Main Submission

CRISPR and TALEN: Facilitating Tailored Genomes of the Future

Caleb Seward, Harry Li, John-Paul Oliveria

Division of Respirology, Department of Medicine, McMaster University

Introduction

Genome modification allows scientists to directly manipulate gene expression through the utilization of novel genetic editing techniques. Notable techniques include clustered interspaced short palindromic repeats (CRISPR) and CRISPR-associated systems (Cas9), and transcription activator-like effector nucleases (TALEN). This article will discuss the strengths and limitations in both design and execution of the aforementioned techniques, and their applicable utility.

CRISPR

The CRISPR method contains viral genetic information within spacer DNA sequences, obtained following viral exposure (1). Cas genes code an endonuclease and helicase protein complex, which initiates double-stranded breaks in highly specific regions in the genome utilizing the spacer sequences as "guiding machinery." These breaks ultimately inactivate the gene of interest, rendering the viral infection ineffective (1). Thus, archaeal and bacterial cells are able to store viral genetic information within their spacer sequences as an adaptive mechanism to avoid viral infiltration (2).

In order to selectively replace a gene of interest, three main components need to be added into a cell: the Cas9 protein, guiding RNA (gRNA) specific for the gene of interest, and single-stranded DNA (ssDNA) containing the desired genetic information (3). When all three components are present, Cas9 will cut the gene at a particular location and replace it with the ssDNA. Cellular machinery then reverse-transcribes the gene segment. The double-stranded breaks created by Cas9 are repaired through non-homologous end-joining or homology-directed repair, thereby incorporating the added genetic information (4).

TALEN

TALEN has a DNA-binding domain and a nuclease effector domain. The DNA-binding domain is composed of 18 subunits, known as transcription activator-like effector (TALE) repeats, each recognizing one DNA base. Every repeat consists of 33 or 34 amino acids, and the 12th and 13th amino acids are known as the repeat variable diresidue (RVD), which varies between each TALE (5). There are 4 common RVDs, each preferentially binding to one of the four DNA bases, and each TALE is directly engineered to target one particular base. There are 2 sets of 18 TALEs bound to opposite sides of the target DNA region, hence one complete DNA-binding domain is able to recognize a 36 base-pairing sequence (6). Thus, the DNA-binding domain can be engineered to specifically target known sequences in the human genome.

Regarding its nuclease activities, TALEN utilizes Fok1 enzymes adjoined to each end of the TALE repeats. Two Fok1 enzymes will line up at opposite sites of the target DNA sequence, dimerize, and cleave out a particular section of nucleotides. Since site recognition is carried out by the highly accurate TALE repeats, which are engineered to target very specific sequences, this mechanism is applicable to manipulate any sequence in the human genome (7). Other endonucleases, such as EcoR1, are restricted to cut sequences 5 to 6 base pairs long. Therefore, TALEN has been proven to be a great tool in applications that require flexibility (8).

Applications

The field of genetics and genomics is undergoing a transformative phase with the emergence of genome modification technologies, which have unique advantages and drawbacks. For example, when comparing efficiency and ease of design, the CRISPR technique excels over TALEN. However, in terms of specificity and accuracy, the TALEN technique excels over CRISPR.

CRISPR is a more efficient genome editing technique (9), since the CRISPR system only requires the creation of a single gRNA sequence that is both significantly smaller in size and can recognize multiple loci (6). This efficiency makes CRISPR highly suitable for time-sensitive genetic

manipulations. An example of this would be in agriculture, where changing crop-growing conditions requires quick adaptability; thus, the CRISPR system could be used to quickly and efficiently modify the genome of plants to be able to withstand poor growing conditions.

However, the CRISPR system is prone to off-target activity, resulting in lower accuracy and specificity for targeted genome modifications. This is attributed to the gRNA found in Cas9 proteins, which contributes to the specificity and regulation of the CRISPR/Cas9 system (2). This gRNA can accommodate up to 5 mismatched base pairs per target site, making CRISPR prone to off-target activity (2). This is a concern, especially in the context of operations that require a high degree of accuracy. Alternatively, TALEN is a technique that excels in both accuracy and specificity. The DNA binding domain can be engineered to recognize a 36 base pair-long sequence, which is probably unrepeated in the human genome, thus significantly reducing potential errors (10). In addition, TALEN designs are very flexible, albeit lengthy, which may allow this technique to be applied to any sequence in the human genome. TALEN has been successfully implemented to correct genetic mutations that cause disease, engineer stably modified human stem cells, and hone the immune system to combat cancer (8).

As expected with many novel technologies, TALEN also has limitations, notably in its ease of design. At a single target site, 36 subunits (18 on each side) are required to be assembled, with each subunit consisting of 33-34 amino acids. In other words, it would be necessary to design a different DNA-binding domain of more than 1000 amino acids for each target site, which can be a highly tedious and technical process (6).

Conclusion

Currently, genome editing technologies are nowhere near perfect. However, with time, they are sure to be utilized in novel high-risk procedures. For operations demanding high accuracy, such as human genome editing, it would be best to use TALEN. However, for cruder genetic engineering, CRISPR would be a more effective technique. Nevertheless, gene therapy is an evolving field, and extensive investigation into both techniques must be conducted in order to elucidate their efficacious utility in varying fields of science and bioengineering. ■

References

1. Doudna JA, Charpentier E. The new frontier of genome engineering with CRISPR-Cas9. *Science*. 2014;346(6213):1258096.
2. Mali P, Yang L, Esvelt K, Aach J, Guell M, DiCarlo J et al. RNA-guided human genome engineering via Cas9. *Science*. 2013;339(6121):823-826.
3. Li D, Qiu Z, Shao Y, Chen Y, Guan Y, Liu M et al. Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. *Nat Biotechnol*. 2013;31(8):681-683.
4. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, WuX, Jiang W, Marraffini LA, Zhang F. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013;339(6121):819-823.
5. Bleijs DA. *TALEN (Transcription activator-like effector nuclease) as gene editing tool.* Available from: http://www.genetherapynet.com/gene-editing-tools/talen.html [Accessed 2017 Feb 13].
6. Boettcher M, McManus MT. Choosing the right tool for the job: RNAi, TALEN, or CRISPR. *Mol Cell*. 2015;58(4):575–585.
7. Ain QU, Chung JY, Kim YH. Current and future delivery systems for engineered nucleases: ZFN, TALEN and RGEN. *J Control Release*. 2015;205:120-127.
8. Joung JK, Sander JD. TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol*. 2012;14(1):49–55.
9. Maruyama T, Dougan SK, Truttmann MC, Bilate AM, Ingram JR, Ploegh HL. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat Biotechnol*. 2015;33(5):538-542.
10. Guilinger JP, Pattanayak V, Reyon D, Tsai SQ, Sander JD, Joung JK, et al. Broad specificity profiling of TALENs results in engineered nucleases with improved DNA-cleavage specificity. *Nat Methods*. 2014;11(4):429–435.

List of Abbreviations

CRISPR – clustered regularly interspaced short palindromic repeats

Cas9 - CRISPR-associated systems

gRNA - guiding RNA

RVD - repeat variable diresidue

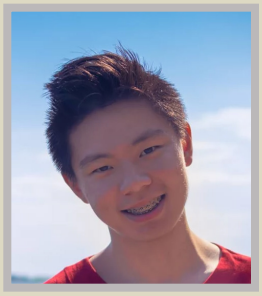
TALE - transcription activator-like effector

TALEN - transcription activator-like effector nuclease

ssDNA – single-stranded DNA

Caleb Seward

Caleb is a 3rd year undergraduate student in McMaster University's Bachelor of Life Sciences program. He currently volunteers at the McMaster Cardio-Respiratory Research Lab under the mentorship of Dr. Gail Gauvreau and John-Paul Oliveria. Caleb is interested in immunology, molecular genetics, and exercise physiology.



Harry Li

Harry is a 1st year undergraduate student in McMaster University's Health Sciences program. He currently volunteers at the McMaster Cardio-Respiratory Lab under the mentorship of Dr. Gail Gauvreau and John-Paul Oliveria. He particularly enjoys learning about molecular biology, immunology, and genetics.

John-Paul Oliveria

John-Paul is a 4th year Doctor of Philosophy candidate in the Department of Medical Sciences at McMaster University. He began his MSc in the McMaster Cardio-Respiratory Research Lab in 2011 and transferred to his PhD in 2013. His research focuses on the biology of B cells (IgE+ B cells and regulatory B cells) and the pathogenesis of allergic asthma. In the fall of 2017, he will be commencing a post-doctoral scholar appointment at Stanford University.