

## Advances in Genome Engineering Approaches

Arvind Chhabra\*

Department of Medicine, University of Connecticut Health Center, Farmington, Connecticut, USA

### Abstract

Targeted genome editing is essential for functional characterization of a gene of interest. Targeted gene inactivation via homologous recombination made it feasible to create gene knockout animal models to ascertain the physiological role of the target genes; however, lower efficiency of site specific insertion of the genetically modified construct through homologous recombination has limited a wider applicability of this approach. Development of targeted gene knockdown through RNA interference (RNAi) offered a cost effective, high-throughput alternative to homologous recombination, however, RNAi-mediated gene knockdown is incomplete, produces experiment to experiment variation, and could provide only a temporary inhibition of the gene function. Development of genome engineering methodologies utilizing nucleases linked to the guide sequences targeting a gene of interest, such as Zinc Finger Nucleases (ZFN), Transcription Activator like Effector Nucleases (TALEN) and Clustered Palindromic Repeats (CRISPR), are quite encouraging. A brief overview of recent advances in genome engineering approaches is provided with their respective advantages and limitations.

**Keywords:** Genome engineering; MicroRNAs; ZFN; TALEN; CRISPR

**Abbreviations:** RNAi: RNA interference; ZFN: Zinc Finger Nucleases; TALEN: Transcription Activator like Effector Nucleases; CRISPR: Clustered Palindromic Repeats

### Small RNA-Mediated Genome Engineering

RNA has long been considered a messenger of the genetic code imprinted in the DNA of a cell. However, the discovery of first small RNA based gene silencing of *lin-4* gene in *Caneorhabditis elegans* [1,2], offered novel insights towards a regulatory role of RNA in the cellular physiology, and set the foundation for the identification of small RNAs modulating gene expression. It is now well known that eukaryotes produce many different kinds of small RNAs (19-30 nucleotides in length) that play essential roles in regulation of developmental cascades as well as cellular stress conditions, for example viral infections, through multiple mechanisms, such as direct mRNA degradation or translational repression. Among the most common small RNAs include small interfering RNAs (siRNA) and microRNAs. After initial processing in the nucleus by Drosha, precursors RNAs (pre-RNAs) are transported to the cytoplasm, where Dicer cleavage generates mature microRNAs and siRNAs [3-5]. While miRNAs are generated from the double stranded RNA (dsRNA) region of the hairpin-shaped precursors, siRNAs are generated from long ds RNAs. These double-stranded products then assemble with Argonaute proteins in a protein complex known as the RNA-induced silencing complex (RISC) [3-5], such that their one strand is preferentially selected to guide the sequence-specific modulation of complementary target mRNA through mRNA cleavage, translational repression, or cleavage independent mRNA decay [3-5]. Systematic studies characterizing the sequence requirements for developing an efficient siRNA resulted in development of tools for creation of custom siRNAs for targeting gene of interest [6,7], and this coupled with the development of vector as well as non-vector based small RNA delivery methods, made siRNAs a preferred method to ascertain the functional relevance of any given gene of interest, *in-vitro* as well as *in-vivo*. However, RNAi-mediated approaches provide an incomplete ablation of target gene and potential off-target effects make it difficult to draw accurate conclusions.

Although several natural siRNAs have been reported in different eukaryotic species, in mammals, microRNAs represent the major natural small RNAs that have been identified. Among the

approaches used for the identification of microRNAs include northern hybridization, microarray based approaches and small RNA cloning etc. Several microRNAs have been found to be expressed in embryonic stem cells as well as in distinct hematopoietic cell populations [8,9]. Many microRNAs have been associated with modulation of different development pathways [10,11]. MicroRNAs have also been suggested to modulate functional characteristics of host immune system, for example microRNAs have been identified in different effector T cell population with distinct functional profiles, such as Th1/Th2/regulatory CD4 T cells as well as CD8 effector T cell [12-14]. While on one hand these findings demonstrate an active role for microRNAs in modulation of cellular physiology, it also offers a unique opportunity for targeted genome engineering. Identification of naturally occurring microRNAs that can regulate specific physiologic pathways would be quite significant in developing microRNA based genome engineering methodologies. For example, identification of naturally occurring microRNAs that can regulate specific developmental pathways from human hematopoietic cells (HSC) can be utilized for generating cells of a specific lineage from engineered HSCs. Similarly, identification of microRNAs regulating the generation of specific effector function in T cells can help generate immune effectors with pre-defined functional profile. However, identification of specific microRNAs that can meet these functional end points without any off-target effects is essential to achieve these goals.

### Nuclease-Mediated Genome Engineering Approaches

Targeted genome editing in embryonic stem cells through homologous recombination revolutionized the field by allowing the creation of different knock-in/knockout mice strains to ascertain the functional relevance of a gene of interest [15,16], however, as

\*Corresponding author: Arvind Chhabra, Department of Medicine, University of Connecticut Health Center, Farmington, Connecticut, USA, Tel: 860-679-1447; E-mail: [arvindac@yahoo.com](mailto:arvindac@yahoo.com)

Received December 27, 2013; Accepted December 27, 2013; Published January 02, 2014

Citation: Chhabra A (2014) Advances in Genome Engineering Approaches. Adv Genet Eng 3: e107. doi:10.4172/2169-0111.1000e107

Copyright: © 2014 Chhabra A. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

mentioned before, low efficiency of site specific recombination limited a widespread applicability of this approach. To address this limitation, several nuclease based targeted genome engineering approaches such as Zinc finger nucleases (ZFNs), Transcription Activator Like Effector Nucleases (TALENs), Clustered Palindromic Repeats (CRISPRs) etc., have recently been developed that utilize a custom guide sequence for targeting a gene of interest coupled with a nuclease [17], discussed in details below:

### ZFNs-mediated genome engineering

Cys<sub>2</sub>-His<sub>2</sub>zinc-finger domain is among the most common types of DNA-binding motifs found in eukaryotes and represents the second most frequently encoded protein domain in the human genome. ZFNs utilize the DNA binding ability and specificity of zinc finger proteins (ZFPs), along with the nuclease function of FokI nuclease for targeted genome engineering. The zinc finger domain is comprised of 3-6 zinc finger modules linked together, with each module comprised of 30 amino acids and recognizing 9-18 base pairs. Since DNA cleavage activity of FokI nuclease functions only in dimer form, two zinc finger modules are created recognizing 9-18 base pairs on opposite sides of a target gene, allowing on target specificity. Binding on the zinc finger modules to the target sequence results in FokI dimerization, activation of its nuclease activity and double stranded break (DSB) in the target sequence spacer region. Once DSB is introduced, it leads to activation of cellular DNA repair mechanisms, such as the error prone non homologous end joining (NHEJ) or the DNA homology mediated repair by homologous recombination (HR), culminating in gene disruption by homologous template introduced or NHEJ mediated nucleotide insertion, deletion or frame-shift mutations. Among the key constraint in designing an effective ZFN is their natural preference for G-rich sequences. Furthermore, making a single designer DNA-binding domain is easier than joining two domains in an appropriate orientation, with correct spacing, to yield a functional nuclease. These technological challenges make in house development of an effective ZFN quite difficult for most laboratories.

### TALEN-mediated genome engineering

Transcription Activator Like Effectors (TALE) are natural DNA binding proteins of plant pathogenic bacterium, *Xanthomonas*, that utilize DNA binding domain comprised of 33-35 amino acid tandem monomer repeats, each recognizing a single base pair [18]. TALEN utilizes the DNA binding ability of TALEs with the cleavage function of FokI nuclease for targeted genome editing, with two custom FokI fused TALENs binding in the opposite directions of the target gene and resulting in dimerization of FokI and cleavage of the target DNA [17,19,20]. The key difference between ZFNs and TALENs is that each ZFN finger module (30 amino acids) recognizes three nucleotides while each TALEN repeat (33-35 amino acids) recognizes only 1 nucleotide. The DNA recognition specificity is determined by two hyper variable amino acids, known as the repeat variable diresidues (RVDs), at position 12 and 13. For example HD (His, Asp) targets Cytosine (C), NI (Asn, Ile) targets adenine (A), NG (Asn, Gly) targets thymine (T), and NN (Asn, Asn) targets guanine (G) and adenine (A). Following RVDs binding to the complementary sequences, FokI monomers on opposite TALENs dimerize to cause DSB at the spacer region. Repair of DSBs through externally introduced homologous elements can lead to site-specific insertions, while repair of DSBs through NHEJ pathway can lead to nucleotide deletions, insertions, causing frame shift mutations and gene disruption. While single base recognition requirement for TALEN-DNA binding repeat provides it greater flexibility than the triplet confined zinc-finger proteins, high sequence similarity between

each repeat (difference of only two RVD residues between TALEN repeat modules) present a challenge to assemble a functional TALEN. However, methods such as "Golden Gate" molecular cloning have made rapid assembly of TALENs possible [19] and several TALENs have been shown to produce site specific genome editing in different cell types.

### CRISPR/Cas9 based system for genome engineering

CRISPRs are short direct repeats of 21-47 nucleotides length interspersed with short intervening spacers in the bacterial genome. The CRISPR repeats are identical but spacers sequences vary. The CRISPR locus is surrounded by a cohort of CRISPR-associated (*Cas*) genes. The transcribed CRISPR RNA (crRNA) is processed and the mature crRNA binds to the Cas protein/nuclease and guides the complex to DNA target, complementary to the spacer region, to cause Cas-mediated cleavage, followed by DNA repair through NHEJ or HR, as in ZFNs and TALENs. Several types of CRISPR/Cas systems, comprised of small, non-coding CRISPR RNAs and a set of Cas proteins, have been identified in different bacterial species [17,21,22]. Of these, Type II CRISPR/Cas system uses less components than the others, such that it requires a small CRISPR RNA (consisting of palindromic repeats flanking sequence recognizing spacers) and a partially complementary trans-acting (trac RNA) that associate with a single Cas9 protein to mediated target degradation. Availability of CRISPR based dual recombinant vector system that expresses the target gene specific customized short crRNA/tracRNA under polymerase III promoter in one vector, and the Cas9 protein in the second vector, makes it feasible to design custom CRISPRs for targeting a gene of interest. This gives CRISPR an edge over technological challenges under ZFNs and TALENs, as it requires construction of only one variable vector encoding target specific crRNA, with Cas9 being the common moiety between different CRISPRs. However, additional studies are needed to ascertain the efficiency of CRISPR-mediated gene targeting approaches as well as their off-target effects.

### Concluding Remarks

MicroRNAs, ZFNs, TALENs and CRISPRs are efficient targeted genome engineering tools. Characterization of microRNAs that can facilitate the development of a specific cell lineage or imprint pre-defined functional characteristics in a cell of choice will have enormous translational implications. While limited studies with TALENs and CRISPRs improved genome engineering efficiency, additional studies are needed for establishing their toxicity and off-target effect profiles. However, despite these challenges, it is suffice to say that the field of targeted genome engineering has made significant progress in the last few years, and with systematic cross-disciplinary studies their potential could be translated into therapeutic approaches for the benefit of the society.

### References

1. Lee RC, Feinbaum RL, Ambros V (1993) The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75: 843-854.
2. Wightman B, Ha I, Ruvkun G (1993) Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* 75: 855-862.
3. He L, Hannon GJ (2004) MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* 5: 522-531.
4. Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. *Cell* 136: 215-233.

5. Shukla GC, Singh J, Barik S (2011) MicroRNAs: Processing, Maturation, Target Recognition and Regulatory Functions. *Mol Cell Pharmacol* 3: 83-92.
6. Yuan B, Latek R, Hossbach M, Tuschl T, Lewitter F (2004) siRNA Selection Server: an automated siRNA oligonucleotide prediction server. *Nucleic Acids Res* 32: W130-W134.
7. Paddison PJ, Silva JM, Conklin, Schlabach M, Li M, et al. (2004) A resource for large-scale RNA-interference-based screens in mammals. *Nature* 428: 427-431.
8. Houbaviy HB, Murray MF, Sharp PA (2003) Embryonic stem cell-specific MicroRNAs. *Dev cell* 5: 351-358.
9. Petriv OI, Kuchenbauer F, Delaney AD, Lecault V, White A, et al. (2010) Comprehensive microRNA expression profiling of the hematopoietic hierarchy. *Proceedings of the National Academy of Sciences of the United States of America* 107: 15443-15448.
10. Zhao Y, Srivastava D (2007) A developmental view of microRNA function. *Trends BiochemSci* 32: 189-197.
11. Baltimore D, Boldin MP, O'Connell RM, Rao DS, Taganov KD (2008) MicroRNAs: new regulators of immune cell development and function. *NatImmunol* 9: 839-845.
12. Baumjohann D, Ansel KM (2013) MicroRNA-mediated regulation of T helper cell differentiation and plasticity. *Nat Rev Immunol* 13: 666-678.
13. Dudda JC, Salaun B, Ji Y, Palmer DC, Monnot GC, et al. (2013) MicroRNA-155 is required for effector CD8+ T cell responses to virus infection and cancer. *Immunity* 38: 742-753.
14. Khan AA, Penny LA, Yuzepolskiy Y, Sarkar S, Kalia V (2013) MicroRNA-17~92 regulates effector and memory CD8 T-cell fates by modulating proliferation in response to infections. *Blood* 121: 4473-4483.
15. Thomas KR, Capecchi MR (1987) Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* 51: 503-512.
16. Austin CP, Battey JF, Bradley A, Bucan M, Capecchi M, et al. (2004) The knockout mouse project. *Nat Genet* 36: 921-924.
17. Gaj T, Gersbach CA, Barbas CF 3rd (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol* 31: 397-405.
18. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, et al. (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326: 1509-1512.
19. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, et al. (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res* 39: e82.
20. Clark KJ, Voytas DF, Ekker SC (2011) A TALE of two nucleases: gene targeting for the masses? *Zebrafish* 8: 147-149.
21. Mali P, Yang L, Esvelt KM, Aach J, Guell M, et al. (2013) RNA-guided human genome engineering via Cas9. *Science* 339: 823-826.
22. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, et al. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337: 816-821.