High-efficiency and heritable gene targeting in mouse by transcription activator-like effector nucleases

Zhongwei Qiu1, Meizhen Liu1, Zhaohua Chen1, Yanjiao Shao1, Hongjie Pan1, Gaigai Wei1, Chao Yu2, Long Zhang1, Xia Li1, Ping Wang1, Heng-Yu Fan2, Bing Du1, Bin Liu3, Mingyao Liu1,4,* and Dali Li1,*

1Shanghai Key Laboratory of Regulatory Biology, Institute of Biomedical Sciences and School of Life Sciences, East China Normal University, Shanghai 200241, China, 2Life Science Institute, Zhejiang University, Hangzhou 310058, China, 3Shanghai Institute of Materia Medica, Chinese Academy of Sciences, 555 Zu Chong Zhi Road, Shanghai 201203, China and 4The Institute of Biosciences and Technology, Texas A&M University Health Science Center, Houston, TX 77030, USA

Received January 7, 2013; Accepted March 20, 2013

ABSTRACT

Transcription activator-like effector nucleases (TALENs) are a powerful new approach for targeted gene disruption in various animal models, but little is known about their activities in Mus musculus, the widely used mammalian model organism. Here, we report that direct injection of in vitro transcribed messenger RNA of TALEN pairs into mouse zygotes induced somatic mutations, which were stably passed to the next generation through germ-line transmission. With one TALEN pair constructed for each of 10 target genes, mutant F0 mice for each gene were obtained with the mutation rate ranged from 13 to 67% and an average of ~40% of total healthy newborns with no significant differences between C57BL/6 and FVB/N genetic background. One TALEN pair with single mismatch to their intended target sequence in each side failed to yield any mutation. Furthermore, highly efficient germ-line transmission was obtained, as all the F0 founders tested transmitted the mutations to F1 mice. In addition, we also observed that one bi-allele mutant founder of Lepr gene, encoding Leptin receptor, had similar diabetic phenotype as db/db mouse. Together, our results suggest that TALENs are an effective genetic tool for rapid gene disruption with high efficiency and heritability in mouse with distinct genetic background.

INTRODUCTION

Gene targeting, based on the technique of site-directed mutagenesis in mouse embryonic stem (ES) cells, facilitates researchers to generate virtually any desired modification of the mouse genome (1,2). Conventional gene targeting, by use of homologous recombination (HR) in mouse ES cells to modify specific gene, has been proven to be a valuable tool in mouse genetic research (1,3–5). Although thousands of genes have been disrupted in mice by conventional gene targeting (6), investigators’ growing experimental requirements are still to be addressed. There are several limiting factors of conventional gene targeting in mice, such as the time-consuming process of targeting vector construction, low spontaneous HR rate in mouse ES cell and usually low germ line transmission efficiency after injection of targeted ES cells into wild-type blastocysts (7). In addition, when ES cells are not available from the desired strain, it is necessary to take a backcrossing strategy, which is a really time-consuming process.

Genome editing with engineered nucleases, an emerging technique, was designated the method of the year 2011 by the Nature Methods (8). Zinc-finger nucleases (ZFNs) have overcome the limitations of ES cell-based conventional gene targeting to a certain degree. ZFNs are generated by fusing a zinc finger DNA-binding domain to a DNA-cleavage domain, such as FokI (9,10). After binding of Zinc-finger proteins to their respective target sites, the corresponding FokI domains form heterodimers and possess DNA cleavage activity (11). Once double-strand breaks (DSBs) are generated, eukaryotic cells...
mainly responded by one of two pathways, either error-prone non-homologous end joining (NHEJ) or high-fidelity HR, depending on the presence of a donor template, to repair the breaks (12,13). Once DSBs are repaired by the error-prone NHEJ, deletions, substitutions or insertions (indels) at cleavage sites are frequently introduced into the genome. ZFNs have been used to modify endogenous genes in multiple cell types (14). Moreover, by microinjection of DNA or messenger RNA (mRNA) of customized, sequence-specific ZFNs into fertilized eggs, germline transmittable founders have been successfully obtained in >10 different species, including not only the widely used model organism, such as mouse, rat, zebrafish, frog, fruit fly, but also livestock (14–20). However, owing to the nucleotide triplet-recognition nature of the zinc finger motif, ZFNs have less flexibility in target sequence selection. In addition, because of the so-called neighbor effect between motifs, ZFNs require considerable optimization to achieve specific binding to new target sequences (14). In recent years, several new methods have been available to facilitate motif assembly and screening, but the nuances of zinc finger protein engineering made the method not easy to be adopted in new laboratories (20). Although engineered zinc finger proteins for custom targets are commercially available, it is still expensive to impede ZFN as routine strategies for majority of laboratories worldwide (14,20).

Another method in genome editing with engineered nucleases is the transcription activator-like effector nucleases (TALENs), which has distinct DNA recognition and binding domain compared with the artificial zinc finger-mediated DNA binding. The DNA-binding domain of TALENs is derived from transcription activator-like (TAL) effector proteins, which are originally discovered from the plant pathogenic bacteria Xanthomonas (21–23). The DNA-binding domains of TAL effectors contain different numbers and ordering of repeats, with each repeat ¬34–35 amino acids long (21). Owing to the one-to-one correspondence between the amino acids 12 and 13 (known as repeat variable di-residues (RVDs)) of the TAL effector module and single target nucleotides, TAL effectors have more predictable DNA-binding rules, which allow researchers to access and engineer virtually any genomic DNA sequence with no need for optimization (21,22,24,25). Recently, TALENs have been successfully applied to introduce precise genomic modifications (21,22,24,25). The target site was divided into several parts constructed simultaneously to accelerate the process.

**Materials and Methods**

**Construction of TALEN expression vectors**

Target sites were identified using the online software published by Cermak et al. (28) (http://boglabx.plp.iastate.edu/TALENT/help.php). Nucleotide-recognizing TALE single unit vectors and TALEN expression vectors are kindly gifted from Bo Zhang (Peking University). TALE repeat arrays and TALE nuclease expression vectors were constructed as described by Huang et al. (29). Briefly, two adjacent single unit vectors were double digested with Nhel+HindIII (for the upstream module) or SpeI+HindIII (for downstream module). The two DNA fragments were ligated to generate a double unit vector, which then served as an intermediate vector for the next round of the digestion-ligation cycle. The target site was divided into several parts constructed simultaneously to accelerate the process.

**In vitro transcription and preparation of TALEN-coding mRNA**

To prepare mRNA for microinjection, TALEN expression vectors were linearized by NotI, followed by phenol/chloroform extraction and ethanol precipitation. Then the linearized DNA was transcribed using Sp6 mMESSAGE mMACHINE Kit (Ambion) following the manufacturer’s instructions. mRNA was recovered by lithium chloride precipitation and resuspend in nuclease-free water.

**Microinjection of zygotes**

Mouse zygotes were obtained by superovulation of FVB/N or C57BL6N females (SLAC Shanghai) mating with the males of the same strain. The zygotes were cultured in KSOM embryo culture medium (Millipore) before injection. In all, 40 ng/µl mRNA of each TALEN pair was injected into the cytoplasm of one-cell stage embryos through the injection needle. Injections were performed using an Eppendorf transferMan NK2 micromanipulator. Injected zygotes were transferred into pseudopregnant ICR female mice immediately after injection or after overnight culture in KSOM medium. Mice were housed in standard cages in a specific pathogen-free facility on a 12h light/dark cycle with ad libitum access to food and water. All animal experiments conformed to the regulations drafted by the Association for Assessment and Accreditation of Laboratory Animal Care in Shanghai and were approved by the East China Normal University Center for Animal Research.

**Founder identification by T7E1 mismatch sensitive assays**

Toe or tail clips were subjected to standard DNA extraction procedure. The extracted DNA was amplified using the following primers around the target sites to produce amplifications at sizes of 400–800 bp: *Lepr*: F-tgaagagttccgacactca, R-gaatgctggagcaacaccacaac; *Paklpl*: F-ttgcccttggtcttcctc, R-tccacagctgtgtgcttc; *Gpr55*: F-tatgcatgtgctgtgtgcttc, R-gttaagacacccatcacatgc; *Rprm*: F-gcctgtaagagctgtgcttc, R-gaagtgtgtagagactggtgct; *Wdr20a*: F-gacttagggcaggatgaa,
Genetic code: wild-type DNA: 98 generate mismatches into heteroduplexed DNA with poly-R-ctgtggtaagggccaaatcatg. After amplification was finished, the following program was carried out to couple receptor 55 (Gpr55), Reprimo (Rprm), SMAD specific E3 ubiquitin protein ligase 1 (Smurf1), WD repeat domain 20A (Wdr20a), Transmembrane protein 74 (Tmem74), F-box only protein 6 (Fbxo6), Dcaf13 DDBI and CUL4-associated factor 13 (Dcaf13) and Family with sequence similarity 73, member A (Fam73a) range from 22 to 27 bp. In longer spacer group, a perfect match with the target DNA. TALE nuclease experiments would affect TALEN activity, the last nucleotide mismatch between TALENs and their target DNA sequences was not ended with a cytosine. To identify the modifications in founders, PCR products from each founder using the primers aforementioned were cloned using TA cloning kit (Takara) following the manufacturer’s instructions. At least six colonies were picked from each transformation and then sequenced with RV-M primer.

Glucose tolerance tests
After overnight (14–16 h) fasting, the mice were weighed and injected intraperitoneal with glucose (2 mg/g body weight). Blood samples were collected from the tail immediately before and 15, 30, 60, 120 and 180 min after injection. Blood glucose levels were measured at each time point using an ACCU-CHEK® Performa Nano (Roche).

Insulin tolerance tests
Mice were fasted for up to 4 h, after which they were weighed and injected intraperitoneal with insulin (1.5 U/kg). Blood samples were collected from the tail immediately before and 15, 30, 60 min after injection. Blood glucose levels were measured at each time point using an ACCU-CHEK® Performa Nano.

Oil Red O staining
In all, 5 μm frozen sections from liver were fixed in formalin and rinsed in three changes of distilled water. For staining, sections were placed in Oil Red O solution (0.5% Oil Red O in propylene glycol) for 10 min. Then sections were rinsed in 85% propylene glycol for 5 min, rinsed in distilled water and counterstained in hematoxylin for 3 min. Finally, sections were rinsed under running water for 15 min and mounted with aqueous solution.

RESULTS
Design and assembly of TALENs
To investigate the possibility and the efficiency of TALENs to generate gene mutant mice through direct TALEN microinjection in zygotes, we first selected 10 target mouse genes located in eight different chromosomes. The protein products of the genes are varying from transmembrane, cytoplasmic to nuclear proteins. Specific recognition sites of the target genes, including Leptin receptor (Lepr), p21-activated protein kinase-interacting protein 1 (Pak1ip1), G protein-coupled receptor 55 (Gpr55), Reprimo (Rprm), SMAD specific E3 ubiquitin protein ligase 1 (Smurf1), WD repeat domain 20A (Wdr20a), Transmembrane protein 74 (Tmem74), F-box only protein 6 (Fbxo6), Dcaf13 DBB1 and CUL4-associated factor 13 (Dcaf13) and Family with sequence similarity 73, member A (Fam73a) were identified by using the online software published by Cermak et al. (28), and one pair of TALENs for each gene was designed. As shown in Figure 1a, all pairs of TALENs were generated using the +63 TALEN scaffold with the heterodimeric FokI domains (29).

To test whether the activity of TALENs is significantly affected by the length of the spacer as reported previously (30), we designed the TALENs and grouped them into two sets depending on the length of the spacer. For the Lepr, Pak1ip1, Gpr55, Rprm, Smurf1 and Fbxo6 gene, the length of the spacer sequence is <20 bp (from 16–19 bp). In the other group, the spacers for Wdr20, Tmem74, Dcaf13 and Fam73a range from 22 to 27 bp. In longer spacer group, a SacI restriction endonuclease site is included in spacers for Wdr20 and Tmem74 (Figure 1b). To test whether the mismatch between TALENs and their target DNA sequences would affect TALEN activity, the last nucleotide of both sides of the recognition site was not ended with a thymidine (T) for the Smurf1 gene, labeling as Smurf1(m) (Figure 1b, red character). In the TALEN expression vector we used, the last 0.5 unit encoded the RVD binding T (29); thus, the TALEN for Smurf1(m) is not a perfect match with the target DNA. TALE nuclease expression vectors were constructed with unit assembly method developed by Huang et al. (29).

Successful generation of targeted gene disruption with TALENs in mice
In vitro transcribed mRNA for each TALEN-pair was injected into the cytoplasm of one-cell stage mouse embryos. After injection, the embryos were either transferred into pseudopregnant female mice immediately or following overnight culture. The genomic DNA of the pups from each TALEN injected embryos was extracted from tail or toe clips, and the sequence span target site of each gene locus was amplified by PCR. For the Lepr, Pak1ip1, Gpr55, Rprm, Fbxo6, Smurf1, Dcaf13 and Fam73a genes, PCR products were melted and annealed with a similar amount of PCR products amplified from wild-type mouse genomic DNA to form heteroduplex DNA. The DNA heteroduplex was then subjected to a mismatch sensitive endonuclease (T7E1) assay to detect genomic modifications (Figure 2a–c). The ratio of T7E1 digestion assay positive offspring ranged from 24 to 67% in the C57BL/6 background, and similar (41–62%) efficiency was obtained in the FVB/N background (Table 1). For the Wdr20a and Tmem74 genes, purified PCR products were subjected to SacI digestion, with the expectation that some of the potential mutations would destroy...
the SacI site and give rise to SacI-resistant DNA fragments during DNA electrophoresis. For Wdr20a gene, 3 of 23 (13%) newborns revealed SacI-resistant DNA fragments and subsequent sequencing confirmed the SacI-disrupting mutations (Figure 2d). With T7E1 digestion assays (data not shown), no additional mutation was identified. For the Tmem74 gene, none of 16 newborns showed SacI-resistant DNA fragments (data not shown), possibly owing to the SacI site not being at the center of spacer region where TALEN-mediated cleavage is likely to happen. However, when PCR products of 16 Tmem74 newborns were subjected to mismatch sensitive endonuclease (T7E1) assays, two of them (13%) showed a mismatch to wild-type genomic DNA, and the sequences of the mutations were confirmed by subsequent sequencing. These results suggest that TALENs are effective in mouse embryos in distinct genetic background with high efficiency.

### Analysis of gene mutations generated by TALENs

To determine the mutations generated in individual founders, PCR products were TA cloned, and at least six colonies for each founder mouse were sequenced. All the mutations identified by sequencing are listed in Table 1. Similar to other species, some of the founders were biallelically modified. For instance, founders 1, 4, 6, 10, 12, 13 of the Rprm gene had the same biallelic non-frame-shifting 3-bp deletions. Founders 2, 3, 7, 12 of the Gpr55 gene had a combination of frame-shifting and non-frame-shifting deletions (Figure 3a). Also, there are mosaic founders where more than two allele types have been identified from a single mouse, such as founder 8 of the Lepr gene in the FVB/N background, founders 3 and 7 of the Rprm gene and founder 19 of the Fbxo6 gene (Figure 3a). However, biallelic mutations were not observed in the founders of other genes tested in this study. In line with the indels induced by TALEN in zebrafish (31), our TALENs generated indels with a broad range of lengths in mice, from 1 to 173 bp, with an average length of 15.4 bp (n = 61 independent indels in total) (Figure 3b). In accord with the deletions induced by ZFNs in mice (18), the deletions generated by TALENs contain a microhomology of 1–4 bp at the deletion boundary (Figure 3a). The statistic of the injections and mutation rates is shown in Table 1.

To test whether the specificity of TALENs with the intended binding target will greatly affect the enzyme activity, we generated TALEN pairs against Smurf1,
mutations. M, DNA marker.

Figure 2. Identification of founder mice using T7 endonuclease assay or SacI digestion. (a–c) PCR primers were designed flanking the TALENs binding sites of each gene. PCR products were melted and annealed with a similar amount of PCR products amplified from wild-type mouse genomic DNA to form heteroduplex DNA. The DNA heteroduplex was then subjected to mismatch sensitive endonuclease T7E1 assays, and the products were resolved in agarose gels. SacI uncut mutations were indicated. Underlined data indicated mice bearing TALENs against the C57BL/6 strain (a) or FVB/N strain (b) mice injected with TALENs (T7E1) assays, and the products were resolved in agarose gels. A heteroduplex was then subjected to mismatch sensitive endonuclease digestion. (d) The PCR products from newborn mice injected with TALENs against Fam73a gene were subjected to SacI digestion, and the products were resolved in agarose gels. SacI uncut mutations were indicated. Underlined data indicated mice bearing mutations. M, DNA marker.

Table 1. Injection statistics

<table>
<thead>
<tr>
<th>Target</th>
<th>Strain</th>
<th>Dose (ng/µl)</th>
<th>Injected/total (%)</th>
<th>Total newborns</th>
<th>Mutants number</th>
<th>Mutants rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lepr</td>
<td>C57BL/6</td>
<td>40</td>
<td>152/93 (61)</td>
<td>14</td>
<td>8</td>
<td>57%</td>
</tr>
<tr>
<td>Lepr</td>
<td>FVB/N</td>
<td>40</td>
<td>84/57 (67)</td>
<td>13</td>
<td>8</td>
<td>62%</td>
</tr>
<tr>
<td>Pak1ip1</td>
<td>C57BL/6</td>
<td>40</td>
<td>88/45 (51)</td>
<td>6</td>
<td>3</td>
<td>50%</td>
</tr>
<tr>
<td>Gpr5</td>
<td>FVB/N</td>
<td>40</td>
<td>135/84 (62)</td>
<td>22</td>
<td>9</td>
<td>41%</td>
</tr>
<tr>
<td>Rphpm</td>
<td>C57BL/6</td>
<td>40</td>
<td>145/88 (68)</td>
<td>15</td>
<td>10</td>
<td>67%</td>
</tr>
<tr>
<td>Fbxo6</td>
<td>C57BL/6</td>
<td>40</td>
<td>200/130 (75)</td>
<td>47</td>
<td>24</td>
<td>51%</td>
</tr>
<tr>
<td>Smurf1</td>
<td>C57BL/6</td>
<td>40</td>
<td>185/168 (91)</td>
<td>33</td>
<td>15</td>
<td>45%</td>
</tr>
<tr>
<td>Tmem74</td>
<td>C57BL/6</td>
<td>40</td>
<td>97/70 (72)</td>
<td>16</td>
<td>2</td>
<td>13%</td>
</tr>
<tr>
<td>Wdr20a</td>
<td>C57BL/6</td>
<td>40</td>
<td>123/92 (81)</td>
<td>23</td>
<td>3</td>
<td>13%</td>
</tr>
<tr>
<td>Dcaf13</td>
<td>C57BL/6</td>
<td>40</td>
<td>182/123 (68)</td>
<td>33</td>
<td>11</td>
<td>33%</td>
</tr>
<tr>
<td>Fum75a</td>
<td>C57BL/6</td>
<td>40</td>
<td>118/75 (64)</td>
<td>13</td>
<td>3</td>
<td>24%</td>
</tr>
</tbody>
</table>

Statistics for TALEN injection. TALENs mRNA was injected into the cytoplasm of mouse zygotes, which were transferred into pseudopregnant ICR female mice. The third column indicates the number of injected zygotes/the number of zygotes transferred to foster mothers. Mutations of the newborns were confirmed by sequencing after weaning.

which containing mismatches as aforementioned. The last 0.5 TALE repeat recognizing nucleotide T at the 3' end of the TALE binding site was pre-fixed into the backbone of the pCS2-FokI vector, which restricted the selection of TALEN targets. For the mismatched TALENs against Smurf1 gene, both TALEN binding sites did not end with nucleotide T (Figure 1b). All of 18 newborns injected with the mismatched TALEN pair against Smurf1 turned out to be wild-type animals in the T7E1 assay (data not shown), and no mutation was identified by DNA sequencing (data not shown). However, we also used a TALEN pair whose target is adjacent to that of the mismatched pair and generated 45% mutant pups. These data indicated that the mismatch of the last 0.5 TALE repeat to the target sequence has the potential to impair TALEN activity.

Phenotypes of Lepr somatic mutations induced by TALEN

One of our targeted genes, Lepr, encodes a receptor for Leptin, which has been shown to regulate food intake, body weight and glucose homeostasis and also has the potential for treating diabetes (32). The Lepr mutant db/db mice demonstrate obesity by the time they are 3–4 weeks old and have been widely used as the mouse model for type 2 diabetes (33–35). Biallelic mutation of the Lepr gene in mice will not cause embryonic lethality, and the obesity phenotype makes them easy to be distinguished, which is why we chose the Lepr gene as the target to test whether TALENs can generate somatic mutations with a similar phenotype. Our biallelically modified Lepr founder 8 (Lepr8) in the FVB/N strain, with different frame-shifting deletions, exhibited an obese phenotype by 4 weeks of age (Figure 4a). The body weight was approximately double the weight (44 versus 23 g) of wild-type littermates (Figure 4b). The glucose homeostasis of Lepr8 was severely impaired. In glucose tolerance tests, Lepr8 showed a marked higher fasting glucose level and glucose intolerance compared with its wild-type littermate (Figure 4c). During insulin tolerance testing, Lepr8 was unresponsive to insulin administration (Figure 4d), indicating resistance to insulin treatment. Moreover, examination of white adipose tissue (WAT) from Lepr8 revealed a considerable cellular hypertrophy compared with the wild-type littermate (Figure 4e). In addition, Oil-Red-O staining of frozen sections from Lepr8 liver showed strong signals indicating lipids, whereas the counterpart section from the littermate control was barely stained, indicating a marked increase of lipid accumulation in the liver of Lepr8 (Figure 4f). These results suggested that the somatic mutations induced by TALEN effectively impaired target gene function.

Highly efficient germline transmission for TALEN-induced mutations

Establishment of knockout mouse lines is dependent on successful germ-line transmission. To determine whether TALEN-induced mutation can be effectively passed through the germline, one or more founders of each gene were backcrossed to wild-type mice. All founders
Figure 3. Somatic mutations induced by TALENs in mice. (a) Frequencies and sequences of TALEN-induced mutations. The binding sites of the TALEN half-site are shown in bold. A dash represents a single base deletion. Insertions or substitutions are indicated in red letters. Microhomology that was likely used by NHEJ is underlined. The length and frequency of indels are labeled to the right of each sequence. (b) TALENs induce a wide range of indel length. The percentage is from the number of clones of different indel lengths divided by the number of total clones sequenced.
that give birth to next generation successfully transmitted at least one mutant allele to their offspring, with the germline transmission rate ranging from 18 to 89% (Figure 5a). As observed during ZFN-mediated knockout in mice (18), mutant alleles that were not initially identified were also transmitted and identified in the F1 generation in founder 2 of Pak1ip1, founders 3, 7, 8, 11 of Gpr55 and founder 16 of Wdr20a, suggesting incomplete sequencing of the TA clones or an uneven distribution of certain mutation-bearing cells between the toes and germ cells (Figure 5b). Successful germ-line transmission was observed in all founders by test breeding. In addition, we observed that the chimerism in the tail of the founders is not directly indicating the germline transmission rate (Supplementary Table S1). These results indicate that TALEN-induced gene mutation is heritable in mice of different genetic backgrounds.

**DISCUSSION**

TALEN has been proven to be a robust and valuable methodology for species in which it was difficult to generate targeted mutations for desired genes (29,36–38). In this study, we successfully generated mutant mouse strains for 10 genes and demonstrated that the mutations were heritable.

A number of studies demonstrated that TALENs exhibited strong specificity for the intended binding sequence (39,40). TALEN activity could be greatly reduced by as few as a two-nucleotide difference...
Figure 5. Efficient germline transmission of TALEN-induced mutations in mice. (a) Effective transmission of TALEN-induced mutation through germline. Alleles that appeared in F1 but were not originally identified in founders are highlighted in yellow. (b) Alignment of TALEN-induced mutations in the F0 and the F1 generation of selected genes. The binding sites of the TALEN monomers are shown in bold. A dash represents a single base deletion. Insertions or substitutions are indicated in red letters.
between the alternative target and the intended binding site (40,41). When the 2-bp mismatch was located within the target half-site of the TALENs, the mutation frequency was dropped from >50% (the same site with no mismatch) to ~2% in zebrafish embryos (41). In this study, to test what extent the mismatch (especially the mismatch of the 3' most RVD binding) of TALENs to the target DNA sequence would influence enzyme activity, we designed one pair of TALEN for Smurf1 (Figure 1b). No mutation was detected from 18 pups injected with the TALEN pair mismatched to the target sequence. Although there was no perfect control of the TALEN pair, we used a TALEN pair whose target is adjacent to the mismatched pair and generated 45% mutant pups. Our data illustrated that a match at the 3'-end of the target site is important for TALEN activity.

Besides the specificity of TALENs to their intended binding site, the spacer length between TALEN targets is also critical for the activity. For the +63 TALEN scaffold architecture, when the optimal length of spacer between two half-sites is 12–21 bp, it yielded at least 5% modification frequency, whereas no activity was observed once the spacer length was outside this range in human cell lines (30). We also confirmed that the spacer length was important for TALEN activity in mouse embryos. All 6 nuclease pairs separated by 15–19 bp resulted in at least a 41% mutation rate in newborns with average mutation rate of 53% (62 mutants among 117 newborns). In the other four pairs, the spacer length was longer than 22 bp, and these TALENs yielded a 13–33% modification frequency (22% in average), which was below the average (54%) of the above TALENs with shorter spacers. It is not to say the optimal spacer length is from 15 to 19 bp, as the number of TALENs used in this study is too limited, and the nucleotide composition of the spacer could also affect the activity. Nevertheless, we suggest that it is better to limit the spacer length to <19 bp when designing the target sites to generate knockout mouse strains.

As observed by others in generating knockout animals through artificial nucleases in mice and rats, some deletions were unexpectedly commonly identified in multiple founders (18,42). As the deletions were shared in founders 1–4 bp at the deletion boundary, which was consistent with the observation of Carbery et al. (18) and again supported the notion that microhomology at the ends of a DSB promotes NHEJ (43).

It is interesting to investigate when the TALEN functions in the embryos after mRNA injection. Recently, Sung and colleagues reported that TALEN created mutations of mouse genome in 12 h after mRNA injection (27). They found that mosaicism was rarely detected in the mutants generated by two TALEN pairs and claimed that TALEN activity is not likely to be maintained after the first cleavage of one-cell embryos (27).

However, in this study, we found that mosaicism was detected in majority of the founders. And in founder number 7 of Gpr55 knockout mouse strain at least five genotypes were observed in its offspring containing new genotype, which was not identified in the founder tail DNA. In total, we found six genotypes from this founder or its offspring (Figure 5a). It suggests that TALEN activity maintained at least to two-cell stage embryos, and the genotype of all F1 mouse need to be identified by sequencing.

Although gene targeting through HR in ES cells to generate knockout mouse strains is a well-established technology in developmental and genetic studies, it is still worthwhile to introduce TALENs into mouse genetics owing to its multiple advantages. In our studies, it usually takes 4 months to get heterozygous mice for desired genes through TALENs, which will save at least 5–6 months compared with using a conventional strategy. It is worth noting that in this study, none of the 10 TALEN pairs underwent any pre-selection and optimization in cells, and multiple founders of each gene were obtained following only one round of microinjection. Moreover, owing to the early onset of TALEN activity after injection of TALEN mRNA, mutation-bearing cells are likely to contribute to a considerable percentage of germ cells, thus overcoming the low transmission efficiency issue that usually occurred with injection of targeted ES cells into wild-type blastocysts. We also demonstrated comparable TALEN activities in mice from different genetic backgrounds with high germline transmission rates. TALENs will be a powerful tool for gene targeting in mouse strains whose ES cell line is not easily available. It is convenient to introduce exogenous DNA sequences and to make conditional knockout mice through HR in ES cells. However, to our knowledge, no long DNA fragments knockin mammals were generated through TALENs, although knockin of both short or long DNA sequences have been reported in zebrafish (44,45). As knockin reporter and conditional knockout mouse strains are important tools for research, how to generate large fragment knockin and conditional knockout strains through TALENs is the next important bottleneck to be broken through.

SUPPLEMENTARY DATA
Supplementary Data are available at NAR Online: Supplementary Table 1 and Supplementary Figure 1.

ACKNOWLEDGEMENTS
The authors thank Dr. Bo Zhang in Peking University for kindly providing nucleotides-recognizing TALE single unit vectors and TALEN expression vectors. They also thank Stefan Siwko for comments and advice.

FUNDING
State Key Development Programs of China (in part) [2012CB910400 to M.L., 2010CB945403 to D.L.].
National Natural Science Foundation of China [no. 31171318 to D.L. and 30930055 to M.L.]; Science and Technology Commission of Shanghai Municipality [11DZ2260300]. Funding for open access charge: National Natural Science Foundation of China [no. 31171318].

Conflict of interest statement. None declared.

REFERENCES


