Paper: **Therapeutic applications of CRISPR/Cas9 system in gene therapy**
Hasan Mollanoori.Shahram Teimourian

## 1. What is the significance of this research topic?

The CRISPR/Ca9 technology is one of the most significant breakthroughs in the last 10 years in genome modification. Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR)-Cas system is part of the immune system found in some bacteria and archaea. When a virus invades these organisms, their DNA is "captured" and incorporated in the host DNA as spacers within the CRISPR array, each spacer is separated from the other by repeat units. The next step is transcription of CRISPR repeats and spacers into CRISPR RNA (crRNA). The last step is the activation of the Cas9 protein, which upon conformational change, finds the target DNA which then is followed by the attachment using base pair complementarity of the crRNA to viral RNA; and the cleavage of the foreign RNA complex and its elimination. Cas9 nuclease activity can be easily replaced by a sequence of interest making Cas9 compared to other DNA-binding platforms (meganucleases, zinc fingers (ZF), transcription activator-like effectors (TALEs)) simple and very efficient. Targeting efficiency for Cas9 system has been reported up to > 70% compared to ZFNs and TALENs which could only achieve efficiencies raging from 1% to 50% for similar species.

Another important advantage of the CRISPR-Cas9 system, is its ability to cleave multiple distinct targets sequences in parallel, inducing multiple mutations in different genes which extend the possibilities for modeling complex diseases when using mice from single gene knockouts without lengthy breeding strategies and with less animal sacrifice. Other reasons to add multiples gRNAs are: 1) using dual nicknases to generate a knockout or edit to reduce off-target activity, and 2) deleting large region of the genome. Since its first application in genome engineering, CRISPR-Cas technology has continued to evolve, leading to many more discoveries:

- Type VI CRISPR systems include enzymes (Cas13) that target RNA without the need for spacer sequence and can be used to reduce RNA levels.
- In Epigenetic modifications (the combined genetic modifications across the genome), a catalytic dead dCas9 protein is fused to a variety of epigenetic modifier; that is the genes which modify the epigenome through DNA methylation, alteration of the structure of the chromatin or its posttranslational modification. Fusing dCas9 to an epigenetic modifier can repress or activate transcription. One major advantage of this combination is its reversibility once the effector is inactivated from the system.
- CRISPR-Cas9 has been used for fluorescent tagging of protein, whereby a fluorescent protein such as GFP is attached to the protein of interest, to provide insight for the protein's function. Also, gRNAs can be fused to protein-interacting RNA aptamers to visualize targeted genomic loci, as well detecting the chromatin dynamics in living cells. The CRISPR tagging system is compatible with fluorescence microscopy; and produces less artifacts attributed to exogeneous overexpression of a protein fused to a fluorescent tag.
- CRISPR for locus-specific genomic DNA purification: the locus-specific purification of genomic regions enables the isolation of a genomic region of interest in vivo or in vitro for identification of bound interacting proteins to provide insights into pathogenesis of diseases

caused by abnormal gene regulations. The CRISPR DNA purification utilizes DNA-binding ability of catalytically inactive dCas9 enzyme and chromatin immunoprecipitation (ChIP) to isolate telomeres or centromeres for identification of their associated proteins.

**Who is working in this area?**

3 major players in commercializing CRISPR-gene editing are Intellia Therapeutics, Editas Therapeutics and CRISPR Therapeutics, companies which were started by scientits as a spin-off of their research. Other significant players in the CRISPR field are the Korean company Toolgen, Millipore Sigma, Vilnius University, Cellectis, the French company which has been developing CAR-T cell therapies, DowDuPont, MPEG-LA. The Broad Institute of MIT and Harvard has been granted 31 CRISPR patents, other patents have been allowed to UC Berkely, University of Vienna and Emmanuelle Charpentier.

Following the CRISPR timeline, the top research centers in CRISPR are: Broad Institute, UC Berkeley, Max Planck Institute, Vinius University, and many other academic institutions. Interestingly, in 2021, the United States have published the most CRISPR related papers, but China is catching up.

**2. What methods are used to study the concepts described in the paper?**

Among the different proteins involved in the CRISPR complex (Cas enzymes), Cas9 is the only enzyme within the Cas gene cluster that plays a role in locating and DNA cleavage. The Cas9 protein has 6 domains:

- **Rec 1** and **Rec 2** domain binds the complementary region of the guide RNA. Rec1 role is essential compared to Rec 2 for the binding of repeat/anti-repeat target DNA.
- **Bridge helix (NH)** arginine-based structure which modulates target DNA cleavage and mismatch tolerance.
- **Photospacer Adjacent Motif (PAM)-Interacting (PI)** domain confers PAM specificity, and is responsible for initiating binding to target DNA.
- **HNH and RuvC** domains are nuclease domains that cut single stranded DNA (specifically between the 3$^{rd}$ and 4$^{th}$ nucleotides from the PAM).

A single guide RNA (sgRNA) can be engineered by fusing a crRNA containing the target DNA sequence to a noncoding trans-activating crRNA (tracrRNA) to activate the Cas9 protein. The guide RNA forms a T-shape comprised of one tetraloop and 2 or 3 stem loops, and is constructed to have a 5' end complimentary to the target DNA sequence. One key factor in CRISPR-Cas9 efficiency is the Pam binding step which allows the Cas9 protein to quickly screen for potential target with appropriate PAM before melting. Another key advantage of CRISPR-Cas9 system is its flexibility, several studies have shown that single, mismatches at the 5' end of the sgRNA are well-tolerated, double mismatches can still result in cleavage for some sequences, and beyond this number sgRNA activity is suppressed; Cas9 can also tolerates up to 5 base mismatches within the protospacer region, or a single base difference in the PAM sequence.

CRISPR/Cas9 is used to create double strand breaks (DSBs) to perform desired indels or knock out existing genes in the target DNA strand. Homologous directed-repair (HDR) and non-homologous end joining (NHEJ) are the two major pathways to resolve DSBs. Between the two,

HDR is favored as it requires higher sequence homology between the severe and donor strands of DNA and reduce the risks of genomic instability or cell deaths. Strategies have been developed to promote HDR over NHEJ such as inhibiting NHEJ pathway using inhibitors (CYREN) and HDR factors (e.g., Rad18) to activate HDR. HDR methods are further categorized into conservative methods: double-strand break repair (SDBR), synthesis-dependent strand-annealing (SDSA), break-induced repair (BIR) and non-conservative: single-strand annealing (SSA) pathway. Insertion sites of the gene modification have to be less than 1-bp away from the DSB. Also, HDR template have to be designed to prevent the Cas9 enzyme to keep cutting and repairing beyond the targeted sequence.

3. **Does the review article lead to new questions or hypotheses in this technical area (by the authors, by other researchers)?**
   There are many existing challenges faced by CRISPR/Cas9 system before it can be used commonly in clinical applications:
   - Recent studies have focused in reducing the rate at which Cas9 binds to unintended genomic sites for cleavage, termed as off-target events. More than 3 mismatches between target sequences and 20 nucleotides of gRNA can cause off-target effects; whereas 4 mismatches in PAM induces off-target effects.
   - Off-target effects could result in loss of gene function, ultimately leading to cancer or severe problems in the receiving organism. Decreasing enzymatic concentration improves on-to off target ratio at the expense of the efficiency of on-target cleavage. Duration of Cas9 expression is another factor to impact off-target activity. Due to its larger genome, there are high chances of off-target events in mammalian genome than bacteria. To address this issue, researchers have developed in vivo/vitro biochemical assays and in-silico approaches to detect and quantify off target effects to increase gene editing efficiency. GC content, and length of the gRNA can mitigate off-target effects. Improved variants SpCas9 and other Cas9 orthologous have been designed to reduce off-target mutations but also to overcome the size limitation when using adeno-associated virus (AAV) for delivery. One strategy has been to find smaller Cas9 orthologs; about 1000 amino acids in size (CjCas9, NmeCas9, Cas12b, dCs13b) or to package the gene payload in two AAV. When a cell is transduced by both AAVs, the full gene or protein is reconstituted. Other mitigation methods include non-viral delivery machinery (RNP complex), genome editing without DSBs (prime editing is a gene-editing technique which does not use DSB), broader PAM recognition sites, or anti-CRISPR proteins (CRISPR/Cas inhibitors).

4. **What are some practical applications of the research discussed in the article?**

**6. How does this topic relate to other areas of cell biology, bioengineering, or medicine?**

https://www.jax.org/news-and-insights/jax-blog/2014/march/pros-and-cons-of-znfs-talens-and-crispr-cas
https://sites.tufts.edu/crispr/crispr-mechanism/rna-binding/
https://star-protocols.cell.com/protocols/934