**CRISPR-cas9: A revolution in genome engineering**

**Mahender Miland Lakeshar\*, Jyoti Choudhary, Manoj kalwaniya, Kamal Sokhal, Parteek Balana, Sudip Solanki**

**\*Asst. Prof., Department of Veterinary Microbiology,**

**Sri Ganganagar Veterinary college, Sri Ganganagar**

[**Email-drmlakeshar@gmail.com**](mailto:Email-drmlakeshar@gmail.com) **Mb. - 9887822696**

Abstract:

Gene therapy has long held promise to correct and identify a variety of human, veterinary diseases and defects. Discovery of the Clustered Regularly-Interspaced Short Palindromic Repeats (CRISPR) hint of their existence first came in 1987, later on discovery continues. The mechanism of the CRISPR-based prokaryotic adaptive immune system (CRISPR-associated system, Cas9-protein), and its repurposing into a potent gene editing tool has revolutionized in the field of molecular biology, immunology and generated excitement for new and improved gene therapies. Additionally, the simplicity and flexibility of the CRISPR/Cas9 site-specific nuclease system has opened its widespread use in many biological research tools including bacteria, development of model cell lines, discovering mechanisms of disease, identifying disease targets, development of transgene animals and plants, and transcriptional modulation. Doudna and Charpentier review the history of genome editing technologies, including oligonucleotide coupled to genome cleaving agents that rely on endogenous repair and recombination systems to complete the targeted changes, self-splicing introns, and zinc-finger nucleases and TAL effector nucleases. The clustered regularly interspaced palindromic repeats (CRISPRs), and their associated (Cas) nucleases, were discovered from an adaptive immune system in bacteria. The development of the CRISPR-Cas system into a facile genome engineering tool that is revolutionizing all areas of molecular biology.

Keywords: CRISPR, Gene editing, ZFN, TALEN, Crispr Cas9 Mechanism

1. Introduction:

Genome editing is an insertion, deletion or replacement of any gene for eliminating or inducing specific and desired characters in genome. Three genome editing techniques are already present that are commonly used ZFNs, TALEN and CRISPR/ Cas9. The ZFN technology was first reported in the 1990s (Kim *et al*., 1996). Genome editing has been done previously with engineered nucleases, i.e. all three major classes of these enzymes—zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and engineered mega nucleases.

In general, the precise editing or regulation of genomic information of a gene at the DNA level requires the action of a molecular machine composed of two major parts: a DNA-binding domain that mediates sequence-specific DNA recognition and binding, and an effector domain that enables DNA cleavage or regulates transcription near the binding site. Creating a double-stranded break (DSB) in a gene by using a sequence-specific endonuclease can stimulate the DNA repair pathway and greatly increase the rate of gene modification at the desired sequence.

The gene gain/loss-of-function comprises through (i) the generation of double-stranded breaks (DSBs) in defined regions of the genome, (ii) correction/replacement of the defective endogenous genes or introduction of exogenous genes and (iii) DSB repair. DSBs in eukaryotes are repaired by one or two endogenous repair mechanisms i.e. non-homologous end joining (NHEJ) or homology-directed repair (HDR). In NHEJ is a primary cellular DSB repair mechanism that can occur at any phase of the cell reproduction cycle. In this the broken DNA strand re-ligated either directly or by including nucleotide insertion or deletion without a homologues DNA template which is from their cognate gene. In contrast, HDR typically occurs in late S- or G2-phase. It uses a homologous repair template to precisely repair the DSB. Therefore, insertion of exogenous DNA (knock-in) or replacement of genomic segments for knock in via HDR, it requires a short exogenous sequences DNA as donors based upon frequency of DSB.

At present the Zinc finger nucleases (ZFNs), Transcription activator-like effectors (TALENs) and CRISPR/Cas9 are utilized as programmable nucleases that generate a site specific DSBs with greater fidelity and specificity. To follow this DSBs repair, a donor DNA sequence (either dsDNA donor or ssDNA donor) made available to insert new gene. CRISPR systems were classified into six types (Type I-VI) that act as adaptive immune system in prokaryotes. Type II CRISPR system, CRISPR/Cas9 is one of RNA guided endonuclease that mostly used in gene editing. The genomic CRISPR locus is comprised of three components: the trans-activating CRISPR RNA (tracrRNA) gene, promoter, the Cas gene and the spacer sequences. These are transcribed into tracrRNA, Cas9 protein, and pre-crRNA. Following transcription it cleavages into tracrRNA and pre-crRNA which are stabilized by Cas9 and base pair. After that the RNase III enzyme convert the pre-crRNA into crRNA results in formation of active crRNA-guided endonuclease that is crRNA, tracrRNA, Cas9 complex. This three complex are required for initiation of editing in our choice of gene.

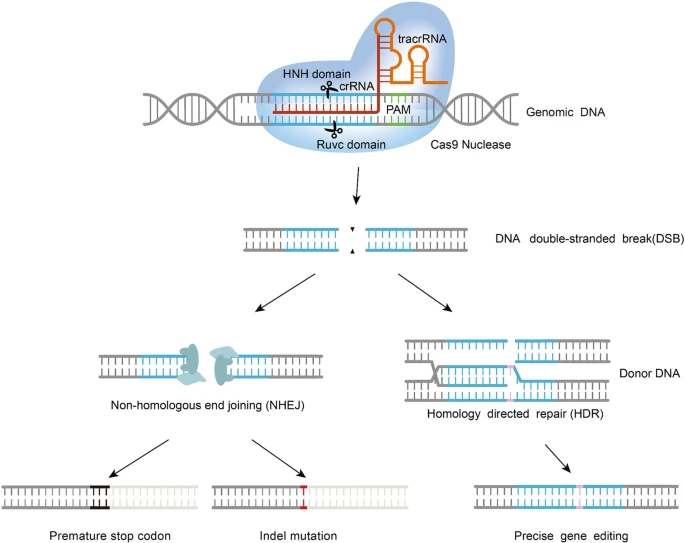
CRISP/Cas9 mainly comprises a nuclease termed Cas9 and a single-guide RNA (sgRNA) mostly made up of 20nucleotide that is complementary to the target sequence. On the presence of a protospacer-adjacent motif on the opposite strand, sgRNA recognizes the target strand via PAM and base pairing and thus guides Cas9 to bind to and cut target DNA sequence (Hsu *et al*., 2014). This result in DNA double-strand breaks that is repaired by either nonhomologous end joining (NHEJ) or homologous-directed repair HDR (Cong *et al*., 2013). NHEJ is error-prone in that case it may generate unpredictable indel mutations, whereas HDR can produce desired gene replacement (Bibikova *et al*., 2003). CRISPR-Cas9 is easy and safe to use because retargeting requires only the redesign of a sgRNA with the help of various online tool, which is much easier than the synthesis of a bulking guiding protein as required in conventional programmable nucleases such as zinc finger nuclease and transcription activator-like effector nuclease.

1. Type of CRISPR system

There are two classes of CRISPR systems, each containing multiple CRISPR types. Class 1 contains the CRISPR type I and type III systems that are commonly found in Archaea. Class 2 contains type II, IV, V, and VI CRISPR systems (Koonin *et al*. 2017). Type I and Type III CRISPR systems both utilize sets of Cas proteins and multiple effector molecules. In Type I systems, a multi-protein CRISPR RNA (crRNA) complex known as Cascade recognizes the target DNA, which is then cleaved by Cas3. In Type III systems, Cas10 assembles into a Cascade-like complex that recognizes and cleaves the target.

Type II CRISPR systems bear only one protein, Cas9 found in bacteria that is compatiable with involvement of RNaseIII enzyme, bind and stick the target DNA sequence (Makarova *et al*., 2011). The genomic CRISPR locus is comprised of three factors that are trans- cranking CRISPR RNA (tracrRNA) gene, the Cas gene, and the CRISPR reprise and spacer sequences (Chylinski *et al*., 2014). These are transcribed into different substitute i.e. tracr RNA, Cas9 protein, and pre-crRNA. Following recap, the tracrRNA and pre-crRNA are stabilized by Cas9 and base brace, and RNase III processes the pre-crRNA into crRNA by adhering it at the reprise (Deltcheva *et al*., 2011). This dependence on RNase III likely explains why Type II systems are set up in bacteria and not archaea because of RNase III isn't set up in archaea (Garrett *et al*., 2015). The crRNA tracrRNACas9 complex forms the active crRNA- guided endonuclease in CRISPR system (Chylinski *et al*., 2014).

The Cas9RNA complex interrogates with DNA in the cell, searching first for the applicable protospacer adjacent motif (PAM), which is specific because without correct identification of PAM sgRNA will not work. PAM is a short motif (5 ′- NGG- 3 ′ for Cas9 from Streptococcus pyogenes) conterminous to the target sequence (Chylinski *et al*., 2013). Upon recognition of the PAM sequence, the Cas9RNA complex unwinds the DNA from the first 10 – 18 nucleotides following the PAM sequence, nominated the seed region (Szczelkun *et al*., 2014) still, the HNH nuclease sphere of the Cas9 nuclease lobe cleaves the target beachfront while the RuvC- such like nuclease sphere of the Cas9 α- spiral lobe cleaves the non-target or off target. (Jinek *et al.,* 2014; Nishimasu *et al.,* 2015).



1. SgRNA

Single Guide RNA (SgRNA) a chimeric RNA composed of crRNA and tracrRNA, connected by a short RNA linker and can be designed by various online tool. The guide RNA is a specific RNA sequence that recognizes the target DNA region of interest through PAM and directs the Cas nuclease there for cutting (Paul *et al.,* 2019). The sgRNA is made up of two parts: (a) crispr RNA (crRNA)- a 17-20 nucleotide sequence complementary to the target DNA of host, (b) tracr RNA- binding scaffold optimate for the Cas nuclease enzyme (Zhang *et al*., 2016)

1. CAS9 Protein:

Cas (CRISPR-associated protein) directed to the specific DNA locus by a gRNA, where it makes a double-strand break by HNH & RuvC domain (Doudna *et al*., 2014). Cas9 is a nuclease, an enzyme specific for cutting DNA when complexed with both crRNA and tracrRNA. The most commonly used which are in trade i.e. Cas9 for genome engineering has been adapted from the type II-A CRISPR system from *Streptococcus pyogenes* (Sp). Simplicity of Sp cas9 is that it has simple PAM (NGG, or a weaker NAG, where N is any nucleotide, which has length of upto1368 AA). Cas9 has Bi-lobed structure having active sites and two grooves for binding of nucleic acid (DNA) i.e. (i) REC (large recognition lobe) and (ii) NUC (small nuclease lobe), Both are connected by a helix bridge. Specificity of Cas9 revealed insight is REC. NUC has 2 domains. (a) RuvC & HNH and (b) PAM interacting domain (PI). (Guan *et al*., 2018)

1. PAM (PROTOSPACER ADJACENT MOTIF):

PAM is a short signature sequence (of 2–5 nucleotides) that flanks the protospacer region (the target DNA sequence that is complementary to the spacer of the CRISPR RNA) in invading host DNA and are specifically recognized the particular sight and required by Cas9 for DNA cleavage (Van Der Oost *et al*., 2014). The sequence of the PAM influences the activity of SgRNA. The very much important thing is that protospacer-adjacent motif (PAM) is been strictly required to be immediately next to the 3′ end of the target sequence.

NHEJ dominates DNA repair occurs during G1, S and G2 phases of cell cycle whereas HDR is restricted to late S and G2 phases when DNA replication is completed inside cell and sister chromatids are available to serve as repair templates. Because HDR requires template DNA. By using repair process Precise deletion, Base substitution, Insertion of coding sequence of interest can be introduced to our desired genome of host (Zhang et al., 2017). It can be mediated by ssODNs (Single stranded oligodeoxy-nucleotides), Plasmid delivery, Double cut donors, PCR products and liposomes in cell lines (Hisano *et al*., 2015). The length of each homology arm is dependent on the size of the change being introduced, there is particular guideline. Where we have required if larger insertions, we need longer homology arm (Huang *et al*., 2015). ssODNs- used as a DNA template, homology of 30-70 base pair on either arm is sufficient for high-level HDR and it must be complementary. When using the length of homology arm 60bp, the HDR efficiency is decreased about 19% while 90 bp homology arm used, it can increase efficiency up to 20-30% (Cristea *et al*., 2013). For larger fragment insertion the homology arm is 0.2-0.8kb up to 2kb in plasmid mediated delivery (Byrne *et al*., 2015). There are various protocols to enhancing the HDR efficiency by using Chemical and Genetic Modulation, Timed Delivery of the CRISPR-Cas9 System (Yang *et al*., 2016), using Overlapping Sequences (Irion *et al*., 2014), using Modified Cas9 (Komor *et al*., 2016). Rational design of homology repair arms through online software templates strongly enhances HDR efficiency (Renaud *et al*., 2016). By using a linear repair mechanism template with homologous flanks in zebrafish, HDR efficiency can increase by almost 10-fold (Irion *et al*., 2014).

The Double Stranded break is either lethal for smaller insertion the homology arm (HA) or should be 30-50 nucleotide and for large DNA insertion the homology arm should be 400-1000bp (Byrne *et al.,* 2015). The length of each homology arm is dependent on the size of the change being introduced/required, with larger insertions require longer homology arms comparatively (Zhang *et al*., 2017).

1. Factors affecting efficacy of the CRISPR/Cas9 system

While the CRISPR/ Cas9 system has been demonstrated great efficiency for point-specific gene editing and other operations, there are several factors that impact its efficacity which must be addressed, especially if it's to be used for in vivo modal gene remedy. There are several factors which include target DNA point selection, sgRNA design, off- target slicing, prevalence/ effectiveness of HDR & NHEJ, Cas9 exertion, and the system of delivery. As delivery remains the most important factor for use of CRISPR for in vivo operations, sweats addressing other factors will be compactly epitomized then.

1. Conclusion

In a brief time, CRISPR/Cas9 has been modified to create different mutations either deletions or insertion in a variety of organisms including humans. As a result, this technology has the much potential to treat and cure diseases by editing the gene which are responsible for a particular disease. This system has made significant progress in the history of molecular biology and it has been widely used in genome editing to create gene knock-ins, knock-outs, and point mutations. It can be also used in hematopoietic stem and progenitor cells (HSPCs) transplantation therapy, plant and soil science and has an inherent benefit in the treatment of many bacterial and viral infection and it is known as a new type of antiviral therapy against various incurable viral infections. Pig is the most genetically modified livestock till date. CRISPR-Cas9 plays a key role in improvement of livestock as the most prominent gene-editing technology today.

References :

Bibikova, M., Beumer, K., Trautman, J. K., & Carroll, D. (2003). Enhancing gene targeting with designed zinc finger nucleases. *Science*, *300*(5620), 764-764.

Byrne, S. M., & Church, G. M. (2015). Crispr‐mediated gene targeting of human induced pluripotent stem cells. *Current protocols in stem cell biology*, *35*(1), 5A-8.

Byrne, S. M., & Church, G. M. (2015). Crispr‐mediated gene targeting of human induced pluripotent stem cells. *Current protocols in stem cell biology*, *35*(1), 5A-8.

Chylinski, K., Makarova, K. S., Charpentier, E., & Koonin, E. V. (2014). Classification and evolution of type II CRISPR-Cas systems. *Nucleic acids research*, *42*(10), 6091-6105.

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., & Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science*, *339*(6121), 819-823.

Cristea, S., Freyvert, Y., Santiago, Y., Holmes, M. C., Urnov, F. D., Gregory, P. D., & Cost, G. J. (2013). In vivo cleavage of transgene donors promotes nuclease‐mediated targeted integration. *Biotechnology and bioengineering*, *110*(3), 871-880.

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., & Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, *471*(7340), 602-607.

Doudna, J. A., & Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science*, *346*(6213), 1258096.

Guan, L., Zhu, S., Han, Y., Yang, C., Liu, Y., Qiao, L., & Lin, J. (2018). Knockout of CTNNB1 by CRISPR-Cas9 technology inhibits cell proliferation through the Wnt/β-catenin signaling pathway. *Biotechnology letters*, *40*, 501-508.

Hisano, Y., Sakuma, T., Nakade, S., Ohga, R., Ota, S., Okamoto, H., & Kawahara, A. (2015). Precise in-frame integration of exogenous DNA mediated by CRISPR/Cas9 system in zebrafish. *Scientific reports*, *5*(1), 8841.

Hsu, P. D., Lander, E. S., & Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, *157*(6), 1262-1278.

Irion, U., Krauss, J., & Nüsslein-Volhard, C. (2014). Precise and efficient genome editing in zebrafish using the CRISPR/Cas9 system. *Development*, *141*(24), 4827-4830.

Jinek, M., Jiang, F., Taylor, D. W., Sternberg, S. H., Kaya, E., Ma, E., & Doudna, J. A. (2014). Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science*, *343*(6176), 1247997.

Koonin, E. V., Makarova, K. S., & Zhang, F. (2017). Diversity, classification and evolution of CRISPR-Cas systems. *Current opinion in microbiology*, *37*, 67-78.

Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., & Koonin, E. V. (2011). Evolution and classification of the CRISPR–Cas systems. *Nature Reviews Microbiology*, *9*(6), 467-477.

Nishimasu, H., Cong, L., Yan, W. X., Ran, F. A., Zetsche, B., Li, Y., & Nureki, O. (2015). Crystal structure of Staphylococcus aureus Cas9. *Cell*, *162*(5), 1113-1126.

Paul, A., Bharati, J., Punetha, M., Kumar, S., Vidyalakshmi, G. M., Chouhan, V. S., & Sarkar, M. (2019). Transcriptional regulation of thrombospondins and its functional validation through CRISPR/Cas9 mediated gene editing in corpus luteum of water buffalo (Bubalus bubalis). *Cell Physiol Biochem*, *52*(3), 532-52.

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

Van Der Oost, J., Westra, E. R., Jackson, R. N., & Wiedenheft, B. (2014). Unravelling the structural and mechanistic basis of CRISPR–Cas systems. *Nature Reviews Microbiology*, *12*(7), 479-492.

Zhang, R., Miner, J. J., Gorman, M. J., Rausch, K., Ramage, H., White, J. P., & Diamond, M. S. (2016). A CRISPR screen defines a signal peptide processing pathway required by flaviviruses. *Nature*, *535*(7610), 164-168.