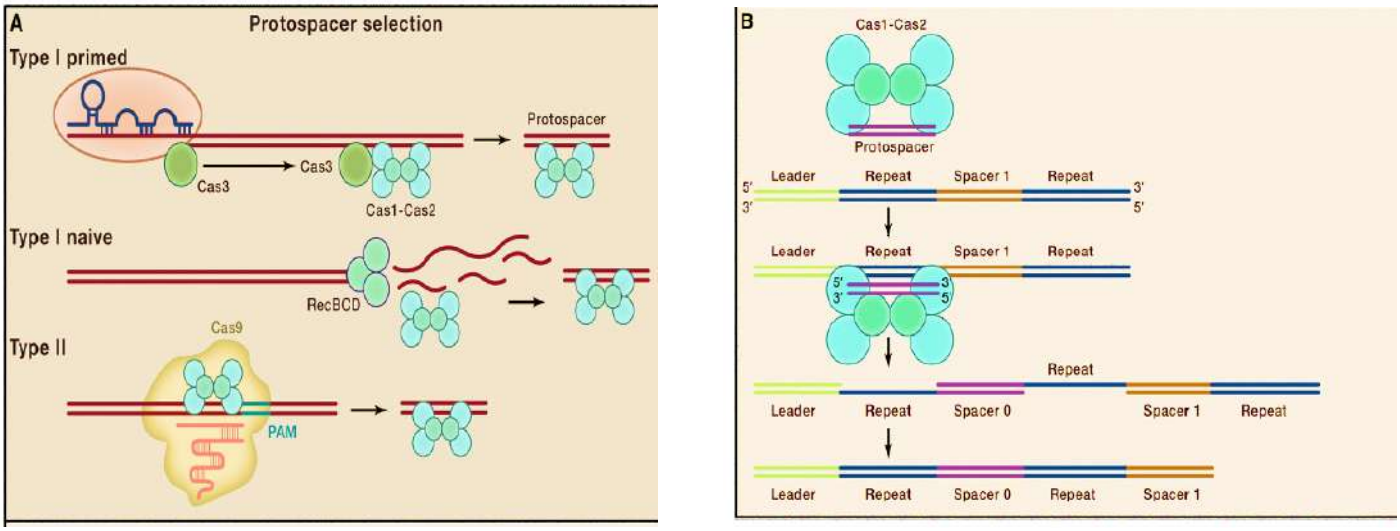


Fig: Process of Protospacer selection and insertion as Spacer

The first component CRISPRs is termed as Clustered Regularly Interspaced Short Palindromic Repeats which is a distinct type of repetitive DNA stretches. This This CRISPR array is all in one part of the chromosome, gathering in one cluster.

.

- Regularly interspaced is referring the spaces that are regularly placed between these repeats all along the CRISPR region. **Spacers** are highly variable sequences that comprise the prokaryotes 'immunological memory
  Generation of spacers: when a bacteriophage injects its DNA inside a bacteria, a pair of enzymes (Cas1 and Cas2) works together and cuts off a region of viral DNA. This piece of viral DNA is called a **protospacer**. This protospacer is then get inserted into the 5' teminal end of the CRISPR arrangement as a **spacer.** In this way, these protospacers get inserted into the array as spacers eventually build a new repeat region.

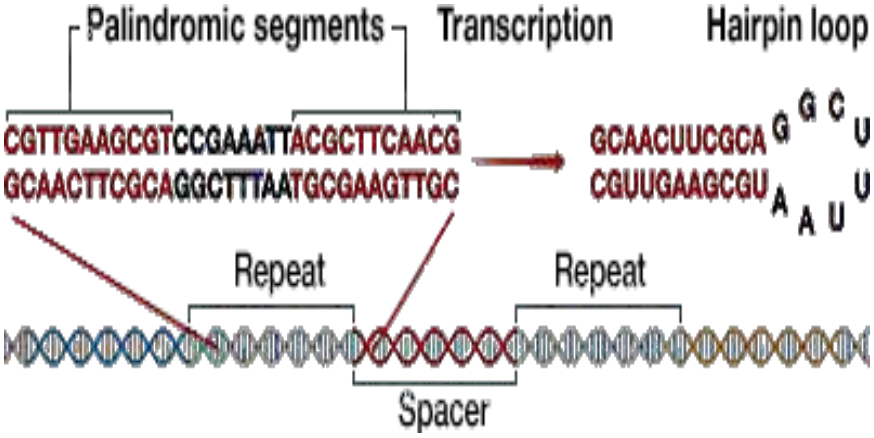


Fig: Parts of CRISPE array

- Another distinctive and essential element within the CRISPR array is **palindromic repeats**. In other words, these are regions in which a DNA or RNA sequence reads the same both forward and backwards. They play a fundamental role in the system's functioning. These palindromic repeats provide a distinctive pattern that is translated into RNA in conjunction with the spacer sequences. The Cas proteins, in particular the Cas9 enzyme, are directed by these RNA molecules, known as "guide RNAs" (gRNAs) or "crRNAs," to recognise and target particular viral DNA sequences for elimination. The complementary nature of these spacer sequences in the guide RNA and the target virus DNA is what enables identification. The presence of these repeats emphasises about the support provided by the structural components to the CRISPR-Cas system which are conserved across diverse bacterial species.

- A crucial part of the CRISPR-Cas9 system is the **PAM region**. It ensures that only the other DNA sequences in the genome cuts off by the Cas9 protein. Fundamentally this region ensures the specificity of the CRISPR-Cas9 system. The **protospacer adjacent motif**, or **PAM region**, is a short DNA sequence that is present right next to the target DNA sequence that the CRISPR-Cas9 system will cut. Normally this region consists of 2–6 base pairs and is palindromic, which means that it may be read both forward and backwards. The sequence 5'-NGG-3', where "N" can be any nucleotide, is the most frequent PAM sequence. However, other PAM sequences, including 5'-NG-3' and 5'-NAG-3', are also feasible.

  The two steps of the CRISPR-Cas9 system are **adaptation** and **interference**. The interference stage is responsible for the identification and chopping of invasive DNA, while the adaption stage is responsible for acquiring new spacers from the invasive DNA. The PAM sequence is necessary for both the adaptation and interference processes, but the degree to which it is recognised differs for each stage. This is so because the steps of adaptation and interference use various molecular processes.

  PAM elements have been separated into target interference motifs (TIMs) and spacer acquisition motifs (SAMs) to accommodate this variation. SAM element is functional motif connected to the protospacer sequence before the excision that is the spacer acquisition machinery.The in-depth knowledge about CRISPR Cas-9 and gene editing is solely based on the understanding of Pam regions.

- The CRISPR-Cas9 system exploits pre-CRISPR RNA (pre-crRNA) which is a precursor of CRISPR RNA (crRNA), a short RNA molecule that targets and cleaves DNA. In the process of transcription of CRISPR and Cas genes ,**Pre-CRISPR RNA** and **Cas mRNA** are generated respectively. The mature crRNA is subsequently created by processing the **pre-crRNA**. The protospacer sequence is complementary to the short RNA molecule known as the mature **crRNA.** From the matured crRNA a ribonucleoprotein complex is formed and the translated Cas protein functions in cleaving the target viral DNA in a specific sequence.

- **Trans-activating CRISPR RNA (tracrRNA**) is the distict part of the CRISPR/Cas9 system. These are short RNA molecule has a major role in directing the Cas9 protein in the target DNA sequence. The tracrRNA actually "trans-activates" the CRISPR/Cas9 system in developing the concept, which is if it is present in the same cell as the crRNA and Cas9 protein, it can activate the CRISPR-Cas9 system. These tracer RNAs have complementary repeats in them to regions in the CRISPR RNA.

- An enzyme RNAase comes along then simultaneously and cuts through those repeated regions. This action results in producing pieces of RNAs made up of spacers, repeats and tracer RNA. A structure formation takes place within the cell containing

  → one piece of spacer RNA and repeat RNA- which acts as a single polymer of nucleotide and
  → tracrRNA – another polymer of RNA nucleotide held to the previous polymer by hydrogen bonds.

  This structure is called **cr:tracrRNA.** This molecule is then get picked up by an enzyme called Cas-9. After binding of Cas-9, cr:tracrRNA is called **gRNA** or **guide RNA**. The Cas proteins then cut the foreign DNA at a specific location.

Cas9

Cas-9 is one of the protein in Cas protein family of 160 kilodalton (CRISPR-associated protein 9, formerly Cas5, Csn1, or Csx12) is essential for certain bacteria's immune defence against DNA viruses and plasmids. Cas9 is an RNA dependent DNA endonuclease thereby having the ability to cut DNA. A guide RNA (gRNA), a kind of RNA, directs Cas9 to a certain DNA sequence. The HNH (Hen1-NPH2) domain of the Cas9 protein allows it to recognise the PAM region. The Cas9 protein functions in cleavage of the target DNA sequence in the HNH domain, which binds to the PAM region. Cas9 breaks the DNA at a specific region known as the protospacer adjacent motif (PAM) when it binds to the gRNA.

○ STRUCTURAL CONFIGURATION OF CAS-9

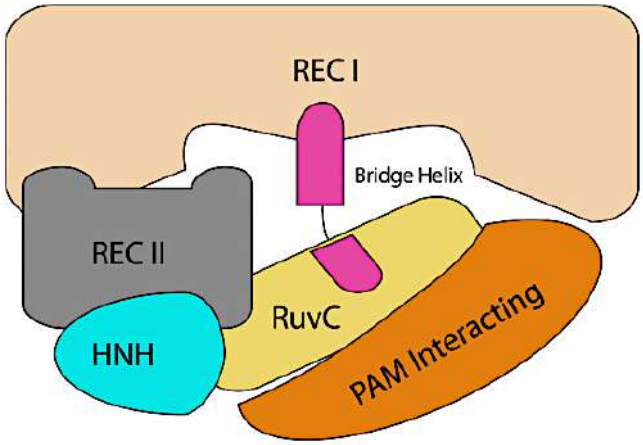
A complex of Cas protein Cas9-sgRNA-DNA ternary complex were present in the crystallographic asymmetric unit. The sgRNA and the DNA are recognised by Cas9

protein in comparable ways despite the two complexes having different conformations. The HNH domain in Mol B is flexible, in contrast to Mol A where it is coupled to the RuvC domain via a disordered linker. Therefore it is important to describe the structural characteristics of Mol A before talking about the structural variations between the two complexes, which is responsible for the conformational flexibility of Cas9.

Cas9 has two lobes: a **recognition (REC) lobe** and a **nuclease (NUC) lobe**, as shown by the crystal structure

REC Lobe: The CRISPR-Cas9 protein has a domain called the REC lobe that is in charge of binding to the guide RNA. The REC lobe is classified into three regions:

→ the REC1 region or domain
→ the REC2 domain, and
→ Bridge helix, which is a lengthy helix



A new study illustrateded that **REC lobe** has no structural similarities to any other well-studied proteins. This proves that it has a functional domain that unique to Cas9 gene. REC lobe is one of the minimum preserved or conserved sectionss in type II CRISPR system's three Cas9 families (IIA, IIB, and IIC). It is found that various Cas9 orthologs also have notably shorter REC lobes.

Fig: Different regions of REC lobe

REC1 Domain: The Cas9 protein's REC1 domain, which is the biggest domain, is in charge of binding to the guide RNA. Through a sequence of interactions, the various amino acids of this protein and the guide RNA's nucleotides accomplish this.Cas9 protein and the gRNA are positioned for DNA cleavage by the REC1 domain. Various positively charged amino acid residues in the REC1 domain specifically bind with the negatively charged phosphate backbone of the gRNA. The hydrophobic bases of the gRNA interacts with by a variety of the hydrophobic residues that are present in it. Cas9 protein and the gRNA combine to create a stable complex due to these interactions.

REC2 Domain: The CRISPR-Cas9 protein has a little domain called the REC2 domain, which is found in the REC lobe. The interaction between Cas9 protein and the gRNA is assumed to be stabilised by the REC2 domain. It accomplishes this by interacting with the bridge helix and the REC1 domain. It is also believed that the regulation of Cas9 activity involves the REC2 domain. Studies have demonstrated that changes to the REC2 domain can cause Cas9 activity to decline. Although the REC2 domain's precise purpose is not yet precisely known, it is assumed to be one of the crucial components of the CRISPR-Cas9 system. The following are some potential REC2 domain uses:

→ preventing the Cas9 enzyme and guide RNA from forming an unstable complex
→ controlling of Cas9 activity
→ interacting with other CRISPR/Cas-9 system proteins

Bridge helix (BH): The CRISPR-Cas9 protein has bridge helix (BH) which is a short, arginine-rich helix that joins the REC and NUC lobe. It is believed to be crucial for directing the Cas9 enzyme and the guide RNA where to cleave the DNA. Six positively charged amino acid residues called arginine residues make up the BH. These arginine residues bind with the guide RNA's negatively charged phosphate backbone, stabilising the Cas9 protein and guide RNA combination. As it connects REC and NUC lobes, direct and indirect interaction occurs with target DNA, tracrRNA, and crRNA; BH plays a crucial part in the function.Cas9 activity can be reduced by creating mutations of the arginine residues in the BH. This is due to the fact that the BH is necessary for both stabilising Cas9 protein and gRNA complex and also for placing the Cas9 protein and guide RNA appropriately for DNA cleavage.

NUC Lobe: C-terminal domain of CRISPR/Cas-9 protein is also known as NUC lobe. NUC lobe contains various domains, like HNH, RuvC and PAM-interacting (PI) domains. A positively charged groove accommodates negatively charged sgRNA: target DNA heteroduplex at the junctional position of REC and NUC lobes. The RuvC domain contains three separate RuvC motifs, designated as RuvC I–III that are put together in the NUC lobe. In the NUC lobe their interaction occurs with the PI domain to create a positively charged surface; afterthat their interaction with the 3′ tail of the sgRNA happens. Only a few interactions exist between the HNH domain and rest of the protein, which is located between the RuvC II and III motifs.

The complementary DNA strand to the gRNA is split apart by HNH domain. The DNA strand which is not complementary with the gRNA is cut by RuvC domain. A short DNA sequence, called PAM sequence where PAM-interacting domain associates. It plays essential role in DNA cleavage mechanism of Cas9. NUC lobe is required for cleavage of the DNA by Cas-9 system. In case of binding to the PAM sequence, The PAM-interacting domain IS IN-CHARGE. On the other hand, cleaving the DNA strand is dependent on the RuvC nuclease domains and HNH domains.

PAM Interacting Domain: The domain for CRISPR/Cas-9 protein binding to the PAM sequence is named as PID or PAM-interacting domain. BH also engage in interaction. The PAM sequence, short DNA sequence with which the PAM-interacting domain binds to, is necessary for Cas9 to cleave DNA. Cas9 protein and gRNA are positioned correctly for DNA cleavage thanks to the contact between the BH and the PAM-interacting domain.

Carboxyl-terminal of Cas9 protein is the place where PID is found. A number of conserved amino acid residues help to stabilise the hairpin shaped structure. The PAM sequence and these conserved amino acid residues work together to bind the DNA-binding PAM-interacting domain. The PID is crucial to the CRISPR-Cas9 system's specificity. This is so because each CRISPR-Cas9 system has a different PAM sequence. The PID makes sure that Cas9 only cleaves DNA that has the PAM sequence by attaching to it. A decrease in Cas9 specificity can result from PID mutations. This is because binding to the PAM sequence requires the PID. Cas9 may cleave DNA that lacks the PAM sequence if the PID is altered because it may no longer be able to connect to the PAM sequence.

HNH Domain: CRISPR/Cas-9 protein has a nuclease domain called the HNH domain, which aids in the cleavage of complementary DNA strand to the gRNA. It can be detected in the Cas9's C-terminus. There are two subdomains: HNH1 and HNH2 subdomains.

→ HNH1 Subdomain: The smallest of the two HNH domain subdomains is the HNH1 subdomain. It is in charge of attaching to the guide RNA's complementary DNA strand. It accomplishes this through several interactions between its amino acid residues and the DNA strand's nucleotides. A number of conserved amino residues help to stabilise the -hairpin structure of the HNH1 subdomain. There are 80 amino acid residues in it. Ten hydrogen bonds help to stabilise its -hairpin structure. The DNA strand and these conserved amino acid residues interact for attaching HNH1 subdomain to the DNA. The CRISPR-Cas9 system cannot modify DNA without the HNH1 subdomain. Cas9 may become less active or may cleave the incorrect DNA strand as a result of mutations in the HNH1 subdomain.
The DNA strand is depicted in green, while the HNH1 subdomain is displayed in blue. The DNA strand's nucleotides are depicted in black, while the HNH1 subdomain's conserved amino acid residues are highlighted in red.

→ HNH2 subdomain: Of the two subdomains of the HNH domain, the HNH2 subdomain is the larger. It is in charge of cleaving the guide RNA's corresponding DNA strand. By first creating a phosphodiester link with the DNA strand and then hydrolysing that bond, it accomplishes this. Numerous amino acid residues make up the HNH2 subdomain and are required for its proper operation. The pockets formed by these amino acid residues can hold the DNA strand. The DNA backbone can be deprotonated by the acidic residues around the pocket, which makes it simpler to rupture the link.

RuvC Domain: Numerous proteins, including the CRISPR-Cas9 protein, have the RuvC domain, a nuclease domain. It is called after the bacterial protein RuvC, which is involved in DNA double-strand breaks repair activity. HNH domain and the RuvC domain share structural and functional similarities. Each domain is made up of a pocket that can hold a DNA strand and several acidic residues that can deprotonate the DNA backbone. The

domains can split the DNA strand more readily as a result. The CRISPR-Cas9 system can modify DNA without the need for the RuvC domain. It can, however, be used to cause double-stranded DNA breaks that are more challenging for cells to heal. Applications like gene editing and cancer therapy may benefit from this. The following information on the RuvC domain is provided:

→ It is comparable to the HNH domain in both structure and function.
→ It exceeds the HNH domain in size.
→ It has a unique pocket architecture.
→ Specific cleavage of the DNA strand which is not complementary to the gRNA

→ Ability to modify DNA is not required for CRISPR/Cas-9 system to function.
→ Also utilised to create more challenging double-stranded DNA breaks.

Chart of domains present in Cas-9 protein

Domain	Function
REC1	Binds to the gRNA
REC2	Unclear
Bridge helix	Positions Cas9 protein and gRNA for DNA cleavage
PAM-interacting domain	Binds to the PAM sequence
HNH domain	Cleaves DNA strand that is complementary with the gRNA
RuvC domain	Cleaves the DNA strand which is not complementary with the gRNA

sgRNA: Single guide RNA, i.e., sgRNA, as the name indicates is a single RNA molecule which associates scaffold tracrRNA sequence with crRNA short sequence. The tracrRNA contributes in stabilising the complex and attracting the Cas9 protein, while the gRNA binds to the DNA target region.

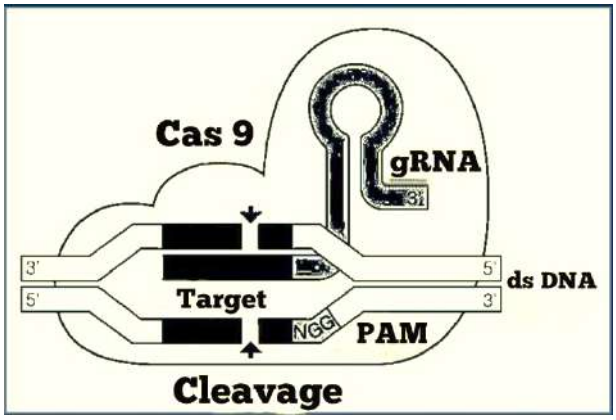


Fig: Diagram of sgRNA

The target DNA sequence and Cas9 protein both are taken into consideration when a computer programme is used to build the sgRNA. The best sgRNA sequence is chosen from within the many generated by the programme, based on factors including binding affinity,

stability, and effectiveness. According to a recent study, these algorithms are categorised into three groups: Regression models include gradient boosting regression trees and extreme gradient boosts (XGBoost), classification models that include both logistic regression, as well as support vector machines, and emerging technologies or hybrid algorithms include a simple average of multiple models and deep learning technology.

The development of effective sgRNAs is becoming a more urgent task as CRISPR/Cas-9 system has spread to become a common gene editing tool in the diverse field of biological research. To deal with this crucial issue, numerous bioinformatics tools have been created. However, a significant drawback of earlier research is to the calibre of the datasets utilised to develop such systems. The majority of experimental techniques are based on phenotypic screening, which makes it difficult to gauge the effectiveness of CRISPR/Cas9 editing. It is uncertain that effective gene editing will result in uniform and several phenotypic changes among all of the examined genes and their target sites. As a result, these indirect methods add unintentional distortion to the datasets used to train machine learning algorithms that may conceal the real properties that are unique to sgRNA-guided Cas9 cleavage. Additionally, sgRNAs studied in functional screens are often created for a limited number of genes and examined in a single cell line. These limitations may introduce biases that are particular to each experimental context, such as biases relating to varying degrees of genomic accessibility or varying reactions to DNA cleavage in a cell line- or gene-specific way. All of such elements might limit the generalizability of the model.

Classification of CRISPR-Cas system:

Two major classes of CRISPR-Cas systems was developed by Makarova and his team and these two major classes of CRISPR-Cas also divided into six primary kinds (I-VI), which are currently divided into approximately 30 subtypes . Class 1 and Class 2 systems are different from one another due to distinct protein-effector complex. Class 2 effector complexes 40 has a single Cas protein along with crRNA, occasionally joined by tracrRNA. Six Cas gene encodes several protein subunits and thr class I effector complex 39 also has crRNA. The first class is further split into type- I,type- III, and type-IV, while class two is divided into type- II, type-V, and type-VI.

- **Cas-I System:** Type-I CRISPR-Cas system is most common to present in sequenced genome and encode multi-subunit effector complexes(as effector complexes93) as opposed to single-effector type-II systems. Type I interference saw the first descriptions of numerous mechanistic findings on interference, including the existence of a seed sequence, the mechanism of R-loop creation, and inhibition by anti-CRISPRs. There are at least six different subtypes of type-I systems, which all contain the distinctive protein Cas3 and make use of the DNA binding complex Cascade. A mature crRNA is needed for the effector complex in type-I system, which is created during pre-crRNA processing. The repeat sequences are broken up into specified stem-loop structures by a designated ribonuclease (often Cas6).

One generalised description cannot effectively capture the variety of Type-I systems due to the most diverse CRISPR-Cas grouping. Each subtype has distinct properties in addition to the presence of a distinctive large subunit protein, such as the split Cas3 organisation, the replacement of Cas6 with Cas5 with respect to their functional role or the presence of their fusion , the fusion of Cas2 and Cas3, and the presence of Cas8 and Cas3. The degree of variation found in these complexes underlines the adaptability of CRISPR-Cas immune systems and illustrates the vast array of options that nature has investigated throughout the course of evolution.

- **Cas-II System:** Cas9 endonuclease, a protein originally known by the names as Cas5 and Csn1, is a component of type II systems is a familiar protein. In DNA targets that have complementarity to the 20-nucleotide crRNA guide sequence and a flanking protospacer-adjacent motif (PAM), Cas9 induces double-strand breaks (DSBs) under the guidance of a dual-RNA substrate made up of crRNA and tracrRNA. Because these sequences lack flanking PAMs, the PAM requirement ensures that Cas9 stays away from complementing targets within the CRISPR array itself. It's interesting to note that the acquisition of additional spacer sequences depends on Cas9-mediated PAM recognition.
  Despite the fact that type II CRISPR-Cas systems offer a strong defence against viral infection, phages can evade interference by employing a variety of evasion techniques. Effective targeting is prevented by spontaneous mutations in the PAM or protospacer89, and DNA modifications like glucosylation can reduce DNA recognition and/or increase mutation frequencies that allow escape.
- **Cas-III System:** The targeting of nucleic acids in type III CRISPR-Cas systems, which involves three different nuclease modes, is likely the most challenging. Years of research suggested that the interference mechanism was not consistent, a type III-A system reporting DNA targeting in vivo studies and a type III-B system showing RNA targeting in vitro studies .A recent study is showing that the additional nuclease activity is stimulated by secondary messangers.The multi-subunit effectors in type-III systems are classified into four categories and are called Csm or Cmr complexes.

  As RNA-guided RNA-targeting effectors, type III complexes recognise single-stranded target RNAs by base-pairing with the crRNA. But a conserved -hairpin found in Cas7 deliberately breaks up the resultant RNA duplex, causing a nucleotide to flip outward at 6-nucleotide intervals. Thus, these particular nucleotides are set up for cleavage, which is made possible by a conserved aspartate residue found in Cas7. Type III immune systems have been demonstrated to target escape mutants from other CRISPR systems, in addition to utilising a formidable arsenal of weaponry for stand-alone defence. Additionally, they distinguish between lysogenic and lytic infections, enabling the host to benefit from possible advantages of

lysogenic infection. These findings highlight the wonderful polishing influence of natural selection on the structure and operation of CRISPR-Cas systems.

- **Cas-IV System:** The recently added system in CRISPR/Cas system is Cas-IV system. They express known multi-subunit protein-crRNA complex components but frequently lack adaptability genes, a potential DNA nuclease, and even CRISPR arrangements. They are not studied thoroughly till now and their functional role is still unknown.

- **Cas-V System:** In 2015, Makarova et al. classified Type V systems as a possible grouping based on the signature gene cpf1, which had been found in multiple bacterial genetic material next to adaptation genes and a CRISPR arrangement. Shortly after, Zetsche et al. published experiments that clearly established this new classification within Class-2 systems, showing that Cpfl also known as Cas12a which works similarly to Cas9 as a single-effector, RNA-guided endonuclease that performs double-stranded DNA cleavage.

- **Cas-VI System:** Subtypes A-E of Type V systems, which encode Cas12a-e and were formerly known as Cpf1 (A), C2c1 (B), C2c3 (C), CasY (D), and CasX (E), have recently been added. Based on the discovery that they have flank CRISPR arrangement that frequently include spacers matching with phage genomes, these putative RNA-guided effectors are projected to be active and exhibit RuvC-like nuclease domains. Members of the Cas13 family consists of two HEPN domains with expected single-stranded RNase activity, and tests using heterologously produced Type VI systems in E. coli showed that these systems provide protection against RNA phages. Upon binding to the target RNA, Cas13 activates, causing multiple-turnover nonspecific RNA cleavage activity to be released in trans rather than within the targeted region itself. Although not always, targeting is aided by the presence of a certain protospacer flanking sequence, and depending on the Cas13 homolog, collateral, nonspecific RNA cleavage is both sensitive to secondary structure and has particular nucleotide preferences.

The collateral, nonspecific cleavage of other cellular RNAs inhibits bacterial growth and may have developed as a method to cause cell dormancy and/or programmed cell death. Cas13 members may provide particular protection against RNA phages when produced heterologously in E. coli. Additional research is necessary to comprehend the physiological effects of interference in Type VI systems, especially involving assays that test Cas13 function in native species.

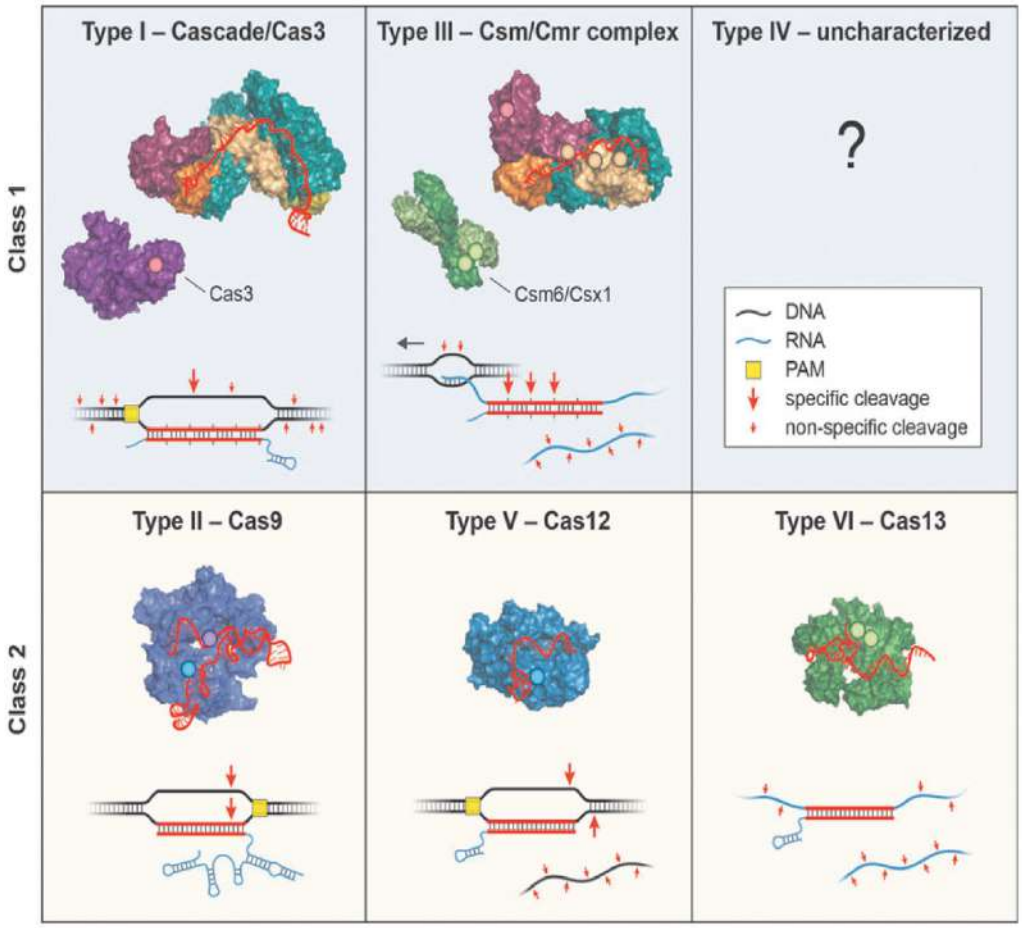


Fig: Different types of CRISPR-Cas system

Breaks repair mechanism: Cas9 may be designed to create DSBs at DNA target locations in conjunction with a natural dual-RNA guide or an artificial single-RNA guide49, leading to DNA repair outcomes that fall into two main classes: non homologous end joining (NHEJ) and homology-directed repair (HDR)185 (Fig. 3A). Despite not being random, repair results from NHEJ frequently lead to tiny insertions or deletions at the target site, which when introduced within exons frequently result in loss-of-function phenotypes. Combining Cas9-gRNA and a donor template enables access to precise genomic changes through HDR.

The potential impact of CRISPR Cas9 system

The CRISPR-Cas9 system is acclaimed to be a revolutionary tool in gene editing in the modern scientific era. Among them, the most benefitted fields are-

Advancements in Treatment Procedures and Medical Research:

Medical research and treatment are improving due to the CRISPR-Cas9 system's capacity to accurately modify DNA sequences. It offers the possibility of specialised therapy for genetic illnesses that were previously thought to be incurable. It may contribute to revolutionary medicines for illnesses including cystic fibrosis, sickle cell anaemia, and particular types of cancer by addressing disease-causing mutations at their source. The CRISPR/Cas9 technology offers a powerful and innovative way to address conditions like diabetes. There are many approaches used, including some cell-based therapies using stem cells and brown adipocytes, focusing on key genes involved in the aetiology of diabetes, and examining the difficulties and restrictions posed by this technology.

In a recent experiment, NBS-CRISPR, a sequence that interacts with NF-B, was recently tacked onto the CRISPR-Cas9 framework to create a system that causes inflammation. This newly created genome-editing tool's genetic scissor function is triggered in the presence of an inflammatory attack and deactivated or scaled back in non-inflammatory settings. Furthermore, by focusing on the MyD88 gene, a significant participant in the NF-B signalling system, they used this platform to correct inflammatory diseases and produced remarkable therapeutic results. Finally, P65 (RELA) can move from the cytoplasm to the nucleus during inflammation. They created an NBS-P65-CRISPR system which expresses the fusion protein Cas9-p65 to prevent Cas9 leaky DNA cleavage activity. This inflammation-inducible Cas9-mediated genome editing approach offers fresh perspectives and access points for investigating pathogenic genes.

→ **Application and progress of Cas-9 based on hepatocellular carcinoma :**

A common cause of cancer death is mostly from Hepatocellular carcinoma (HCC) despite no appreciable advances in survival rate over the last few years. Despite the fact that surgery, chemotherapy, and radiotherapy have been the main traditional therapeutic modalities for HCC patients and have all made noteworthy progress, there are still issues, such as medication resistance and toxicity. The recently developed gene therapy of clustered regularly interspaced short palindromic repeat/CRISPR-associated nuclease 9-based (CRISPR/Cas9) has attracted considerable attention as an alternative to conventional treatment methods for eradicating resistant malignant tumours and controlling numerous crucial events of target gene-editing. At the intersection of several scientific fields, including chemistry, materials science, cancer biology, and genetics, recent discoveries in CRISPR/Cas9-based anti-drugs are explored. The CRISPR/Cas9 technique's basic principles and working examples were presented in some reviews to demonstrate its viability. The applications of CRISPR/Cas9 technology in therapeutic HCC were also highlighted. The design, action mechanisms, and anticancer uses of non-viral delivery systems for CRISPR/Cas9-based anti-drugs in HCC treatment were also

summarised. Finally, the shortcomings and future directions of recent research for the development of anti-drugs were also examined in an effort to give an outline of theoretical foundation.

- **Customizing Drug Development**: The system's precision allows researchers to engineer cellular models with specific genetic traits, accelerating drug development. CRISPR-Cas9 facilitates the creation of more accurate disease models, enabling researchers to study disease mechanisms, test potential treatments, and evaluate drug safety more efficiently.

  Research on CRISPR-Cas9 and transcriptome analysis to develop p53 knockout U87MG cell line for an unbiased drug delivery testing system in 2021. In their study, they used the CRISPR-Cas9 technology to generate various variants of U87MG cell line with mutated TP53 gene. They found that the loss of function mutation in p53 gene is the primary cause of their significant transcriptional changes. This information made them predict that the mutant clones among all which will have less contrasting characteristics than wild type. Based on the transcriptome data they are the most suitable candidate for drug delivery system. Their expectations were to generate the ideal cell shape, proliferative rate and target antigen-mediated uptake. The experiment was a success in choosing the top qualifying mutant clone on the basis of the results of combined analysis. This work establishes in the methodology of viability and lays the door for expanding to new cell types and target genes.

- **Agriculture and Food Security**: The CRISPR/Cas9 system has the capability to grow crops with increased resistance to pests and diseases, enhanced nutritional value and longer shelf life. This method can be helpful in addressing challenges with food security globally and may lead to lowering the impact of agriculture on the environment while reducing the need of chemical pesticides and insecticides.

  A recent study shows the progression of genome editing based on CRISPR/Cas-9 for improving disease resistance properties in plants. They are using strategies like transgenic technology, and traditional breeding. In recent years production of new lines with improved pathogen resistance was done. But there have been some drawbacks, as during the transfer of desired resistance genes; some undesireable genes also gettransferred, the cost of labour and the inherent genetic variability in plant populations. It's interesting that recent breeding techniques have made it possible to specifically modify one or more genes in plants. In the case of viral illnesses, the CRISPR/Cas9 technology was successfully exploited to increase biotic resistance of plants via mutation of susceptibility (S) genes, detrimental regulators of the plant immune system, as well as pathogen genomes themselves.

- **Industrial and biotechnological applications**: CRISPR/Cas-9is a unique technology that broadens the field of biotechnology. It renders the ability to more effectively manipulate microbes to produce biofuels, medicines, and chemicals. This technology has the ability to foster environmentally friendly and sustainable solutions across numerous industries.

  Recent studies give an outline of the theories and most recent advances in CRISPR/Cas-9 tools for multiplex genome editing, precise alteration of endogenous genes, and total knockout of their expression. The CRISPR-Cas9 system's present and future uses for cell imaging, metabolite engineering, biocontrol, chromatin dynamics, various signalling cascades, and deciphering fungal disease are also emphasised. Their contribution in discussing the difficulties encountered in the design and use of the CRISPR/Cas-9 system is revolutionary in providing the problems of using these methods to study filamentous fungi.

- **Responsibilities and ethical considerations**: CRISPR/Cas9 system has limitation in the clinical field as the experiments faced many challenges. There are challenges involved with CRISPR/Cas-9 to deliver Cas-9 in multicellular organisms. It is crucial to pick the optimum delivery method for introducing the CRISPR system into cells in-vivo and to focus on the correct nucleus sequences.

  Important moral and societal issues are raised with the revolutionary potential of CRISPR-Cas9 system. Editing the human germline, for instance, raises concerns over the unexpected implications of genetic manipulation and its ethical limits. As this technology develops, it is crucial to ensure responsible use and take ethical issues into consideration.

Using Interference Complexes in the CRISPR-Cas9 system:

The majority of Cas protein tool development has been concentrated on making use of interference complexes' functional, DNA recognition based on genetic sequence. In particular, Cas9 from *Streptococcus pyogenes* has shown to be incredibly helpful for genome engineering. Cas9 is now a two-component system that has simple application tn the field of genome editing, transcriptional regulation etc. Cas9 has been studied from diverse cell types from different creatures like human T- cells and mice stem cells, cells of bacteria, plants and fungus. Other interference complexes can also be employed to modify the genome and have been used in the past.

Despite PAM sequences being less comformable in genome engineering than Cas9 due to it's multi-subunit structure, due to its great size and stable strong binding, it has been used in cascade of transcriptional silencing process in *E. coli*. Although neither Csm nor Cmr complexes have been employed in published work, both may probably be used for a variety

of cellular RNA manipulation applications. Out of the 16 Cpf1 homologs studied, two homologs has discovered with the ability of genome editing in human cell. The staggered cuts might turn out to favour different methods of DNA repair, and the specificity of alternative PAM sequence of Cpf1 might have valuable role in locating locations lacking an adequate PAM sequence for Cas9.

The future of CRISPR-Cas9:

Cas9 has major application in research area, but its clinical use has remained complicated. Even while doing programmed cuts has mostly become straightforward, it is still difficult to bias DNA repair in favour of more accurate Homology-directed repair (hdr) rather than error-prone non-homologous end joining. Another difficulty faced by researchers is to deliver the Cas9 in to a particular tissues of whole organisms as RNP or an viral vector or as an plasmid. Clinical studies might take place in a few years, and medicines might follow in a decade, as the field is still developing quickly. Although Cas9 has previously been extensively used in research settings, obstacles still stand in the way of its clinical application. Even while doing programmed cuts has mostly become straightforward, it is still difficult to bias DNA repair in favour of more accurate repair methods like Homology-directed repair (hdr) rather than error-prone non-homologous end joining. Another difficulty that needs to be resolved to allow for clinical applications is the delivery of Cas9 to specific tissues in complete organisms, whether it is via an RNP, a plasmid, a viral vector, or both. Clinical studies might take place in a few years, and medicines might follow in a decade, as the field is still developing quickly. Although Cas9 has previously been extensively used in research settings, obstacles still stand in the way of its clinical application. Even while doing programmed cuts has mostly become straightforward, it is still difficult to bias DNA repair in favour of homology-directed repair rather than non-homologous end joining. Another difficulty that needs to be resolved to allow for clinical applications is the delivery of Cas9 to specific tissues in complete organisms, whether it is via an RNP, a plasmid, a viral vector, or both. Clinical studies might take place in a few years, and medicines might follow in a decade, as the field is still developing quickly.

References:

1. Al-Attar, Sinan, Westra, Edze R., van der Oost, John and Brouns, Stan J.J.. "Clustered regularly interspaced short palindromic repeats (CRISPRs): the hallmark of an ingenious antiviral defence mechanism in prokaryotes" , vol. 392, no. 4, 2011, pp. 277-289.

2. Gleditzsch, D., Pausch, P., Müller-Esparza, H., Özcan, A., Guo, X., Bange, G., & Randau, L. (2019). PAM identification by CRISPR-Cas effector complexes: diversified mechanisms and structures. *RNA biology*, *16*(4), 504–517.

3. Boubakri, Hatem. "Recent Progress in CRISPR/Cas9-Based Genome Editing for Enhancing Plant Disease Resistance." *Gene* 866 (May 25, 2023): 147334.

4. "CRISPR and Pre-CRISPR RNA (CrRNAs) | Molecular Biology | JoVE." Accessed September 3, 2023

5. Nishimasu, Hiroshi, F. Ann Ran, Patrick D. Hsu, Silvana Konermann, Soraya I. Shehata, Naoshi Dohmae, Ryuichiro Ishitani, Feng Zhang, and Osamu Nureki. "Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA." *Cell* 156, no. 5 (February 27, 2014): 935–49.

6. Guo, Minghui, Kuan Ren, Yuwei Zhu, Ziyun Tang, Yuhang Wang, Bailing Zhang, and Zhiwei Huang. "Structural Insights into a High Fidelity Variant of SpCas9." *Cell Research* 29 (January 21, 2019.

7. Babu, Kesavan, Nadia Amrani, Wei Jiang, S.D. Yogesha, Richard Nguyen, Peter Z. Qin, and Rakhi Rajan. "Bridge Helix of Cas9 Modulates Target DNA Cleavage and Mismatch Tolerance." *Biochemistry* 58, no. 14 (April 9, 2019): 1905–17.

8. Babu, Kesavan, Nadia Amrani, Wei Jiang, S. D. Yogesha, Richard Nguyen, Peter Z. Qin, and Rakhi Rajan. "Bridge Helix of Cas9 Modulates Target DNA Cleavage and Mismatch Tolerance." *Biochemistry* 58, no. 14 (April 9, 2019): 1905–17.

9. Anders, Carolin, Ole Niewoehner, Alessia Duerst, and Martin Jinek. "Structural Basis of PAM-Dependent Target DNA Recognition by the Cas9 Endonuclease." *Nature* 513, no. 7519 (September 25, 2014): 569–73.

10. Zuo, Zhicheng, and Jin Liu. "Structure and Dynamics of Cas9 HNH Domain Catalytic State." *Scientific Reports* 7 (December 8, 2017): 17271.

11. Glasgow, Anum, Helen T. Hobbs, Zion R. Perry, Malcolm L. Wells, Susan Marqusee, and Tanja Kortemme. "Ligand-Specific Changes in Conformational Flexibility Mediate Long-Range Allostery in the Lac Repressor." *Nature Communications* 14, no. 1 (March 2, 2023): 1179.
12. Synthego. "The Complete Guide to Understanding CRISPR SgRNA." Accessed September 3, 2023.

13. "Cas9." In *Wikipedia*, August 25, 2023.

14. Yuan, Tingting, Honglin Tang, Xiaojie Xu, Jingjing Shao, Gaojun Wu, Young-Chang Cho, Yuan Ping, and Guang Liang. “Inflammation Conditional Genome Editing Mediated by the CRISPR-Cas9 System.” *IScience* 26, no. 6 (June 16, 2023): 106872.

15. Cheng, Yan, Haiyang Wang, and Mo Li. “The Promise of CRISPR/Cas9 Technology in Diabetes Mellitus Therapy: How Gene Editing Is Revolutionizing Diabetes Research and Treatment.” *Journal of Diabetes and Its Complications* 37, no. 8 (August 1, 2023): 108524.

16. Kwon, Andrew Tae-Jun, Kohta Mohri, Satoshi Takizawa, Takahiro Arakawa, Maiko Takahashi, Bogumil Kaczkowski, Masaaki Furuno, et al. “Development of P53 Knockout U87MG Cell Line for Unbiased Drug Delivery Testing System Using CRISPR-Cas9 and Transcriptomic Analysis.” *Journal of Biotechnology* 332 (May 20, 2021): 72–82.

17. Yu, Shijing, Ruirui Zhao, Bingchen Zhang, Chunmei Lai, Linyan Li, Jiangwen Shen, Xiarong Tan, and Jingwei Shao. “Research Progress and Application of the CRISPR/Cas9 Gene-Editing Technology Based on Hepatocellular Carcinoma.” *Asian Journal of Pharmaceutical Sciences* 18, no. 4 (July 1, 2023): 100828.

18. Klompe, Sanne E., and Samuel H. Sternberg. “Harnessing ‘A Billion Years of Experimentation’: The Ongoing Exploration and Exploitation of CRISPR–Cas Immune Systems.” *The CRISPR Journal* 1, no. 2 (April 1, 2018): 141–58.