IOSR Journal Of Pharmacy
(e)-ISSN: 2250-3013, (p)-ISSN: 2319-4219
Volume 9, Issue 9 Series. I (September 2019), PP. 01-22 *www.iosrphr.org*

Review on CRISPR/Cas-9 Gene Editing Technology and Its Potential Clinical Application

Md. Abul Hassan[1], Nafisa Yesmin[2]
Department of Pharmacy, Faculty of Life Sciences, Daffodil International University, Dhaka, Bangladesh
Department of Pharmaceutical Sciences, School of Health and Life Sciences, North South University, Dhaka
Corresponding author: Md. Abul Hassan
Received 17 August 2019; Accepted 03 September 2019

Abstract: The clustered regularly interspaced short palindromic repeats (CRISPR/Cas9) system is a recently discovered powerful and novel technique for genome editing that has led to a new era in molecular biology due to its huge applications in deleting cancer cell lines to genetically modified plant. CRISPR/Cas system is a unique adaptive immunity of 40% bacteria and 90% archea to combat infection against plasmid and virus. It memorizes previous infections by integrating short sequences of invading genomes—termed spacers—into the CRISPR locus. The spacers interspaced with repeats are expressed as small guide CRISPR RNAs (sgRNAs) that are employed by Cas proteins to target invaders sequence-specifically upon a reoccurring infection and efficiently produce breaks in DNA double strands. Over the past 30 years, CRISPR has evolved from the 'curious sequences of unknown biological function' into a promising genome editing tool. The ability of the minimal CRISPR-Cas9 system to target DNA sequences using programmable RNAs has opened new avenues in genome editing in a broad range of cells and organisms with high potential in Cancer Immunotherapy, Tissue Regeneration , Gene Therapy, Malaria and Insect-borne Diseases, HIV and viral diseases , DMD, obesity and metabolism ,sickle cell anemia ,cancer research etc. This genome editing tool has gained its popularity because of its user friendly, easy, cheap, efficient and specific nature in editing gene.Despite the benefits of CRISPR/Cas9, it has several aspects that affect its safety, specificity including Cas9 activity, target site selection and short guide RNA design, delivery methods, and off-target effect etc. Since this versatile and relatively facile technique provides more benefits in comparison to other gene editing technologies like meganucleases, ZFN and TALEN so has and will be tremendously impactful in research areas. This review discusses the mechanism of CRISPR, advantages, disadvantages and its applications in gene editing as one of the most important scientific discoveries of the 21st century.

Keywords: CRISPR/Cas9, ZFN, TALENs

I. INTRODUCTION

The first known human genetic disease was identified as alkaptonuria. Today, we recognize that there are at least 8000 human diseases that are caused by mutations in single genes and the number increases almost every day [1-2] .Genetic diseases may affect over 400 million people worldwide. For a few number of patients, solid organ transplantation or allogeneic hematopoietic stem cell transplantation can be used to treat their genetic disease, but for the vast majority of patients there is no cure and at best they are treated by management of symptoms. Therapeutic genome editing is being developed not only to cure monogenic diseases but also to cure more common diseases that have multiple origins. It is a delicate method that can be used to cure diseases that have both a genetic and environmental component and infection. The exact nature of the therapeutic edit has to be driven by a solid understanding of the interaction between the underlying genetics and the specific pathophysiology of the disease. That is, one editing strategy might be appropriate for one disease but not applicable to another .Genome editing, also previously known as gene targeting, has been a powerful research tool for scientists. In particular, the ease of gene targeting in yeast was one factor that made yeast such an important model organism in studies of the pathophysiology of human disease [3-4] A critical breakthrough of genome editing was the discovery that by creating a site-specific DNA double-stranded break (DSB) in the target gene can stimulate genome editing by mutations such as small insertions/deletions at the site of the DSB and repaired endogenously by homologous recombination(HR) or nonhomologous end-joining (NHEJ) [5-9].

When the editing of a single break occurs by NHEJ, insertions/deletions are created at the site of the break (Fig. 1a) ,the size of deletions tends to be larger than that of insertions, except when extrachromosomal DNA is captured at the site of the break [10-11] When editing of a single break occurs by HR using a provided donor sequence, precise nucleotide changes in the genome range from a single base insertion to the introduction of large cassette of genes When the editing of two breaks occurs by NHEJ, chromosomal deletions, inversions

or translocations can be created (Fig. 1b) .These gross chromosomal rearrangements can be generated intentionally for therapeutic purposes, but they also must be evaluated because any nuclease platform has the potential to produce off-target effects.

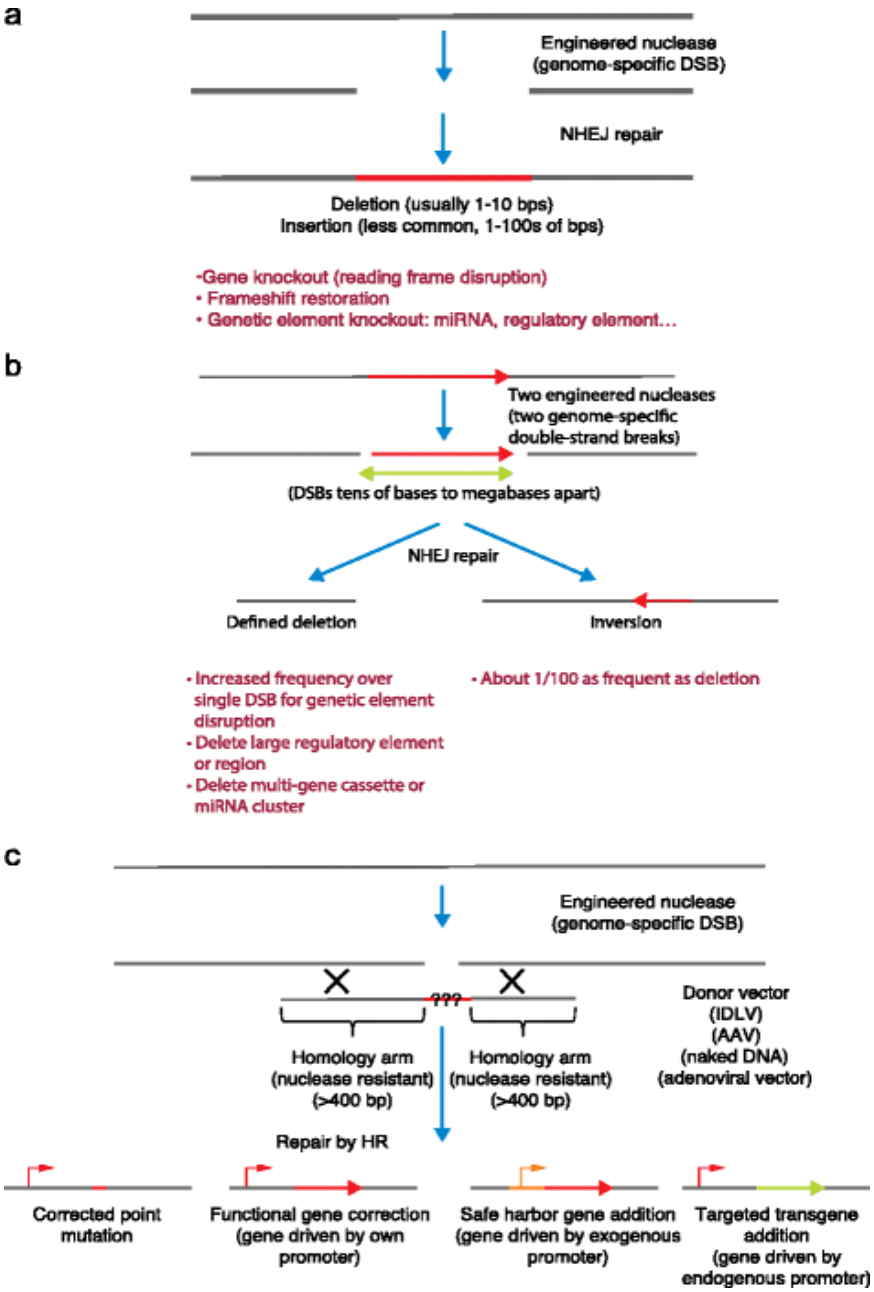


Fig Nuclease-based genome editing creates a specific double-stranded break (*DSB*) in the genome and then allows the cell's own endogenous repair machinery to repair the break.

There are four basic engineered nuclease platforms that include engineered meganucleases, ZFNs, TALENs, CRISPR/Cas9 nucleases, (Table 1).

Engineered meganucleases

These are derived from the large family of natural homing endonucleases (hundreds of members) .A small number of these endonucleases have been designed to recognize natural target sites in the genome using a variety of strategies, including structure-based design and yeast surface display [12] Natural meganucleases have historically been the gold standard for specificity, but the challenge of engineering meganucleases for novel target sites has limited their translational development. Furthermore, the specificity of engineered meganucleases has not been fully evaluated. The first solution to this problem was the development of zinc-finger nucleases (ZFNs) (originally called 'chimeric restriction enzymes', then called 'chimeric nucleases') [13]

1.1 Zinc-finger nucleases (ZFNs)

These are artificial proteins in which a zinc-finger DNA-binding domain is fused to the nonspecific nuclease domain from FokI [14]As the nuclease domain needs to dimerize in order to cut DNA efficiently, a pair of ZFNs needs to be engineered for each target site and these must be oriented correctly to allow dimerization. Engineering ZFNs that have high activity and high specificity to endogenous target sites remains challenging, although ZFN design strategies are continually improving. These ZFNs have entered clinical trials in which the engineered T cells have been shown to be safe [15]

1.2 Transcription activating like effector nucleases (TALANS)

These are also artificial proteins. They share a similar structure to ZFNs in which an engineered DNA-binding domain is fused to the nuclease domain from FokI [16]In TALENs, the DNA-binding domain is engineered by assembling a series of TAL repeats, with each repeat mediating interaction with a single base through a two-amino acid repeat variable di-residue (RVD) that can be described by a simple code [17]Thus, creating a highly active TALEN is much simpler than creating a highly active ZFN and simply involves using the code to assemble the correct TAL repeats needed to recognize a novel target sequence. In addition to the TAL repeats using natural RVDs, TAL repeats using engineered RVDs are now being used to create a TALEN [18]These engineered RVDs might have increased specificity over natural RVDs, although that remains to be further studied. TALENs that use TAL repeats containing natural RVDs may have better specificity than ZFNs.

1.3 CRISPR/Cas9 nucleases (Clustered Regularly Interspaced Short Palindromic Repeats)

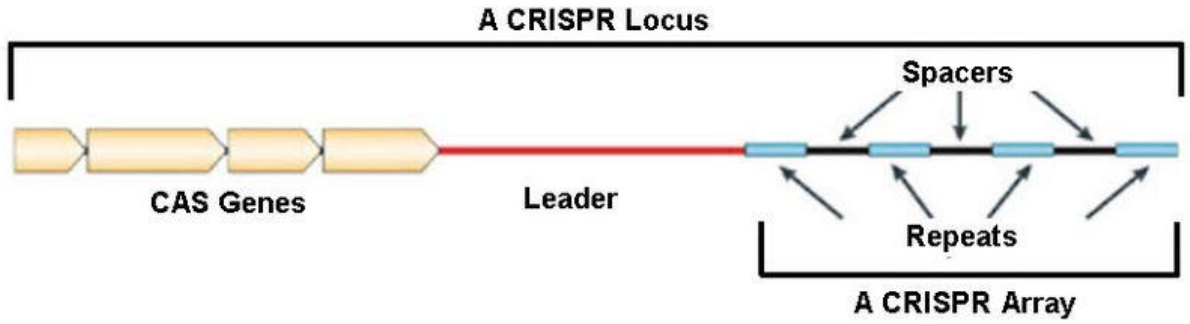


Figure : CRISPR/Cas9

It is derived from a bacteria andarchaea -based adaptive immune system that targets foreign DNA from viruses and plasmids[19]In contrast to the other three platforms, the CRISPR/Cas9 nuclease system does not derive specificity through protein–DNA interaction but instead through RNA–DNA Watson-Crick base pairing.CRISPR/Cas9 is a rapid and easy to use gene editing technology that can selectively delete, modify or correct a disease-causing abnormality in a specific DNA segment. Heterologous CRISPR systems were initially utilized as a genome editing tool to modify genomic architecture and accordingly function [20].

However, CRISPR can also be utilized in functional assays to examine transcriptional regulation and other biological processes [21-27]

The utilization of CRISPR as a versatile tool has revolutionized molecular biology and driven it into a new era. Cas9 is a CRISPR-associated endonuclease (an enzyme), the "molecular scissors" that are easily programmed to cut and edit, or correct, disease-associated DNA in a patient's cell. The location at which the Cas9 molecular scissors cut the DNA to be edited is specified by guide RNA, which is comprised of a crRNA component and a tracrRNA component, either individually or combined together as a 'single guide RNA' (sgRNA).It is designed such that the 20-bp recognition region of the sgRNA is identical to the desired target site (for Cas9, this 20-bp sequence is derived from Streptococcus pyogenes). The target site must be adjacent to a proto-spacer adjacent motif (PAM) sequence, which the Cas9 protein uses to identify target sites [28]For example, a guide RNA can direct the molecular scissors to cut the DNA at the exact site of the mutation present in the genome of patients with a particular genetic disease. Once the molecular scissors make a cut in the DNA, additional cellular mechanisms and exogenously added DNA will use the cell's own machinery and other elements to specifically 'repair' the cut DNA. The multifunctional Cas9 protein, in complex with the sgRNA, is able to unwind double-stranded DNA, then to create a blunt DSB if there is sufficient identity. Thus, CRISPR/Cas9 nucleases can be engineered very easily and between a third and a half of designed nucleases seem to be active at their desired target site. In addition to the four basic platforms described above, other nucleases have been engineered to recognize therapeutically relevant human target sites.

Table 1 Contrasting characteristics of the four standard nuclease platforms

Nuclease	Target site length	Mechanism of recognition	First use in human cells	Ease of design	Number of components	Size of mRNA transcript
Engineered meganuclease	>18 bp	Protein-DNA	1994 (I-SceI)	Extremely difficult	1	Short
Zinc-finger nuclease	18–36 bp	Protein-DNA	2003	Difficult	2	Short
TAL effector nuclease	24–40 bp	Protein-DNA	2011	Easy	2	Long
CRISPR/Cas9 nuclease	19–22 bp (*Streptococcus pyogenes* Cas9)	RNA-DNA Watson-Crick base-pairing	2013	Simple	1 (if using a complex guide RNA with Cas9 protein) or 2(if delivering guide RNA and Cas9 separately)	Long

There are subtle differences between each of these nuclease platforms—for example, the type of break that is created is different: meganucleases and mega-TALs generate breaks with 3′ overhangs; ZFNs create breaks with 5′ overhangs; TALENs create breaks that are variable in position that are usually (but not always) 5′ overhangs, as determined by the properties of the FokI nuclease (Fn); and CRISPR/Cas9 nucleases create blunt breaks. In general, however, each of these platforms mediates their editing effects through the creation of a DSB and thus they share a fundamental mechanism of action.

II. EVOLUTION OF CRISPR

1989 HR-mediated targeting
First study describing genome editing via HR in mouse ES cells

1992 Cre-lox The Cre-lox editing technology was successfully used for site-specific recombination in mice.

1998 Zinc-finger nucleases (ZFNs) Discovery of zinc-finger proteins that can target specific DNA sequences.

2000 Bacterial CRISPR/Cas
The CRISPR defense system is first identified in prokaryotes.

2009 Transcription-like effector nucleases (TALENs)
DNA binding proteins discovered in Xanthomonas bacteria.

2013 CRISPR/Cas genome editing
First demonstration that the CRISPR/Cas system can be used for mammalian cell genome editing.

III. THE DISCOVERY AND BACKGROUNDOF CRISPR/CAS9

In 1987 clusters of repeated DNA with dyad symmetry were noticed in Japan by Atsuo Nakata [29]. Nakata was examining the IAP gene in Gram negative Escherichia coli (E. coli) when unique repetitive sequences were uncovered on the 3’ end of the gene. Nakata identified a consensus sequence: 5’CGGTTTATCCCCGCT –GG -or- AA- CGCGGGGAACTC-3’ that was repeated five times. In between each repeat was a distinct spacer that was non-repetitive and about 31 nucleotides long. Finding repetitive sequences within a genome was not at all surprising as almost all prokaryotic or eukaryotic genomes previously examined have had repeat sequences [30-31]. However, the newly found repeats stood out as being separated by non-repetitive sequences or spacers. This arrangement of DNA sequences later became known as CRISPR, clustered regularly interspaced short palindromic repeats1. Not limited to E. coli, CRISPR was described in gram positive bacteria in 1991 and archaea in 1993 By 2002, CRISPR elements were found in all archaeal genomes examined and about 40% of bacterial genomes. Furthermore, associated genes were universally located in the proximity of CRISPR loci; these associated genes were termed CRISPR associated sequences or Cas genes in 2002 and

displayed motifs characteristic of both helicases and endonucleases [32]. The location of the Cas genes hinted at their potential regulatory mechanisms [32]. By 2005 it was recognized that the spacer (non-repetitive sequences) were very highly homologous to viral and plasmid DNAs (invading DNAs for bacteria). Then, in 2007 Horvath's seminal work showed the incorporation of new spacers into a CRISPR-Cas locus of Streptococcus thermophilus after challenging the bacterium with a bacteriophage. The newly acquired spacers always showed perfect complementarity to sequences on the phage genome and conveyed resistance towards that particular phage upon a subsequent infection [32]. This work linked the spacer DNA to adaptive bacterial immunity and began to delineate the specific roles of Cas proteins in this process. Research interest of the CRISPR field soon accelerated, leading to new discoveries that helped to understand the basic mechanisms of the immune system. In 2008, the processing of the CRISPR transcript into mature crRNAs that guide the Cascade complex of the E. coli type I-E system was experimentally validated, also giving hints that DNA rather than RNA is targeted [33].

The latter was confirmed in the same year as a study demonstrated that indeed DNA is the targeted molecule. This led scientists to think about the potential role that this prokaryotic immune system might play as a DNA manipulation tool. Today, CRISPR-Cas9 is a frequently harnessed tool for genome editing purposes and major progress in understanding the underlying biochemical processes in RNA-guided Cas9 was presented in recent years. In 2010, researchers showed that Cas9 creates a single double-stranded break at a precise position on the target DNA. Further insight into the mechanism was delivered 1 year later as the involvement of another small RNA, called tracrRNA, was shown. The maturation of crRNA requires tracrRNA as well as Cas9 and RNase III. Evidence that the system would function heterologously in other bacteria was demonstrated in 2011, as the S. thermophilus type II CRISPRCas system could provide immunity in *E. coli.* Other research had shown certain elements of the type II system, including the involvement of a PAM sequence in interference but the nature of the cleavage complex remained unknown. In 2012, tracrRNA, which was previously known to be involved in crRNA maturation was shown to also form an essential part of the DNA cleavage complex, with the dual tracrRNA:crRNA directing Cas9 to introduce doublestrand breaks in the target DNA [33]. Further simplification of the programmed targeting was achieved by creating a single-guide RNA fusion of tracrRNA and crRNA that guides Cas9 for sequence-specific DNA cleavage [33]. A few months following the description of the CRISPR-Cas9 technology [33], a number of publications demonstrated its power to edit genomes in eukaryotic cells and organisms, including human and mouse cells [34-36].

IV. CLASSIFICATION OF CRISPR

The Cas proteins are a highly diverse group. Many are predicted or identified to interact with nucleic acids; *e.g.* as nucleases, helicases and RNA-binding proteins. The Cas1 and Cas2 proteins are involved in adaptation and are virtually universal for CRISPR-Cas systems. Other Cas proteins are only associated with certain types of CRISPR-Cas systems. The diversity of Cas proteins, presence of multiple CRISPR loci and frequent horizontal transfer of CRISPR-Cas systems make classification a complex task.

The most adopted classification identifies Type I, II and III CRISPR-Cas systems**,** with each having several subgroups [37]. Type II is the basis for current genome engineering applications which is from *Streptococcus pyogenes* which is orthogonal to the native E. coli system. Different types of CRISPR-Cas systems can co-exist in a single organism. Recently, a Type IV system was proposed, which contain several Cascade genes but no CRISPR, *cas1* or *cas2*. Type IV complex would be guided by protein-DNA interaction, not by crRNA, and constitutes an innate immune system preset to attack certain sequences.

The Type I systems are defined by the presence of the signature protein Cas3, a protein with both helicase and DNase domains responsible for degrading the target [38]. Currently, six subtypes of the Type I system are identified (Type I-A through Type I-F) that have a variable number of *cas* genes. Apart from *cas1*, *cas2* and *cas3*, all Type I systems encode a Cascade-like complex. Cascade binds crRNA and locates the target, and most variants are also responsible for processing the crRNA. Cascade also enhances spacer acquisition in some cases. In the Type I-A system, Cas3 is a part of the Cascade complex.

The Type II CRISPR-Cas systems encode Cas1 and Cas2, the Cas9 signature protein and sometimes a fourth protein (Csn2 or Cas4). Type II systems have been divided into subtypes II-A and II-B but recently a third, II-C, has been suggested. The *csn2*and *cas4* genes, both encoding proteins involved in adaptation [39-43, are present in Type II-A and the Type II-B, respectively, while Type II-C lacks a fourth gene.

Three types of CRISPR mechanisms have been identified, of which type II the most is studied. In this case, invading DNA from viruses or plasmids is cut into small fragments and incorporated into a CRISPR locus amidst a series of short repeats (around 20 bps). The loci are transcribed, and transcripts are then processed to generate small RNAs (crRNA – CRISPR RNA), which are used to guide effector endonucleases that target invading DNA based on sequence complementarity [16–20]

One Cas protein, Cas9 (also known as Csn1), is a key player in certain CRISPR mechanisms. The type II CRISPR mechanism is unique compared to other CRISPR systems, as only one Cas protein (Cas9) is required for gene silencing In type II systems, Cas9 participates in the processing of crRNAs and is responsible for the

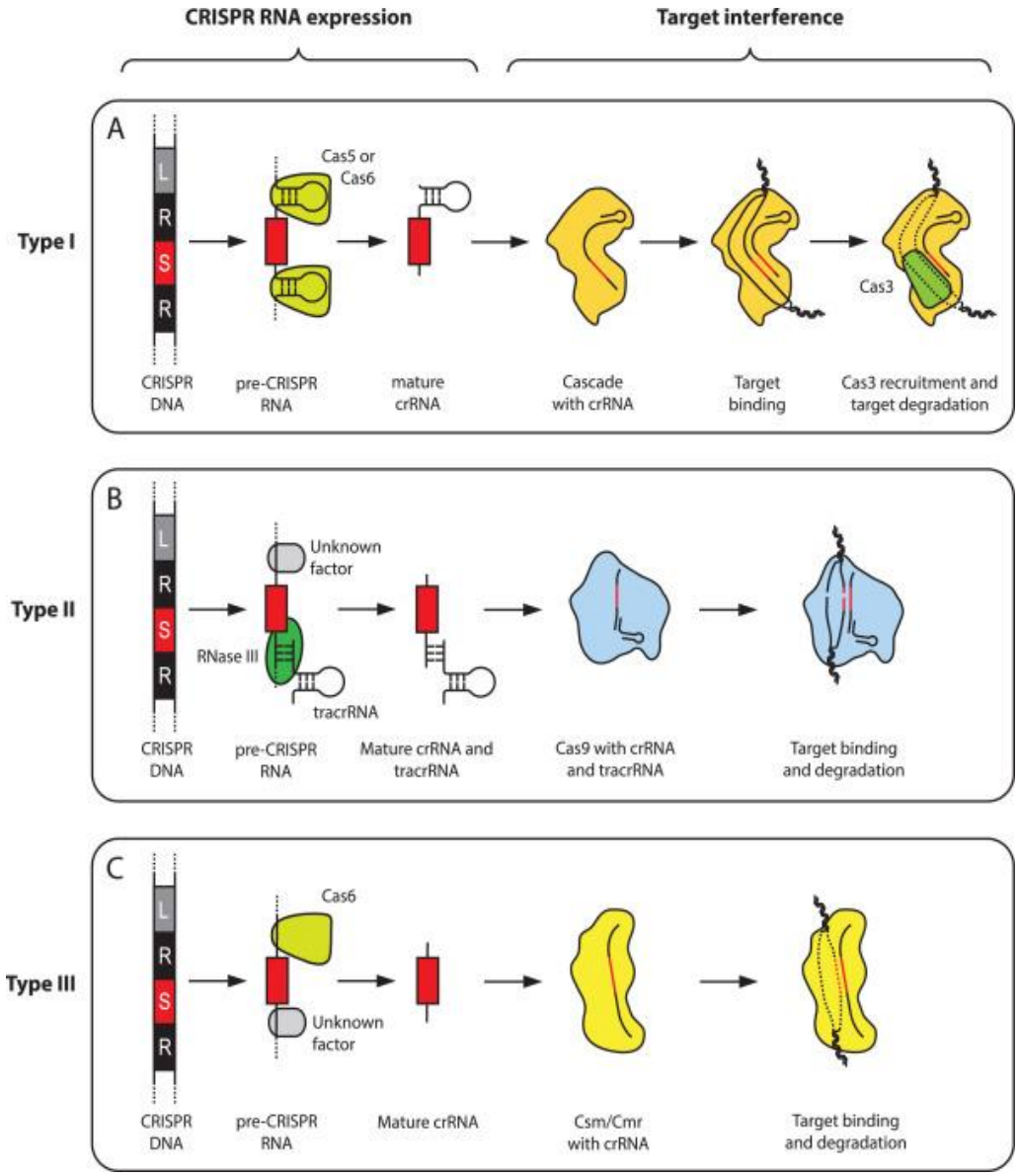
destruction of the target DNA (11). Cas9's function in both of these steps relies on the presence of two nuclease domains, a RuvC-like nuclease domain located at the amino terminus and a HNH-like nuclease domain that resides in the mid-region of the protein

To achieve site-specific DNA recognition and cleavage, Cas9 must be complexed with both a crRNA and a separate trans-activating crRNA (tracrRNA or trRNA), that is partially complementary to the crRNA .The tracrRNA is required for crRNA maturation from a primary transcript encoding multiple pre-crRNAs. This occurs in the presence of RNase III and Cas9

During the destruction of target DNA, the HNH and RuvC-like nuclease domains cut both DNA strands, generating double-stranded breaks (DSBs) at sites defined by a 20-nucleotide target sequence within an associated crRNA transcript .The HNH domain cleaves the complementary strand, while the RuvC domain cleaves the noncomplementary strand.

The double-stranded endonuclease activity of Cas9 also requires that a short conserved sequence, (2–5 nts) known as protospacerassociated motif (PAM), follows immediately 3´- of the crRNA complementary sequence In fact, even fully complementary sequences are ignored by Cas9-RNA in the absence of a PAM sequence [44]

The Type III CRISPR-Cas systems contain the signature protein Cas10 with unclear function. Most Cas proteins are destined for the Csm (in Type III-A) or Cmr (in Type III-B) complexes, which are similar to Cascade. Interestingly, while all Type I and II systems are known to target DNA, Type III systems target DNA and/or RNA So far, the Type II systems have been exclusively found in bacteria while the Type I and Type III systems occur both in bacteria and archaea [45].



V. MECHANISM OF CRISPR IN BACTERIA

The CRISPR-Cas system acts in a sequence-specific manner by recognizing and cleaving foreign DNA or RNA. The defence mechanism can be divided into three stages:

(i) Adaptation or spacer acquisition to the CRISPR locus post foreign DNA exposure

(ii) Biogenesis or expression of the locus which is then transcribed into a long precursor then processed into crRNA by Cas proteins and accessory components, and

(iii) Target interference or, target nucleic acids are recognized and destroyed by crRNA combined with traRNA and Cas proteins.

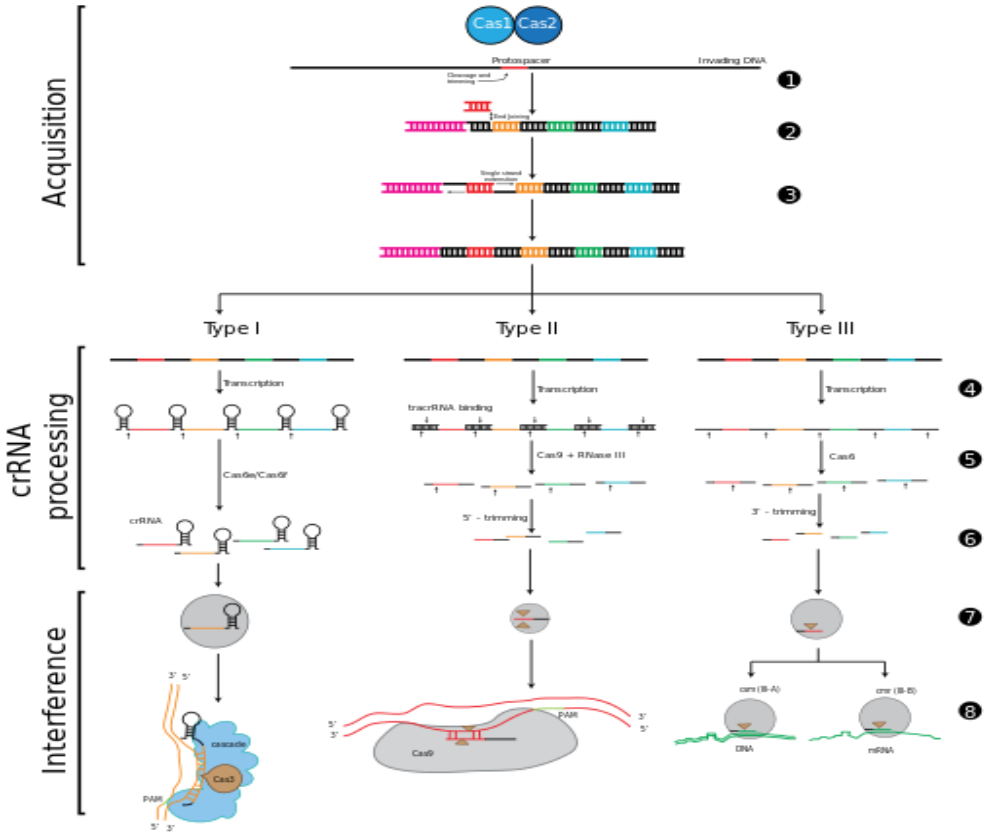


Figure: CRISPR mechanism in bacteria

CRISPR-Cas immunity is a natural process of bacteria and archaea. CRISPR-Cas prevents bacteriophage infection, conjugation andnatural transformation by degrading foreign nucleic acids that enter the cell [46]

5.1 Spacer acquisition

When a microbe is invaded by a virus, the first stage of the immune response is to capture viral DNA and insert it into a CRISPR locus in the form of a spacer. Cas1 and Cas2 are found in all three types of CRISPR-Cas immune systems, which indicates that they are involved in spacer acquisition. Mutation studies confirmed this hypothesis, showing that removal of cas1 or cas2 stopped spacer acquisition, without affecting CRISPR immune response.

Multiple Cas1 proteins have been characterised and their structures resolved. Cas1 proteins have diverse amino acid sequences. However, their crystal structures are similar and all purified Cas1 proteins are metal-dependent nucleases/integrases that bind to DNA in a sequence-independent manner. Representative Cas2 proteins have been characterised and possess either (single strand) ssRNA or (double strand) dsDNA- specific endoribonuclease activity.

In the I-E system of *E. coli* Cas1 and Cas2 form a complex where a Cas2 dimer bridges two Cas1 dimers In this complex Cas2 performs a non-enzymatic scaffolding role, binding double-stranded fragments of invading DNA, while Cas1 binds the single-stranded flanks of the DNA and catalyses their integration into CRISPR arrays. New spacers are always added at the beginning of the CRISPR next to the leader sequence creating a chronological record of viral infections In *E. Coli* a histone like protein called integration host factor (IHF), which binds to the leader sequence, is responsible for the accuracy of this integration[47]

5.2 Protospacer adjacent motifs

Bioinformatic analysis of regions of phage genomes that were excised as spacers (termed protospacers) revealed that they were not randomly selected but instead were found adjacent to short (3 – 5 bp) DNA sequences termed protospacer adjacent motifs (PAM). Analysis of CRISPR-Cas systems showed PAMs to be important for type I and type II, but not type III systems during acquisition. In type I and type II systems, protospacers are excised at positions adjacent to a PAM sequence, with the other end of the spacer cut using a ruler mechanism, thus maintaining the regularity of the spacer size in the CRISPR array. The conservation of the PAM sequence differs between CRISPR-Cas systems and appears to be evolutionarily linked to Cas1 and the leader sequence.

New spacers are added to a CRISPR array in a directional manner occurring preferentially, but not exclusively, adjacentto the leader sequence. Analysis of the type I-E system from *E. coli* demonstrated that the first direct repeat adjacent to the leader sequence, is copied, with the newly acquired spacer inserted between the first and second direct repeats.

The PAM sequence appears to be important during spacer insertion in type I-E systems. That sequence contains a strongly conserved final nucleotide (nt) adjacent to the first nt of the protospacer. This nt becomes the final base in the first direct repeat. This suggests that the spacer acquisition machinery generates single-stranded overhangs in the second-to-last position of the direct repeat and in the PAM during spacer insertion. However, not all CRISPR-Cas systems appear to share this mechanism as PAMs in other organisms do not show the same level of conservation in the final position. [48]It is likely that in those systems, a blunt end is generated at the very end of the direct repeat and the protospacer during acquisition.

5.3 Biogenesis

CRISPR-RNA (crRNA), which later guides the Cas nuclease to the target during the interference step, must be generated from the CRISPR sequence. The crRNA is initially transcribed as part of a single long transcript encompassing much of the CRISPR array.This transcript is then cleaved by Cas proteins to form crRNAs. The mechanism to produce crRNAs differs among CRISPR/Cas systems. In type I-E and type I-F systems, the proteins Cas6e and Cas6f respectively, recognise stem-loops created by the pairing of identical repeats that flank the crRNA. These Cas proteins cleave the longer transcript at the edge of the paired region, leaving a single crRNA along with a small remnant of the paired repeat region.

Type III systems also use Cas6, however their repeats do not produce stem-loops. Cleavage instead occurs by the longer transcript wrapping around the Cas6 to allow cleavage just upstream of the repeat sequence.

Type II systems lack the Cas6 gene and instead utilize RNaseIII for cleavage. Functional type II systems encode an extra small RNA that is complementary to the repeat sequence, known as a trans-activating crRNA (tracrRNA). Transcription of the tracrRNA and the primary CRISPR transcript results in base pairing and the formation of dsRNA at the repeat sequence, which is subsequently targeted by RNaseIII to produce crRNAs. Unlike the other two systems the crRNA does not contain the full spacer, which is instead truncated at one end.

CrRNAs associate with Cas proteins to form ribonucleotide complexes that recognize foreign nucleic acids. CrRNAs show no preference between the coding and non-coding strands, which is indicative of an RNA-guided DNA-targeting system.The type I-E complex (commonly referred to as Cascade) requires five Cas proteins bound to a single crRNA.[49]

5.4 Interference

During the interference stage in type I systems the PAM sequence is recognized on the crRNA-complementary strand and is required along with crRNA annealing. In type I systems correct base pairing between the crRNA and the protospacer signals a conformational change in Cascade that recruits Cas3 for DNA degradation.

Type II systems rely on a single multifunctional protein, Cas9, for the interference step. Cas9 requires both the crRNA and the tracrRNA to function and cleaves DNA using its dual HNH and RuvC/RNaseH-like endonuclease domains. Basepairing between the PAM and the phage genome is required in type II systems. However, the PAM is recognized on the same strand as the crRNA (the opposite strand to type I systems).

Type III systems, like type I require six or seven Cas proteins binding to crRNAs. The type III systems

analysed from *S. solfataricus* and *P. furiosus* both target the mRNA of phages rather than phage DNA genome, which may make these systems uniquely capable of targeting RNA-based phage genomes.

The mechanism for distinguishing self from foreign DNA during interference is built into the crRNAs and is therefore likely common to all three systems. Throughout the distinctive maturation process of each major type, all crRNAs contain a spacer sequence and some portion of the repeat at one or both ends. It is the partial repeat sequence that prevents the CRISPR-Cas system from targeting the chromosome as base pairing beyond the spacer sequence signals self and prevents DNA cleavage. RNA-guided CRISPR enzymes are classified as type V restriction enzymes.

VI. CRISPR IN GENE EDITING

CRISPR technology has the ability to efficiently modify endogenous genes in various species and cell types, and may even serve as potential therapy for genetic diseases. Before CRISPR, genomic alterations were limited to certain model organisms such as yeast and mice CRISPR/Cas9 genome editing is carried out with a Type II CRISPR system. When utilized for genome editing, this system includes Cas9, crRNA, tracrRNA along with an optionalsection of DNA repair template that is utilized in either Non-Homologous End Joining (NHEJ) or Homology Directed Repair (HDR).

6.1 Major components

Component	Function
crRNA	Contains the guide RNA that locates the correct section of host DNA along with a region that binds to tracrRNA (generally in a hairpin loop form) forming an active complex
tracrRNA	Binds to crRNA and forms an active complex.
sgRNA	Single guide RNAs are a combined RNA consisting of a tracrRNA and at least one crRNA
Cas9	Protein whose active form is able to modify DNA. Many variants exist with differing functions (i.e. single strand nicking, double strand break, DNA binding) due to Cas9's DNA site recognition function
Repair template	DNA that guides the cellular repair process allowing insertion of a specific DNA sequence

CRISPR/Cas9 often employs a plasmid to transfect the target cells. The main components of this plasmid are displayed in the image and listed in the table. The crRNA needs to be designed for each application as this is the sequence that Cas9 uses to identify and directly bind to the cell's DNA. The crRNA must bind only where editing is desired. The repair template is designed for each application, as it must overlap with the sequences on either side of the cut and code for the insertion sequence.

Multiple crRNAs and the tracrRNA can be packaged together to form a single-guide RNA (sgRNA)[51] This sgRNA can be joined together with the Cas9 gene and made into a plasmid in order to be transfected into cells.

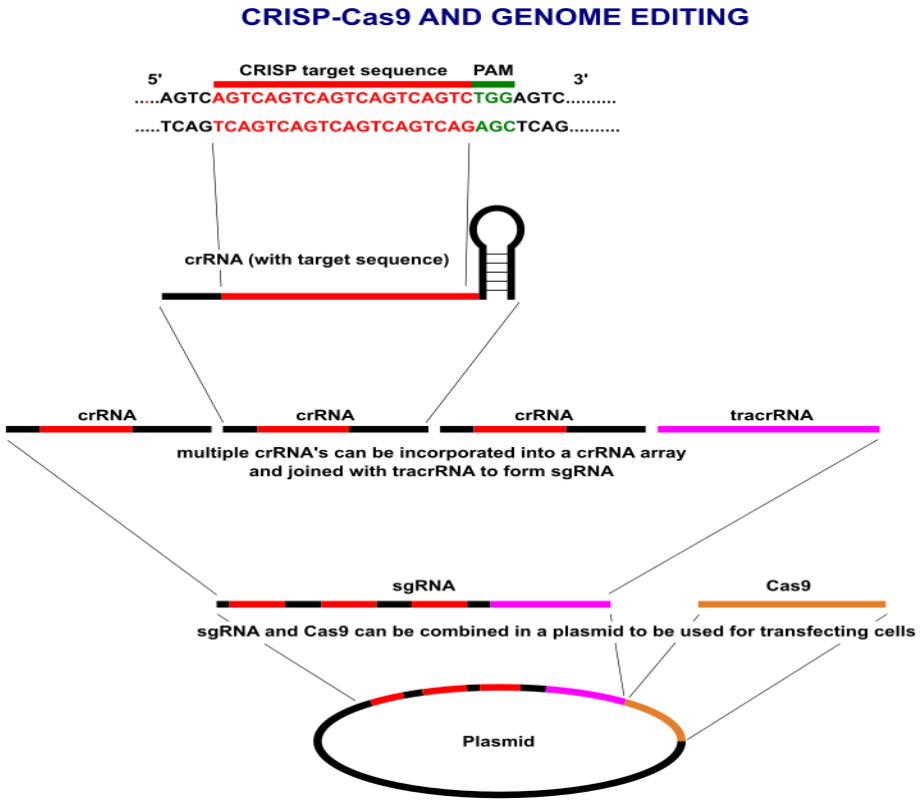


Figure: Overview of CRISPR Cas9 plasmid construction

6.2 Structure

CRISPR/Cas9 offers a high degree of fidelity and relatively simple construction. It depends on two factors for its specificity: the target sequence and the PAM. The target sequence is 20 bases long as part of each CRISPR locus in the crRNA array. A typical crRNA array has multiple unique target sequences. Cas9 proteins select the correct location on the host's genome by utilizing the sequence to bond with base pairs on the host DNA. The sequence is not part of the Cas9 protein and as a result is customizable and can be independently synthesized.

The PAM sequence on the host genome is recognized by Cas9. Cas9 cannot be easily modified to recognize a different PAM sequence. However this is not too limiting as it is a short sequence and nonspecific (e.g. the SpCas9 PAM sequence is 5'-NGG-3' and in the human genome occurs roughly every 8 to 12 base pairs).

Once these have been assembled into a plasmid and transfected into cells the Cas9 protein with the help of the crRNA finds the correct sequence in the host cell's DNA and – depending on the Cas9 variant – creates a single or double strand break in the DNA

Properly spaced single strand breaks in the host DNA can trigger homology directed repair, which is less error prone than the non-homologous end joining that typically follows a double strand break. Providing a DNA repair template allows for the insertion of a specific DNA sequence at an exact location within the genome. The repair template should extend 40 to 90 base pairs beyond the Cas9 induced DNA break. [52] The goal is for the cell's HDR process to utilize the provided repair template and thereby incorporate the new sequence into the genome. Once incorporated, this new sequence is now part of the cell's genetic material and passes into its daughter cells.

Many online tools are available to aid in designing effective sgRNA sequences

6.3 Delivery

Scientists can use viral or non-viral systems for delivery of the Cas9 and sgRNA into target cells. Electroporation of DNA, RNA or ribonucleocomplexes is the most common and cheapest system. This technique was used to edit CXCR4 and PD-1, knocking in new sequences to replace specific genetic "letters" in these proteins. The group was then able to sort the cells, using cell surface markers, to help identify successfully edited cells.[53] Deep sequencing of a target site confirmed that knock-in genome modifications had occurred with up to ~20% efficiency, which accounted for up to approximately one-third of total editing events. However, hard-to-transfect cells (stem cells, neurons, hematopoietic cells, etc.) require more efficient delivery

systems such as those based on lentivirus (LVs), adenovirus (AdV) and adeno-associated virus (AAV).

6.4 Editing

CRISPRs have been used to cut five to 62 genes at once: pig cells have been engineered to inactivate all 62 Porcine Endogenous Retrovirus in the pig genome, which eliminated transinfection from the pig to human cells in culture CRISPR's low cost compared to alternatives is widely seen as revolutionary. [54]

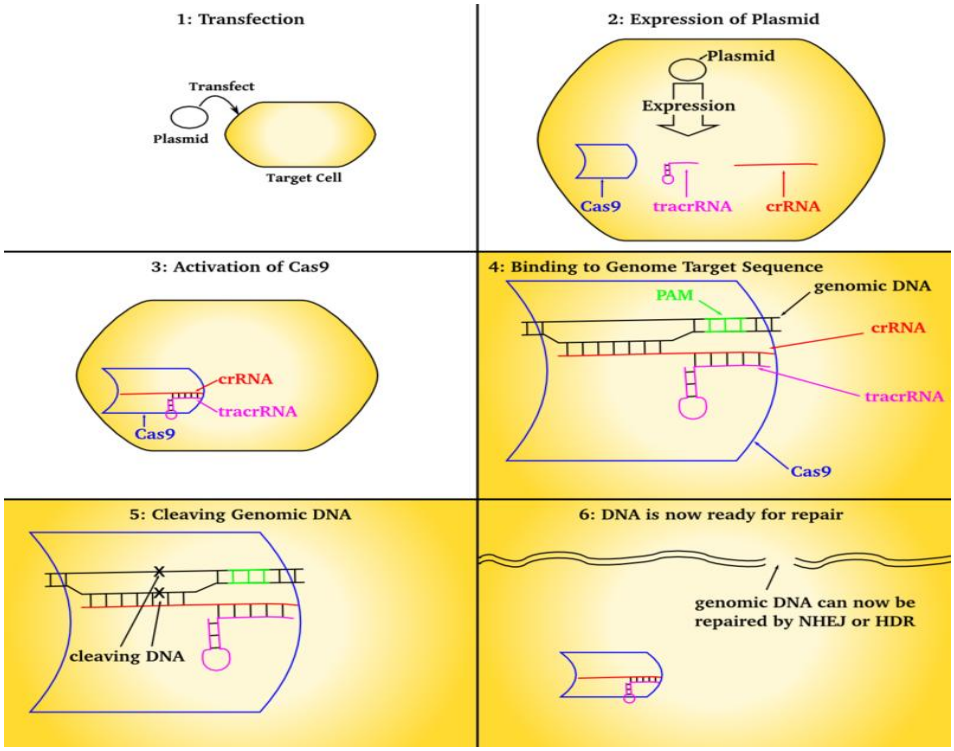


Figure: Overview of the transfection and DNA cleaving by CRISPR Cas9 (crRNA and tracrRNA are often joined as one strand of RNA when designing a plasmid

VII. ADVANTAGES

- The advent of CRISPR has revolutionized genome editing – not only for its cost effective specificity, but also for its ease-of-use in any lab, regardless of molecular biology expertise. Unlike ZF and TALE nucleases, CRISPR/Cas9 does not require protein engineering for each gene being targeted. The CRISPR/Cas9 system requires only a DNA construct encoding the target specific gRNA and Cas9, and if knock-in is being performed, the donor template for HR. Because the target specificity relies on ribonucleotide complex formation and not protein/DNA recognition, gRNAs can be designed readily and cheaply to target nearly any sequence in the genome specifically. E.g knock in/knock out mouse model requires 6- 12 months by traditional methods but with CRISPR requires 2 months and with 90% efficiency.

- The system is super-efficient -Multiple genes can be edited simultaneously with CRISPR, vastly increasing the efficiency of our experiments. Modifications can be introduced by directly injecting RNAs encoding the Cas protein and gRNA into developing mouse embryos. This eliminates the long and laborious processes of transfecting and selecting mouse ES cells that are required to create targeted mutant mice using classical homologous recombination techniques.

- The system is highly specific -The relative ease with which new sites can be targeted, simply by altering the short region of the gRNA that dictates specificity, makes this system a highly attractive method for introducing site-specific DSBs. They also demonstrated that they could co-inject oligonucleotides at the same time as the Cas and gRNAs to introduce specific mutations simultaneously via homologous DNA repair in two genes directly in mouse embryos. The ability to create mice with multiple mutations at the same time offers exciting possibilities to better understand genetic epistasis and to develop more complex disease models without the complicated and lengthy breeding strategies that often are required to create such mice from single gene knockouts. CRISPR-Cas9 is easily programmable by changing the guide sequence (20 nucleotides in the native RNA) of the sgRNA to any DNA sequence of interest.

- Additionally, CRISPR is capable of modifying chromosomal targets with high fidelity whereas ZFN/TALEN are prone to CpG methylation. Additionally, because the Cas9 protein is not directly coupled to the gRNA, this system is highly amenable to multiplexing through the concurrent use of multiple gRNAs to induce DSBs at several loci. CRISPR/Cas system to successfully introducing mutations in five different genes in mouse ES cells simultaneously.

- Last but not least, multiplexed genome editing with CRISPR/Cas9 can be easily achieved with the monomeric Cas9 protein and any number of different sequence-specific gRNAs. Because the rich diversity of natural CRISPR systems has been largely understudied, it is reasonable to expect many new CRISPR-based gene-editing technologies to emerge, including non-Cas9 based type II systems such as the recently described RNA-guided endonuclease Cpf1 and others. [55-56]

VIII. DISADVANTAGES

- Developing a delivery method in which the nuclease is not expressed for a sustained period of time is important both from a genotoxic standpoint and from an immunologic standpoint. It should be anticipated, until proven otherwise, that all of the engineered nuclease platforms will be seen by the immune system as foreign and will elicit a robust immune response that will both eliminate the therapeutically edited cells and perhaps cause toxic organ damage.
- CRISPR/Cas9 technology is changing the field of genome engineering and is also expected to change the treatment of genetic diseases. Achieving this goal requires not only improvements in efficacy and specificity, but also the optimization of delivery methods.The introduction of plasmids that simultaneously encode sgRNA and Cas9 into target cells by electroporation, nucleofection or lipofectamine represents a common and rapid method that can be applied to a wide variety of cell lines. The most commonly used plasmids always express an optimized SpCas9 and chimeric gRNA. With the demand for large fragment expression, multiple plasmids have been used to target different sites. However, all or part of the plasmids are often randomly integrated into the host genome [57] Plasmid DNA can also be inserted into both on-target and off-target sites, which can lead to difficulties in detection.
- Off-target effects - Any nuclease platform has the potential to produce off-target effects. Mutation introduced at non-specific loci with similar, but not identical, homology to the target sites are one of the most important complication of these technologies. These can be difficult to identify and require scanning the genome for mutations at sites with sequence similarity to the gRNA target sequence. Many studies have shown that the CRISPR/Cas9 system is a simple but highly efficient approach for genome editing in a variety of cells and organisms, both ***in vitro*** and ***in vivo***. The Cas9 nuclease can be directed via complementary base pairing between the first 20 nucleotides of sgRNA and target DNA sequences that lie adjacent to PAM. With the cleavage of Cas9, target sequences generate DSBs and introduce new genetic materials. Nevertheless, a number of studies have demonstrated that the CRISPR/Cas9 system can induce a substantial amount of off-target mutagenesis These off-target effects might play a role in recognizing and destroying hypervariable viral nucleic acids or plasmid DNA, which is beneficial to bacteria and archaea However, for biological studies and genetic therapies, off-target phenomena generate undesired mutations at random sites, thus impacting precise gene modification. Previous studies have shown that Cas9-mediated cleavage can be inhibited by a single mismatch in the complementary region of sgRNA and target sequences, especially the PAM-proximal nucleotides .On the basis of whole-genome sequencing, several studies have shown that mismatches at the 5′-terminal of the target sites are generally better tolerated. Moreover, studies in human cells have shown that up to five mismatches will not prevent cutting of target sequences, and that off-target sites can be mutagenized at frequencies higher than the intended on-target sites .Thus, for a better application of the CRISPR/Cas9 system, it is important to evaluate the potential hybrid effects of high-frequency off-target mutations.

- The induction of a DSB through a site-specific engineered nuclease is a critical aspect of gene editing and it is well known that DSBs can generate genomic instability, including chromosomal translocations, chromosome loss and aneuploidy.
- Key aspect in the clinical development of nuclease-mediated genome editing is to establish a series of assays that assess the potential safety of the process. Unfortunately, the field is too young for there to be any single assay or set of assays that has been validated as establishing whether an editing process will be safe in humans.
- Target site selection and sgRNA design - Among the potential pitfalls of CRISPR/Cas9 systems, sgRNA design is a prime concern. Because CRISPR/Cas9 systems are highly programmable, Cas9/sgRNA complexes can be utilized for genome editing or catalytically inactive Cas9 (dCas9)/sgRNA complexes can

be used for gene regulation. These applications require the design of sgRNAs that are efficient and specific. However, because this requires the consideration of many criteria, rational sgRNA design remains a major challenge. Previously, it was assumed that Cas9/sgRNA complexes could cleave double-strand DNA in the presence of PAM and an adjacent complementary target sequence. However, many experiments showed that some sgRNAs were less efficient or even inactive. For genome editing experiments, a pool of sgRNAs first required to be screened for activity; hence, design criteria to maximize sgRNA efficiency are a valuable pursuit.

- **Cas9 activity-** As a genome editing tool, the CRISPR/Cas9 system cleaves specific nucleotides based on sequence complementarity with only two significant components: the Cas9 protein and sgRNA. With the binding of sgRNA and target DNA, the Cas9 protein undergoes a large-scale conformational rearrangement. Accordingly, the catalytic nuclease lobe of Cas9 rotates ~ 100°, generating nucleic acid-cleaving activity .Generally, Cas9 can recognize genomic loci under the guidance of sgRNAs that bind to 20 nucleotide target sequences. However, Hsu *et al*. found that sgRNAs with +85 nucleotide tracrRNA tails increased the activity of Cas9 and induced higher level of indels *in vivo*. They also observed that both concatenated and interspaced two base mismatches, which occurred in the proximal region of PAM, greatly reduced Cas9 activity. This effect was further expanded to three concatenated mismatches, and three or more interspaced and five concatenated mismatches were found to abrogate Cas9 cleavage activity in most genes .Additionally, the results of further investigations suggested that excessively truncated guide RNA would also result in Cas9 losing cutting activity. Therefore, to provide higher Cas9 cutting efficiency, optimization of the sgRNA design and cautious selection of target sites are badly needed. [58]

IX. APPLICATION OF CRISPR

CRISPR/Cas9 technology has been adapted for many research applications beyond than genome editing, such as:

- CRISPR application on plants
- CRISPR/Cas9-mediated Chromatin Immunoprecipitation
- CRISPR Technologies for Transcriptional Activation and Repression
- Epigenetic Editing with CRISPR/Cas9
- Live Imaging of DNA/mRNA with CRISPR/Cas9
- CRISPR/Cas9 Therapeutic Applications

9.1 CRISPR application on plants

Plants can be transformed to produce both Cas9 and a target-specific gRNA, in order to cleave a specified target of invading DNA. For example, a Cas9/gRNA complex can be engineered to target the replicating DNA of Geminiviruses, which are highly destructive to crops in tropical and subtropical climates .Such an engineered Cas9/gRNA complex produces a sequence-specific, targeted immune response which can result in significant host resistance against a DNA virus. These laboratory-based results are exciting if they are reproduced in the field, since conventional breeding has not been universally successful against Geminiviruses .A variety of viral genetic elements can be successfully targeted which would confer long-term utility to this strategy. Crops engineered to express a CRISPR/Cas immune system are transgenic, containing DNA sequences which are bacterial and viral in origin (coding for Cas9 and gRNA, respectively), which may hamper public acceptance.

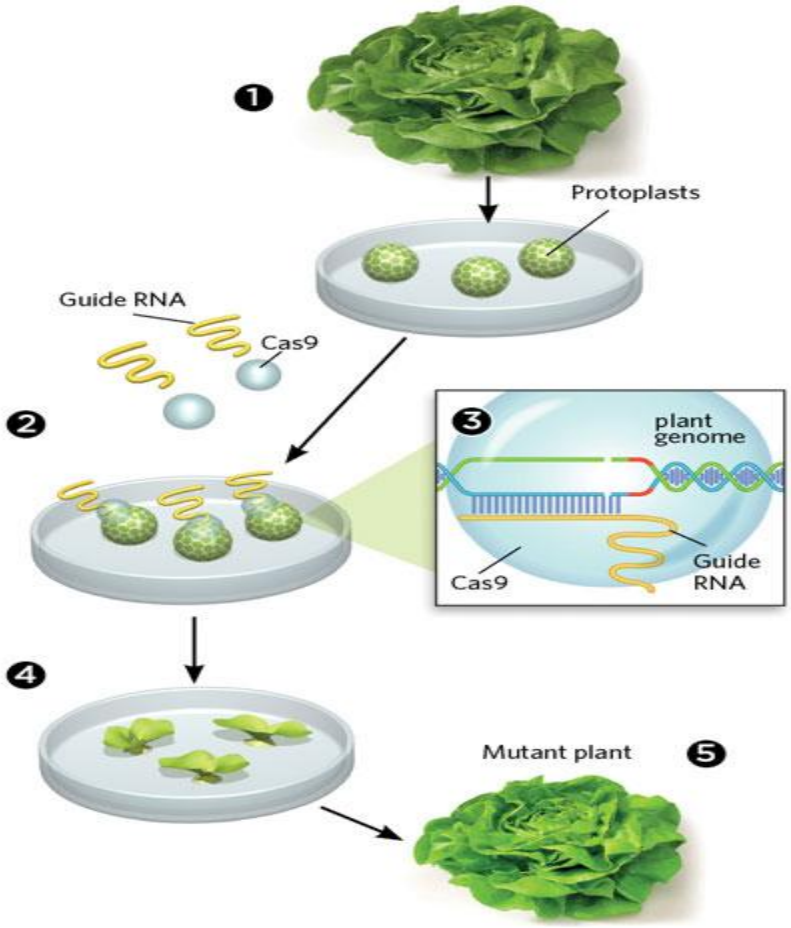


Fig: CRISPR in plants

9.2 CRISPR/Cas9-mediated Chromatin Immunoprecipitation

Purification of specific genomic loci is vital for the characterization of chromatin-associated proteins and RNAs. Modifications to the CRISPR/Cas9 system allow for flexible targeting and isolation of these genomic regions. A nuclear localization signal and epitope tag can be introduced into catalytically inactive Cas9 (dCas9) to create a DNA-binding protein that can be targeted by CRISPR guide RNAs. Existing CRISPR gRNA databases and design tools allow for targeting to any gene of interest. The CRISPR/Cas9-chromatin complex can then be purified with traditional chromatin immunoprecipitation (CHiP) techniques and exposed to mass spectrometry for further characterization. CRISPR-mediated CHiP techniques have been utilized to identify proteins associated with the interferon regulatory factor-1 (IRF-1) promoter region in response to interferon γ stimulation. In this study, researchers purified 15 associated proteins including histone deacetylase complex proteins, which have previously been implicated in interferon γ-mediated gene expression, as well as transcription factors, histones and other DNA-associated proteins. CRISPR-mediated CHiP holds a number of advantages over traditional CHiP methods. While large scale assays require the use of multiple antibodies against each DNA-binding protein or the creation and expression of epitope-tagged proteins, the modular nature of the CRISPR/Cas9 system requires only a single antibody against the tagged-Cas9 protein for purification. In addition, the CRISPR/Cas9 system is unaffected by issues stemming from low, differential or toxic gene expression. [59]

9.3 CRISPR Technologies for Transcriptional Activation and Repression

Several research groups have harnessed the specificity and easy re-programmability of the CRISPR/Cas9 system to create targetable CRISPR/Cas9 ribonucleoprotein complexes that can either activate (CRISPRa) or interfere (CRISPRi) with transcription of any desired coding region within a genome. These systems fuse dCas9 to a well-characterized transcription-regulatory domain, using pre-designed guide RNAs to direct the complex upstream of the transcription start site. By using inactivated dCas9 protein, the complex can be targeted to specific loci without cleaving or altering the genomic DNA. After Cas9 binds the targeted DNA

sequence, the fused transcription-regulatory domains are then able to recruit repressive or activating effectors to modify gene expression. [59]

9.4 Epigenetic Editing with CRISPR/Cas9

Epigenetic modifications to genomic DNA and histone proteins have been shown to play increasingly critical roles in biological processes. Epigenetic marks, such as methylation or acetylation, at specific genomic loci and histone residues can either be inherited or acquired, and can influence gene expression. Recent studies have used CRISPR/Cas9 genome editing to investigate the roles and targets for these epigenetic marks. In one such study, researchers performed CRISPR-mediated knock out of all three active DNA methyltransferases present in human embryonic stem cells, to characterize viable, pluripotent cell lines and study the distinct effects on the DNA methylation landscape. But researchers increasingly need methods for introducing epigenetic modifications at desired genomic loci, in order to model diseases and test hypotheses regarding potential therapeutic strategies. For example, specific epigenetic alterations are often necessary and sufficient to drive the transformation of normal cells into cancerous cells, and play roles in later steps of carcinogenesis. Using the CRISPR/Cas9 system, epigenetic editing has now become feasible. Utilizing inactive dCas9 as a DNA-binding domain platform, fused enzymes such as DNA methylases, histone acetyltransferases, and deacetylases (HATs or HDACs), can be targeted to alter the epigenetic state at precise locations within the genome. Researchers have used this approach, to fuse the catalytic core of human acetyltransferase p300 with dCas9, and shown this system to be sufficient for acetylation of histone H3 lysine 27 at specific target sites and to robustly activate transcription of target genes. Cas9 epigenetic effectors (epiCas9s) can also be used for genome-wide screening to discover novel relationships between epigenetic modifications, chromatin states, and phenotypes such as, cellular differentiation or disease progression [59]

9.5 Live Imaging of DNA/mRNA with CRISPR/Cas9

DNA visualization is an important application in understanding a variety of cellular processes, such as replication, transcription, and recombination, and the interactions between DNA and associated proteins and RNA. Two techniques are commonly used for DNA imaging, fluorescence in situ hybridization (FISH) and fluorescent tagging of DNA-binding proteins. FISH uses fluorescently tagged nucleic acid probes to bind and visualize DNA. While this technique offers the flexibility to target specific sequences through base pairing of the nucleic acid probes, it cannot be used for live imaging because of the requirement for sample fixation. Conversely, proteins tagged with a fluorescent label can be used for live imaging, but are limited by their fixed target sequences, restricting their use mostly to repetitive DNA sequences, such as telomeres. New advances in CRISPR/Cas9 technology offer the benefits of both live imaging and easy target sequence customization and flexibility. Inactivated dCas9 can be tagged with fluorophores for imaging both repetitive DNA elements and protein-encoding genes, enabling us to observe chromatin organization throughout the cell cycle. In addition to live DNA imaging, the CRISPR/Cas9 system can be used for live RNA imaging as well. Modifications to the gRNA sequence allow for mRNA recognition and tracking. Using CRISPR-mediated RNA imaging techniques, researchers have been able to visualize the accumulation of ACTB, CCNA2 and TRFC mRNAs in RNA granules. These new applications improve existing methodologies. [59]

9.6 CRISPR/Cas9 Therapeutic Applications

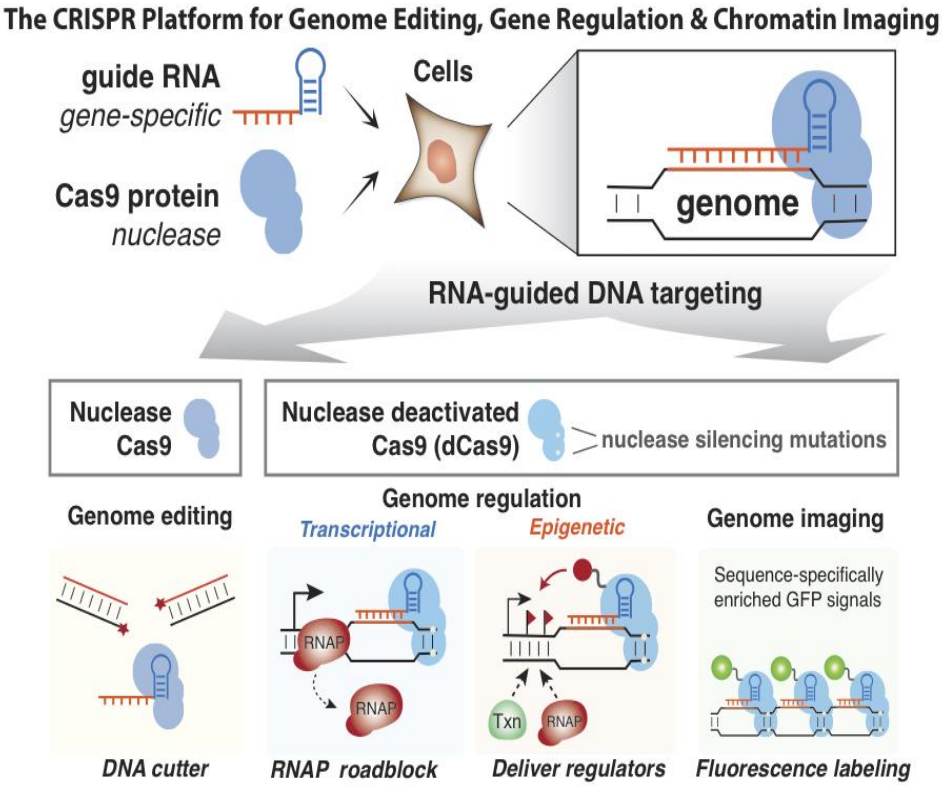
Both well-established pharmaceutical companies and new start-up biotech companies are racing to create CRISPR-based therapeutics. Compared to other strategies for gene therapy, CRISPR genome editing is thought to be faster, less expensive, and potentially far safer. Autologous CRISPR cell therapies that use genome editing to correct a mutation in a patient's own cells hold promise in circumventing the rejection issues present with transplant therapies that require donor matching. CRISPR genome editing is especially promising for diseases that can be tackled by modifying cells that can easily be removed from a patient, which allows for additional screening to ensure no off-target genome modifications during genome-editing.

9.7 Cancer immunotherapy

Cancer immunotherapy has been widely recognized as one of the greatest advances in biomedical research in recent years. In particular, adoptive T-cell immunotherapy, in which autologous T cells are engineered to attack cancer antigens ex vivo and transferred back to the patient, has been impressively successful at treating some cases of lymphoma, leukemia, and melanoma. Despite these successes and promising ongoing clinical trials, there are several areas in which T-cell immunotherapy could be potentially improved by gene editing. Here, both the efficacy against diverse tumor types and the ability to manufacture cell products that can be applied to a broad patient population could be enhanced through gene-editing techniques. For example, a promising strategy for immunotherapy involves engineering T cells to express synthetic receptors known as chimeric antigen receptors, or CARs, that recognize epitopes on cancer cells. Such CAR T cells have been particularly successful in treating B-cell lymphoma by targeting the CD19 cell surface antigen.

However, one limitation of this approach is that these modified T cells express both the endogenous T-cell receptor as well as the engineered CAR. Because these receptors function as dimers, the natural and engineered receptors can dimerize and interact, resulting in unpredictable epitope specificity and potentially reducing therapeutic potency. To address this limitation, several studies have focused on knocking out the endogenous T-cell receptors with engineered nucleases. A major challenge to the development of broadly translatable T-cell immunotherapies is the need to use autologous cellsto avoid immune rejection. To address this, gene editing has been used to knockout the human leukocyte antigen (HLA) by which the immune system discriminates self and foreign cells. Importantly, this approach may be broadly useful for allogeneic cell therapy beyond T-cell immunotherapy. For example, similar approaches have been applied in human pluripotent cells potentially having diverse uses in regenerative medicine as well as in endothelial cells that could be used for allogeneic vascular grafts. Another major obstacle to successful T-cell immunotherapy is the inhibition of T-cell effector functions by the expression of checkpoint inhibitors on the surface of tumor cells. For example, the binding of such inhibitors to the PD-1 receptor on T cells is well documented to block T-cell effector function and induce apoptosis and exhaustion. PD-1 receptor inhibition thus provides a mechanism by which cancer cells successfully evade the immune system. As a strategy to overcome this, gene editing has been used to knockout PD-1 in T cells leading to increased T-cell effector function. The success of this gene-editing strategy is likely extendable to other checkpoint inhibitor pathways that cancer cells exploit to circumvent immunosurveillance, and thus may be a critical technology for broadly enabling immunotherapy for diverse cancer types. Finally, for indications such as glioblastoma, the apoptosis of the engineered T cells resulting from post-surgery anti-inflammatory glucocorticoid steroid treatment severely limits the efficacy of T-cell immunotherapy. In order to create a glucocorticoid-resistant T-cell source, gene editing was used to knockout the endogenous T-cell receptor. This led to successful anti-glioma T-cell therapy in mouse models and was the basis of a subsequent clinical trial

New studies using CRISPR/Cas9-mediated immunotherapy are being used to combat metastatic lung cancer at Sichuan University and myeloma at the University of Pennsylvania. Researchers will be using the CRISPR/Cas9 system to knock out the PD-1 gene in T-cells extracted from patients. PD-1 is an important down-regulator in T-cell activation and functions as an immune checkpoint. After knockout of PD1, the modified T-cells will be released into the bloodstream to target cancer. PD-1 inhibition offers a promising approach for cancer treatment. Last year alone, the FDA approved two new antibody-based therapies which target PD-1, nivolumab and pembrolizumab. [60]



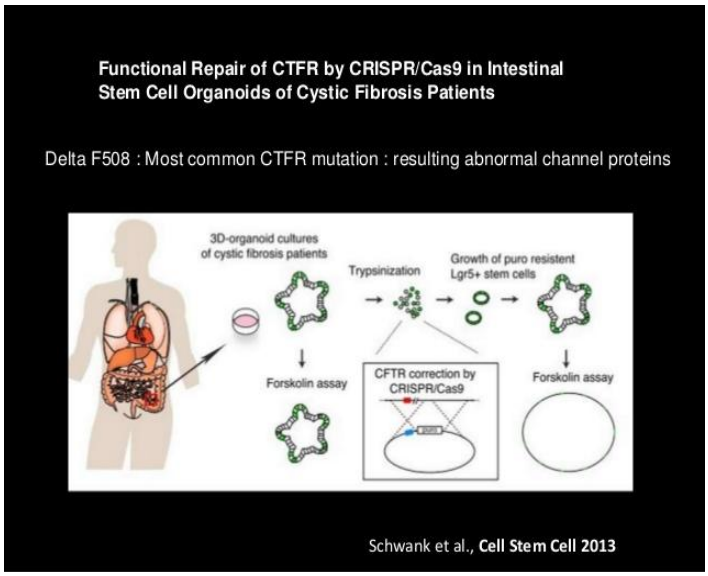
9.8 Gene Therapy

Huntingtons disease is an inherited neurological condition caused by accumulation of mutant Huntingtin protein within the brain which results in cognitive impairment, dementia and death. Using mouse models, researchers have shown that CRISPR/Cas9 gene editing can knock-out production of these mutant proteins. CRISPR/Cas9 editing is sufficient to reduce mutant protein production by up to 90% and shows promise as a therapeutic solution. Additional studies are ongoing using humanized huntington genes.

9.9 Cystic fibrosis

CF is a genetic disease caused by a mutation that occurs in the gene that encodes the cystic fibrosis transmembrane conductance regulator (CFTR) protein. Normal CFTR proteins serve as channels to allow the transport of water and charged ions (e.g., chloride) in and out of cells, creating a thin mucus that protects and lubricates internal organs (like the lungs and pancreas). The defective CFTR protein leads to an imbalance of water and ion flow in and out of cells, resulting in thick mucus that obstructs airways and traps bacteria. Chronic inflammations and infections are common in people with CF for this reason. CF therapies can target DNA, RNA, or proteins, but only at the DNA level can the mutated CFTR gene be replaced and CFTR function be restored.

Editas Medicine's approach relies on gene editing, in which the wrong DNA sequence of the defective CFTR gene is replaced by the correct one using the CRISP/Cas-9 technology. CRISPR/Cas9 uses a protein-RNA complex composed of an enzyme (or protein) — Cas9 — bound to a guide RNA (gRNA) molecule that recognizes the wrong DNA sequence and cuts it out. The cell can then fill the excised portion with the correct gene sequence. The delivery mode is expected to be by adeno-associated virus (AAV) or lipid nanoparticle (LNP)



9.10 Malaria and Insect-borne Diseases

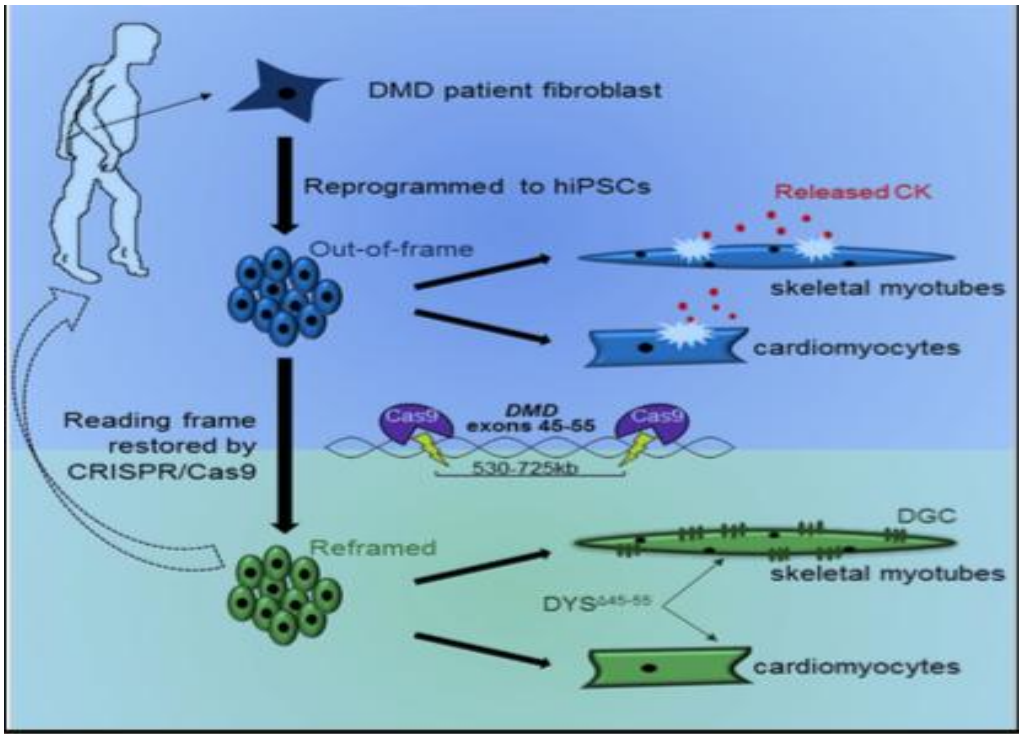
Insect-borne diseases such as malaria and zika pose enormous health concerns across the world. To combat the spread of insect-borne diseases, researchers have modified CRISPR/Cas9 into highly efficient "gene drive" systems which can spread disease resistance genes to entire populations. To create a gene drive, researchers have package disease resistance genes together with CRISPR gRNA and Cas9 components into a single DNA construct. After insertion, the gene drive autonomously replicates into both parental chromosomes, and is inherited by ~99.5% of progeny. Advances in gene drive technology offer immediate solutions for the eradication of these diseases.

9.11 Obesity and Metabolism

FTO is one of the most strongly linked genes to obesity. Certain FTO genetic variants correlate significantly with obesity and heavier weight. Researchers have shown that CRISPR/Cas9-mediated knock-in techniques can convert obesity-promoting FTO variants to normal variants in adipocyte precursor cells. Treated cells display increased metabolic activity and reduced expression of IRX3 and IRX5, genes which determine cell fate as white adipocytes for fat storage.

9.12 Neuromuscular disorders (DMD)

Advances in gene delivery and cell transplantation to the central nervous system and skeletal and cardiac muscle have created new opportunities for gene and cell therapy for many neuromuscular disorders, including DMD, the limb girdle muscular dystrophies, spinal muscular atrophy, Friedreich's ataxia, Huntington's disease, and amyotrophic lateral sclerosis (ALS). Amongst this class of diseases, genome editing has thus far advanced most prominently for DMD, although possible strategies to apply genome editing to other conditions could be envisioned. DMD is caused by mutations to the dystrophin gene, most commonly large deletions that shift the downstream gene fragment out of frame and render the protein product nonfunctional. Because the coding sequence of the dystrophin gene is exceptionally large (14 kb), it cannot be packaged into size-restricted viral delivery vectors. Although truncated minigenes have been developed that do fit into viral vectors, they are only partially functional compared to the full-length gene and their ability to reverse the human disease remains to be determined. For these reasons, and because there is currently no available approved therapy for DMD, gene editing to repair the endogenous gene is particularly compelling. Early reports suggested a mechanism to repair the dystrophin gene with which was followed by proof-of-principle experiments in cultured cells from DMD patients demonstrating dystrophin gene repair by targeted integration of the deleted exons 270 or restoration of dystrophin protein expression by targeted shifting of the reading frame by NHEJ-mediated indels. However, these two strategies suffer from addressing only a limited patient population with any particular gene-editing strategy or lacking predictable and reliable editing outcomes due to the reliance on stochastic NHEJ-mediated DNA repair, respectively. Therefore, more recent studies have focused on deleting one or more exons with a combination of nucleases to generate precisely restored protein products and address larger fractions of the DMD patient population. This includes a single strategy of deleting >300 kb of genomic DNA comprising exons that could be applicable to restoring dystrophin expression in 62% of DMD patients. In order to develop this into an approach that could potentially be applied clinically to DMD patients, recent work has incorporated the CRISPR/Cas9 system into AAV vectors with tropism for skeletal and cardiac muscle. When applied locally via intramuscular injection or systemically via intravenous injection to a mouse model of DMD, gene editing by CRISPR/Cas9 restored expression of the dystrophin protein and improved muscle pathology and strength. Notably, one study showed relatively efficient in vivo gene editing of Pax7-positive muscle progenitor cells that may act as a renewable source of cells in which the dystrophin gene has been repaired. This translational approach builds on demonstration of in vivo gene editing in skeletal muscle with adenoviral delivery and correction of dystrophin mutations in single-cell mouse embryos to reverse disease symptoms. In the future, these efforts may be extended to cell therapies by using patient-derived cell types, such as iPS cells that could be modified by gene correction or targeted dystrophin transgene insertion and expanded to large numbers and efficiently engrafted into muscle tissue.



9.13 Antiviral strategies (HIV treatment)

Gene editing can provide new strategies and therapeutic applications against infectious viral diseases. HIV has been eff ectively eliminated in patients using gene therapy to delete receptors essential for viral cell entry and infection. Recent studies using CRISPR technology have shown that mutations in CCR5 and CXCR4 receptors in both induced pluripotent stem cells (iPSCs) and primary CD4+ cells can lead to HIV resistance in lineages derived from these cells.

The most straightforward application of gene editing is to use the relatively efficient NHEJ mechanism to knockout genes in an ex vivo autologous cell therapy, where somatic cells can be isolated, modified, and delivered back to the patient. Moreover, one of the most compelling applications of gene editing is the prevention of viral infection or replication. Thus the most advanced gene-editing strategy to date is the ex vivo modification of T cells to knock out the CCR5 coreceptor used for primary HIV infection.This early study demonstrated decreased viral loads and increased CD4+ T-cell counts in HIV-infected mice engrafted with T cells in which the CCR5 gene had been knocked out by zinc finger nucleases. This was later followed by demonstration of similar results following gene editing and transplantation of CD34+ HSCs into irradiated mice, allowing for protection of all . Ex vivo and in vivo strategies for therapeutic genome editing. AAV Lipid nanoparticle Direct delivery to patient using viral or non-viral delivery vehicle Ex vivo In vivo Introduce modified cells back into patient Extract stem or progenitor cells Deliver targeted nucleases to cells by physical, chemical, or viral methods Protein DNA. These studies have led to a series of clinical trials evaluating this approach in HIV-positive human patients. Thus far the studies show safe engraftment and survival of CCR5-modified T cells and control of viral load in some patients, providing promising proofof-principle of a gene-editing approach in humans. Interestingly, data from this study showed a greater clinical efficacy in a patient that was already heterozygous for the naturally-occurring Δ32 mutation, suggesting that gene-editing efficiency may be a critical factor for success. Building on these promising studies with ZFNs, several other efforts have developed similar gene-editing strategies to knockout CCR5 with TALENs CRISPR/Cas9 and meganucleases. Other work has expanded beyond targeting only CCR5 to enhance resistance to HIV infection. This includes targeting the CXCR4 coreceptor or PSIP1 gene encoding the LEDGF/ p75 protein required for HIV integration. Some studies have used targeted gene integration into the CCR5 gene by HDR to simultaneously knockout CCR5 and introduce anti-HIV factors. Finally, complete excision of the HIV genome from infected cells using nucleases that target sequences in the long terminal repeats (LTRs) flanking the viral genome has also been reported.Thus, a variety of next-generation gene-editing strategies for preventing HIV infection and replication are on the horizon. Beyond addressing HIV infection, all of the gene-editing platforms have also been applied to various other viral pathogens including hepatitis B virus.herpes simplex virus, and human papilloma virus.These strategies typically involve removing viral genomes by degradation following nuclease cleavage and by targeting genes essential for genome stability, maintenance, and replication. While many of these early studies focused on proof-of-principle reduction of viral load in cell culture or following hydrodynamic plasmid DNA delivery to mice, recent studies using AAV delivery of gene-editing tools directly to the mouse liver provides a plausible path for scalability and clinical translation. A general challenge of antiviral therapies is the high mutability of viral targets. This is a compelling argument in favor of targeting host genes, such as CCR5, but may also be addressed by simultaneous targeting of multiple critical sites in the viral genome.

9.14 RECENT ADVANCES OF CRISPR

Cas9 variants and orthologs for genome engineering
New version of cas 9
Cpf1 (CRISPR from *Prevotella* and *Francisella1)*

In 2015, the nuclease Cpf1 was discovered in the CRISPR/Cpf1 system of the bacterium *Francisella novicida* .Cpf1 showed a 'staggered' cut in double stranded DNA as opposed to the 'blunt' cut produced by Cas9 .Requires only a crRNA for successful targeting in contrast to Cas9 which requires both crRNA and a tracrRNA.It is A New Tool for Plant Genome Editing .It shows robust nuclease activity in human cells.

Besides off-target issues, the PAM requirement also restrains the Cas9 targeting range, even though the PAM sequence of 5′-NGG-3′ appears at high frequency in genomes (5–12 times in every 100 bp for model plant species. To bypass the PAM limitation, a number of Cas9 orthologs from type II CRISPR-Cas systems were characterized and engineered for genome editing). These Cas9 orthologs recognize different PAM sequences and are smaller than the commonly used *S. pyogenes* Cas9, providing useful alternatives for RNA-guided genome engineering. Among these Cas9 orthologs, Cpf1 and FnCas9 recognize a PAM sequence of 5′-TTN-3′ and 5′-YG-3′, respectively which is present in genomes at an equal or higher frequency than 5′-NGG-3′. Because the CRISPR/Cas system is a general immunity system in bacteria and archaea, more Cas9 orthologs and other types of Cas with different PAM specificities will probably be found and engineered for genome editing in the near future.

Of note, the PAM specificities of Cas9 can be modified into other sequences by substituting several amino acid residues in the PAM-binding domain T.hese new Cas9 variants can recognize a broad range of PAM sequences including 5′-NAGN-3′, 5′-NGCG-3′, and 5′-NGNG-3′. These improvements allow the commonly used *S. pyogenes* Cas9 to edit almost any genomic site. After adoption of these Cas9 orthologs or variants, the PAM sequence restraint will be removed.[61]

X. CONCLUSION

CRISPR Cas9 technology has massive potential to revolutionize biological research. Despite major breakthroughs since its discovery, CRISPR Cas9 is still a new frontier in genome engineering [62]. This technology dramatically expanded the ability to manipulate genes and many scientists recognize its potential to help understand and treat diseases. As the utilization of CRISPR Cas9 opens the door to build knowledge it has positive application in many fields of research. Its application in genome wide studies will enable large-scale screening for drug treatments. It can be utilized in the agricultural research and pharmacological studies. Future research is directed to elucidate CRISPR Cas9 mode of action and improve the technology. A large focus on the improvement of CRISPR Cas9 will be on eliminating any off-target effects. It is likely that it will be many years before CRISPR Cas9 is used to directly edit human genomes and produce designer babies having no diseases and with eternal youth. When contemplating the function of repetitive and viral sequences in prokaryotic defense from foreign DNA, one must wonder what the similar sequences in eukaryotes remain to reveal [63]. CRISPR technology will greatly progress knowledge about developmental processes and other basic biological mechanisms. As this technique becomes more commonplace and affordable, CRISPR use will be a staple in reverse genetic, forward genetic and basic cell biological investigations. CRISPR has ushered in an exciting time for science that is ripe for discovery. [64]

REFERENCES

[1]. Rosenberg LE. Legacies of Garrod's brilliance. One hundred years—and counting. *J Inherit Metab Dis. 2008;31:574–9*,

[2]. Johnston JJ, Lewis KL, Ng D, Singh LN, Wynter J, Brewer C, et al. *Individualized iterative phenotyping for genome-wide analysis of loss-of-function mutations.* Am J Hum Genet. 2015;96:913–25.

[3]. Petes TD. Molecular genetics of yeast. Annu Rev Biochem. 1980;49:845–76.

[4]. Orr-Weaver TL, Nicolas A, Szostak JW. Gene conversion adjacent to regions of double-strand break repair. Mol Cell Biol. 1988;8:5292–8

[5]. Jasin M. Genetic manipulation of genomes with rare-cutting endonucleases. Trends Genet. 1996;12:224–8.

[6]. Smih F, Rouet P, Romanienko PJ, Jasin M. Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. Nucleic Acids Res. 1995;23:50

[7]. Rouet P, Smih F, Jasin M. Introduction of double-strand breaks into the genome of mouse cells by expression of rare-cutting endonuclease. Mol Cell Biol. 1994;14:8096–106

[8]. Choulika A, Perrin A, Dujon B, Nicolas J-F. Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of *Saccaromyces cerevisiae*. Mol Cell Biol. 1995;15:1968–73.

[9]. West SC, Chappell C, Hanakahi LA, Masson JY, McIlwraith MJ, Van Dyck E. Double-strand break repair in human cells. Cold Spring Harb Symp Quant Biol. 2000;65:315–21.

[10]. Hendel A, Kildebeck EJ, Fine EJ, Clark JT, Punjya N, Sebastiano V, et al. Quantifying genome-editing outcomes at endogenous loci with SMRT sequencing. Cell Rep. 2014;7:293–305.,

[11]. Lin Y, Waldman AS. Capture of DNA sequences at double-strand breaks in mammalian chromosomes. Genetics. 2001;158:1665–74.].

[12]. Silva G, Poirot L, Galetto R, Smith J, Montoya G, Duchateau P, et al. Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. Curr Gene Ther. 2011;11:11–27.,

[13]. Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci U S A. 1996;93:1156–60.,

[14]. Durai S, Mani M, Kandavelou K, Wu J, Porteus MH, Chandrasegaran S. Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. Nucleic Acids Res. 2005;33:5978–90.

[15]. Tebas P, Stein D, Tang WW, Frank I, Wang SQ, Lee G, et al. Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. N Engl J Med. 2014;370:901–10.

[16]. Bogdanove AJ, Voytas DF. TAL effectors: customizable proteins for DNA targeting. Science. 2011;333:1843–6.

[17]. Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. Science. 2009;326:1501

[18]. Miller JC, Zhang L, Xia DF, Campo JJ, Ankoudinova IV, Guschin DY, et al. Improved specificity of TALE-based genome editing using an expanded RVD repertoire. Nat Methods. 2015;12:465–71.
[19]. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science. 2014;346:1258096.
[20]. Cong, L., et al., Multiplex genome engineering using CRISPR/Cas systems. Science, 2013. 339(6121):819-23
[21]. Feng, Z., et al., Efficient genome editing in plants using a CRISPR/Cas system. Cell Res, 2013. 23(10):1229-32
[22]. Ran, F.A., et al., Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell, 2013. 154(6):1380-9.
[23]. Ran, F.A., et al., Genome engineering using the CRISPR-Cas9 system. Nat Protoc, 2013. 8(11):2281308.
[24]. Mali, P., et al., RNA-guided human genome engineering via Cas9. Science, 2013.339(6121):823-6. 29. Jinek, M., et al., A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science, 2012. 337(6096):816-21
[25]. Hilton, I.B., et al., Epigenome editing by a CRISPRCas9-based acetyltransferase activates genes from promoters and enhancers. Nat Biotechnol, 2015. 33(5):510-7.
[26]. Harrison, M.M., et al., A CRISPR view of development. Genes Dev, 2014. 28(17):1859-72.
[27]. Laufer, B.I. and S.M. Singh, Strategies for precision modulation of gene expression by epigenome editing: an overview. Epigenetics Chromatin, 2015. 8:34.
[28]. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. Cell. 2014;157:1262–78.
[29]. Carte J, Wang R, Li H, Terns RM, Terns MP. 2008 Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. Genes Dev.22, 3489–3496. (doi:10.1101/gad.1742908)
[30]. Haft DH, Selengut J, Mongodin EF, Nelson KE. 2005 A guild of 45 CRISPR-associated (Cas) proteinfamilies and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. PLoS Comput. Biol. 1, e60. (doi:10.1371/journal.pcbi.0010060)
[31]. .Haurwitz RE, Jinek M, Wiedenheft B, Zhou K, Doudna JA. 2010 Sequence- and structure-specificRNA processing by a CRISPR endonuclease. Science329, 1355–1358. (doi:10.1126/science.1192272)
[32]. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S,Romero DA, Horvath P. 2007 CRISPR provides acquired resistance against viruses in prokaryotes.Science 315, 1709–1712. (doi:10.1126/science.1138140)rstb.royalsocietypublishing.org Phil. Trans. R. Soc. B 371: 20150496.
[33]. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S,Romero DA, Horvath P. 2007 CRISPR provides acquired resistance against viruses in prokaryotes.Science 315, 1709–1712. (doi:10.1126/science.1138140)rstb.royalsocietypublishing.org Phil. Trans. R. Soc. B 371: 20150496.
[34]. Brouns SJ et al. 2008 Small CRISPR RNAs guide antiviral defense in prokaryotes. Science 321,960–964. (doi:10.1126/science.1159689)
[35]. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012 A programmable dual-RNAguided DNA endonuclease in adaptive bacterial immunity. Science 337, 816–821. (doi:10.1126/science.1225829)
[36]. Cong L et al. 2013 Multiplex genome engineeringusing CRISPR/Cas systems. Science 339, 819–823. (doi:10.1126/science.1231143)
[37]. K.S. Makarova, *et al.*Evolution and classification of the CRISPR-Cas systemsNat. Rev. Microbiol., 9 (6) (2011), pp. 467-477.
[38]. T. Sinkunas, *et al.*Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system EMBO J., 30 (7) (2011), pp. 1335-1342
[39]. E. Deltcheva, *et al.*CRISPR RNA maturation by trans-encoded small RNA and host factorRNasEIIIINature, 471 (7340) (2011), pp. 602-607
[40]. J.E. Garneau, *et al.*The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA Nature, 468 (7320) (2010), pp. 67-71
[41]. M. Jinek, *et al.*A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity Science, 337 (6096) (2012), pp. 816-821
[42]. Y. Wei, R.M. Terns, M.P. TernsCas9 function and host genome sampling in Type II-A CRISPR-Cas adaptationGenes Dev., 29 (4) (2015), pp. 356-361
[43]. R. Heler, *et al.*Cas9 specifies functional viral targets during CRISPR-Cas adaptation Nature (2015)
[44]. N. Beloglazova, *et al.*A novel family of sequence-specific endoribonucleases associated with the clustered regularly interspaced short palindromic repeatsJ. Biol. Chem., 283 (29) (2008), pp. 20361-20371
[45]. J.K. Nunez, *et al.*Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunityNat. Struct. Mol. Biol., 21 (6) (2014), pp. 528-534

[46]. *.Marraffini LA (October 2015). "CRISPR-Cas immunity in prokaryotes". Nature. **526** (7571): 55–61. PMID 26432244. doi:10.1038/nature15386.*

[47]. *Nuñez, James K.; Bai, Lawrence; Harrington, Lucas B.; Hinder, Tracey L.; Doudna, Jennifer A. (2016-06-16). "CRISPR Immunological Memory Requires a Host Factor for Specificity". Molecular Cell. **62** (6): 824–833. ISSN 1097-4164. PMID 27211867. doi:10.1016/j.molcel.2016.04.027.*

[48]. *Shah SA, Erdmann S, Mojica FJ, Garrett RA (May 2013). "Protospacer recognition motifs: mixed identities and functional diversity". RNA Biology. **10** (5): 891–9. PMC 3737346. PMID 23403393. doi:10.4161/rna.23764.*

[49]. MM, Lundgren M, van Duijn E, Bultema JB, Westra ER, Waghmare SP, Wiedenheft B, Pul U, Wurm R, Wagner R, Beijer MR, Barendregt A, Zhou K, Snijders AP, Dickman MJ, Doudna JA, Boekema EJ, Heck AJ, van der Oost J, Brouns SJ (May 2011). "Structural basis for CRISPR RNA-guided DNA recognition by Cascade". Nature Structural & Molecular Biology. **18** (5): 529–36. PMID 21460843. doi:10.1038/nsmb.2019.

[50]. *Zhang J, Rouillon C, Kerou M, Reeks J, Brugger K, Graham S, Reimann J, Cannone G, Liu H, Albers SV, Naismith JH, Spagnolo L, White MF (February 2012). "Structure and mechanism of the CMR complex for CRISPR-mediated antiviral immunity". Molecular Cell. **45** (3): 303–13. PMC 3381847. PMID 22227115. doi:10.1016/j.molcel.2011.12.013*

[51]. *Ly, Joseph (2013). Discovering Genes Responsible for Kidney Diseases (Ph.D.). University of Toronto. Retrieved 26 December 2016.*

[52]. *Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F (November 2013). "Genome engineering using the CRISPR-Cas9 system". Nature Protocols. **8** (11): 2281–308. PMC 3969860. PMID 24157548. doi:10.1038/nprot.2013.143*

[53]. Scientists successfully edit human immune-system T cells | KurzweilAI". www.kurzweilai.net. Retrieved 2016-01-01

[54]. Ledford, Heidi (3 June 2015). "CRISPR, the disruptor". News Feature. Nature. **522** (7554)

[55]. Zetsche, B, Gootenberg, JS, Abudayyeh, OO, Slaymaker, IM, Makarova, KS, Essletzbichler, P et al. (2015). Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. Cell 163: 759–771

[56]. Ousterout, DG, Perez-Pinera, P, Thakore, PI, Kabadi, AM, Brown, MT, Qin, X et al. (2013).

[57]. M Ullah, Y Kuroda, T J Bartosh, F Liu, Q Zhao, C Gregory, R Reger, J Xu, R H Lee, D J Prockop, iPS-derived MSCs from an expandable bank to deliver a prodrug-converting enzyme that limits growth and metastases of human breast cancers, Cell Death Discovery, 2017, 3, 16064

[58]. Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, Zheng Z, Gonzales AP, Li Z, Peterson RT, Yeh JRet al. (2015) Engineered CRISPR-Cas9 nucleases with altered PAM specificities. Nature523, 481–485.

[59]. www.Handbook of genome editing.com

[60]. Si Chen*, Heng Sun*, Kai Miao*, and Chu-Xia Deng CRISPR-Cas9: from Genome Editing to Cancer Research Int. J. Biol. Sci. 2016, Vol. 12

[61]. Ran et al., 2015.

[62]. Whalley, K., Techniques: A CRISPR method of localization. Nat Rev Neurosci, 2016. 17(7):399.

[63]. Honda, T. and K. Tomonaga, Endogenous nonretroviral RNA virus elements evidence a novel type of antiviral immunity. Mob Genet Elements, 2016. 6(3):e1165785.

[64]. Bikard, D., et al., Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic Acids Res, 2013. 41(15):7429-37.

IOSR Journal of Pharmacy (IOSR-PHR) is UGC approved Journal with Sl. No. 3365, Journal No-62875

Md. Abul Hassan. "Review on CRISPR/Cas-9 Gene Editing Technology and Its Potential Clinical Application." IOSR Journal of Pharmacy (IOSRPHR), vol. 9, no. 9, 2019, pp. 01-022.