Methods 69 (2014) 85–93

Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

Generation of knockout mice using engineered nucleases

Young Hoon Sung [a,1], Young Jin [a,1], Seokjoong Kim [b], Han-Woong Lee [a,*]

[a] Department of Biochemistry, College of Life Science and Biotechnology, Laboratory Animal Research Center, Yonsei University, Seoul 120-749, Republic of Korea
[b] ToolGen, Inc., Byucksan Kyoungin Digital Valley 2-Cha, Geumcheon-Gu, Seoul 153-023, Republic of Korea

ARTICLE INFO

Article history:
Received 3 November 2013
Revised 1 February 2014
Accepted 6 February 2014
Available online 18 February 2014

Keywords:
Engineered nuclease
ZFN
TALEN
RGEN
NHEJ
Knockout mouse
Microinjection

ABSTRACT

The use of engineered nucleases in one-cell stage mouse embryos is emerging as an efficient alternative to conventional gene targeting in mouse embryonic stem (ES) cells. These nucleases are designed or reprogrammed to specifically induce double strand breaks (DSBs) at a desired genomic locus, and efficiently introduce mutations by both error-prone and error-free DNA repair mechanisms. Since these mutations frequently result in the loss or alteration of gene function by inserting, deleting, or substituting nucleotide sequences, engineered nucleases are enabling us to efficiently generate gene knockout and knockin mice. Three kinds of engineered endonucleases have been developed and successfully applied to the generation of mutant mice: zinc-finger nuclease (ZFNs), transcription activator-like effector nucleases (TALENs) and RNA-guided endonucleases (RGENs). Based on recent advances, here we provide experimentally validated, detailed guidelines for generating non-homologous end-joining (NHEJ)-mediated mutant mice by microinjecting TALENs and RGENs into the cytoplasm or the pronucleus of one-cell stage mouse embryos.

1. Introduction

Because of the genetic and physiological similarities between humans and mice, genetically-engineered mouse (GEM) models, including gene knockout and knockin strains, provide us with an invaluable opportunity to gain molecular and physiological insights into orthologous human genes [1,2]. After completion of the human genome project, phenotypic analysis of GEM models has been recognized as an essential tool for uncovering *in vivo* gene functions. Furthermore, GEM models have been continuously developed to generate orthologous models of human diseases in mice. As appreciated by the Nobel Prize in Physiology or Medicine 2007, gene targeting in embryonic stem (ES) cells has been central to the genetic manipulation of the mouse genome. Furthermore, the International Knockout Mouse Consortium (IKMC) was assembled to facilitate the generation and phenotypic analysis of gene knockout mice [3]. However, in spite of many public resources supplying targeting vectors and gene-targeted ES cell clones, the generation of live knockout mice is still rate-limiting (*e.g.*, the generation of chimeric mice and germ-line transmission).

Currently, engineered nucleases provide us a shortcut for modifying the mouse genome, possibly avoiding the use of ES cells and thus the generation of chimeric mice. To date, three kinds of engineered nucleases have been developed: zinc-finger nucleases (ZFNs), transcription activator-like effector (TALE) nucleases (TALENs) and RNA-guided endonucleases (RGENs) or clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) systems [4–6]. After *in vitro* validation in cultured cells, these engineered nucleases have been extensively studied and allow the direct generation of live GEMs without producing chimeras.

ZFNs and TALENs have similar structures. The DNA-binding domain, which mainly consists of zinc-finger or TALE arrays, is connected to the non-specific cleavage domain of Fok I, a type II endonuclease through a linker [7,8]. Because the Fok I cleavage domains are activated by heterodimer formation [9], a pair of ZFNs or TALENs should be designed to recognize 5′ and 3′ sequences with a spacer in which the target sequence resides [10,11]. Each ZFN contains three to six zinc finger arrays that are designed to specifically recognize the target region; the protein modules of zinc finger are typically composed of 30 amino acids that coordinate a zinc ion by two cysteine and two histidine residues [10]. TALEs, originally discovered in the plant pathogen *Xanthomonas*, are transcriptional activators that specifically bind to the promoter region of plant genes and regulate their expression during pathogenesis [12]. The central DNA-binding domains of TALEs are composed of tandem repeats of 34 amino acid residues that are highly conserved except for repeat variable di-residues (RVDs) at positions 12 and

* Corresponding author. Fax: +82 2 2123 8107.
E-mail address: hwl@yonsei.ac.kr (H.-W. Lee).
[1] These authors contributed equally to this paper.

http://dx.doi.org/10.1016/j.ymeth.2014.02.009
1046-2023/

13 that determine the sequence specificity (*e.g.*, HD, NG, NI, and NN for C, T, A, and G, respectively) [5,11], thus allowing us to easily design an effective engineered nuclease. In addition, compared to ZFNs, which are constrained by their preference for GNN triplets and a lower affinity for AT-rich target sequences [13], the TALE codes can be used to design more specific engineered nucleases.

Distinct from ZFNs and TALENs, CRISPR/Cas systems or RGENs are derived from a bacterial adaptive immune system, and are originally composed of Cas9, a protein component, and two RNA components, CRISPR RNA (crRNA) and trans-activating crRNA (tracrRNA) [14].

The RGEN target sequence is X_{20}NGG, where NGG represents a protospacer adjacent motif (PAM) that is recognized by Cas9 protein [15,16]. crRNA recognizes the target sequence by base-pairing to a 20-bp complementary sequence upstream of PAM, and thus determines the specificity of the Cas9 protein complex [17]. Cas9 has two distinct endonuclease motifs, HNH endonuclease and RuvC-like domains, and induces a DSB 3-bp upstream of the PAM sequence [15]. A single-guide RNA (sgRNA), an RNA hybrid of CRISPR RNA (crRNA) and trans-activating crRNA (tracrRNA), is also functional [15]. These characteristics define CRISPR/Cas systems as easily reprogrammable RNA-guided endonucleases, RGENs [17–19].

In principle, engineered nucleases are primarily designed to induce double strand breaks (DSBs) at a specific sequence. It is well known that DSBs are generally repaired by non-homologous end-joining (NHEJ) [20]. NHEJ-mediated DNA repair is error-prone and thus frequently introduces insertions, deletions, and/or substitutions that lead to frameshift mutations. To date, diverse mutant animals have been generated by NHEJ induced by ZFNs, TALENs, or RGENs (reviewed in [21]). The first successful application of ZFNs was the generation of *Mdr1a*, *Jag1* and *Notch3* knockout mice [22]. We recently reported *Pibf1* and *Sepw1* knockout mice created using TALENs [23]. Even though RGEN history is much shorter than that of ZFNs and TALENs, RGEN-mediated generation of knockout mice is already established. Shen et al. reported RGEN-mediated disruption of an eGFP transgene [24] and Wang et al. proved the high efficiency of RGENs by simultaneously targeting double genes including *Tet1* and *Tet2* [25]. We also generated *Foxn1* and *Prkdc* knockout mice using RGENs, which were delivered either as RNAs or recombinant Cas9 protein/sgRNA complexes, and demonstrated the germ-line transmission and phenotypic manifestations in F_1 offspring [26].

In contrast to NHEJ, HR-mediated DNA repair is error-free and is stimulated when homologous donor DNA templates (*e.g.*, double-stranded DNA and synthetic oligonucleotides) are available. HR-mediated gene targeting by ZFNs was first developed by Meyer et al. and achieved in 1.7–4.5% of targeted mouse embryos by co-injecting homologous donor DNA [27]. Wefers et al. successfully created knockin mice by co-injection of TALENs and synthetic oligodeoxynucleotides; validated heterozygous mutants were available within 18 weeks [28]. Yang et al. also recently generated mutant mice in which conditional and reporter constructs were integrated in the genome by using the CRISPR/Cas system [29]. Taken together, the evidence described above indicates that engineered nucleases can facilitate the generation of knockout mice by avoiding time-consuming and laborious use of ES cells. Based on the recent reports, here we provide detailed procedures for generating NHEJ-mediated mutant mice using TALENs and RGENs (Fig. 1).

2. Materials

2.1. Preparation of endonuclease components (TALEN mRNA and RGEN complexes)

2.1.1. Reagents

mMESSAGE mMACHINE T7 Ultra kit (Ambion, cat. No. AM1345) for TALEN and Cas9 mRNA synthesis.

MEGAshortscript T7 kit (Ambion, cat. No. AM1354) for sgRNA synthesis.

Syringe-driven Filter Unit, Millex GV 0.22 μm (Millipore, cat. No. SLGV004SL) for RNase-free injection buffer filteration.

RNase-free injection buffer (0.25 mM EDTA, 10 mM Tris, pH 7.4).

A Cas9 expression construct, p3s-Cas9HC with an upstream T7 promoter (Addgene plasmid 43945).

pCAG-RGEN vector (Fig. 2), a modified pSingle-tTS-shRNA vector for RGEN constructs (Clontech, cat. No. 630933).

Pre-annealed oligomers for sgRNA cloning of pCAG-RGEN (target sequence for 5'-tcgaggg + 20-bp target DNA sequence-3' and 5'-aaaac + 20-bp complementary sequence + ccc-3'; guide RNA scaffold for 5'-gttttagagctagaaatagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgctttttttgtgc-3' and 5'-aaaaaaagcaccgactcggtgccactttttcaagttgataacggactagccttatttaacttgctatttctagctct-3').

PCR primers for generating the template of in vitro transcription (the forward primer for 5'-gaaattaatacgactcactatagg + 20-bp target DNA sequence-3' and the reverse primer for 5'-agcaccgactcggtgccact-3').

Recombinant Cas9 protein (Toolgen, Inc.).

2.1.2. Generation of TALEN constructs

In general, to construct a functional TALEN, a 10–20 tandem repeat array of 4 different TALE modules, each recognizing A/T/G/C nucleotide in a pre-determined order for a given target sequence is assembled. While a construction scheme with multiple steps based on conventional molecular cloning techinues were initially utilized, more sophisticated approaches such as Golden Gate cloning [30], solid surface-based assembly, and ligation-independent cloning allowing the assembly of long TAL effector module array in 1–2 steps were developed to facilitate a straightforward and high-throughput production of TALENs [30–33]. Many of these assembly systems or TALENs produced using these systems are available via multiple academic and commercial sources.

Recently, based on a Golden Gate assembly approach, we established a robust TALEN construction system for ready-for-use TALEN expression plasmids in a single step. The TALEN assembly system features a comprehensive 424 mono-/di-/tri-partite TAL effector arrays for 6 different Golden Gate assembly positions allowing the production of TALENs with up to 18 TAL effector modules in one reaction. Additionally, the diversified sequences for each module limits unwarranted recombination from repetitive nature of TAL effector module arrays and allows efficient analysis by Sanger sequencing, which is not as efficient for TALENs with largely uniform coding sequences for all modules. Using this assembly system, we have previously constructed a comprehensive library of TALENs (designed for optimal activity and specificity) that target all human protein-coding genes and many human miRNAs. Additionally, TALENs prepared using this system have been successfully used to produce gene knockout mouse with very high efficiency. TALENs synthesized by this system are available via the Seoul National University (SNU) TALEN library resources (talenlibrary.net).

2.1.3. Construction of RGEN expression vectors and synthesis of sgRNAs

We favor engineered nucleases that destroy the start condon in the target gene and/or inducing frameshifts just downstream of the start codon, and also check for the existence of alternative start codons in target genes to completely eliminate the function of the target gene. Before applying to mouse embryos, RGENs for selected target sequences are tested *in vitro* [34] or using cell lines. For the expression of RGENs in cell lines, we generate a bi-cistronic vector system using pSingle-tTS-shRNA vector (Clontech) as a backbone vector. This construct expresses both Cas9 protein driven by the