*Article*

# Efficient Single Base Editing in Mouse Using Cytosine Base Editor 4

**Salah Adlat[1], Ping Yang[1], Yang Chen[1], Rajiv Kumar Sah[1], Zin Mar Oo[1], May Zun Zaw Myint[1], Farooq Hayel[1], Noor Bahadar[1], Mahmoud Al-Azab[3], Fatoumata Binta Bah[1], Luqing Zhang[4]*, Yaowu Zheng*[1,2] and Xuechao Feng*[1,2]**

[1] Transgenic Research Center, School of Life Sciences, Northeast Normal University, Changchun, Jilin 130024, China, land788@nenu.edu.cn (S.A.); yangp1993@qq.com (P.Y.); cheny339@nenu.edu.cn (Y.C.); rajiv26500@yahoo.com (R.K.S.); zinmaroo23@yahoo.com (Z.M.O.); mayzunzawmyint@yahoo.com (M.Z.Z.M.); farooqhail1@gmail.com (F.H.); noorizever@yahoo.com (N.B.); alazab82@yahoo.com (M.A.); cissehbah@yahoo.fr (F.B.B.)

[2] Key Laboratory of Molecular Epigenetics of Ministry of Education, School of Life Sciences, Northeast Normal University, Changchun, Jilin 130024, China.

[3] Immunology Laboratory, Guangzhou Institute of Pediatrics, Guangzhou Women and Children Medical Centre, Guangzhou, China, alazab82@yahoo.com (M.A.)

[4] Cardiovascular Research Institute, University of California, San Francisco, CA; Zhangluqing@gmail.com (L.Z.)

*Corresponding author: zhengyw442@nenu.edu.cn (Y.W.Z.); fengxc997@nenu.edu.cn (X.C.F.); Tel.:+86-0431-85098583 (Y.W.Z., and X.C.F)

**Abstract:** Most human genetic disease arises from point mutations. These genetic diseases can theoretically be corrected by gene therapy but clinic practice is still far from mature. Nearly half of the pathogenic single-nucleotide polymorphisms (SNPs) are caused by G:C>A:T or T:A>C:G base changes. The best current methods to repair these changes are by base editing without footprint using recently developed CRISPR-Cas9 technology by David Liu's lab. These base editing methods have been confirmed with precision and efficiency in cultured mammalian cells, but it is barely confirmed and the efficiency is still very low. Animal models are important in dissecting pathogenic mechanism for human genetic diseases and efficacy testing of base correction in vivo. Cytidine base editor BE4 is a newly developed version of cytidine base editing system that converts cytidine (C) to uridine (U) in cultured mammalian cells but has not been proven in vivo. In this study, we have tested this system in cells to inactivate GFP gene and in mice by introducing single-base substitution that leads to a stop codon in tyrosinase gene. High percentage albino coat-colored mice were obtained from black coat-colored donor zygotes after pronuclei microinjection. Sequencing results showed that expected base changes were obtained with high precision and efficiency (56.25%). There are no off-targeting events identified in predicted off-target sites. Results confirm BE4 system can work in vivo with high precision and efficacy, and has great potentials in clinic to repair human genetic mutations.

**Keywords:** BE4, CRISPR-Cas9; Tyr; Cytosine base editing; Mouse model

## 1. Introduction

Point mutations are the most common events in human genetic diseases and nearly 50% of disease-associated mutations are C>T and G>A substitutions [1]. Animal modeling of human genetic diseases are valuable in study of pathogenic mechanism, testing of drug efficacy and proof of gene therapy reliability. CRISPR Cas9 system is an adaptive immune system in bacteria that protects its

genome from invading viruses [2,3]. CRISPR Cas9 system has been successfully applied to genetic engineering in variety of cells and organisms. Targeted insertion and mutation usually requires homologous recombination that is accomplished through cultured embryonic stem cells (ES cells), selection for positive clones and ES cell/blastocyst injection. This process is time consuming and costly. CRISPR-Cas9 system has been applied to targeted gene engineering but efficiency in vivo has been very low. Homology-directed repair (HDR) with CRISPR-Cas9 system requires DNA double-strand breaks (DSBs) at the target and a DNA template with homologous arms [4-7]. Cells respond to DSBs often with nonhomologous end joining (NHEJ) that may introduce insertions or deletions (indels) and lead to disruption of corresponding genes and reduce the expected targeting events [8,9]. In addition, HDR often leaves footprint that is not doable for base changes. David Liu's lab has developed many versions of CRISPR-based base editing systems with different accuracy and efficacy that eventually permit precise and efficient base change in cultured cells.

CRISPR/Cas9-based cytidine base editors (CBEs) have recently been developed to generate precise base changes from cytidine to thymidine with high efficiency [1; 10; 11; 12]. CBEs system consists of a CRISPR-Cas9-derived DNA-binding module and a cytidine deaminase, and is able to introduce nucleotide substitutions of C>T [13-15] and G>A [16] without DSBs. It has been demonstrated successful in various cells and some organisms [15].

Base editing systems have gone through various stage of improvement to broaden their applicability and utility in editing of single nucleotide and have been widely applied to cell lines, various animals and plants. The fourth generation of base editor 4 (BE4) is the most advanced system in base editing with high precision and efficiency in cultured mammalian cells. BE4 has a cytidine deaminase (rAPOBEC1) with two copies of uracil glycosylase inhibitor (UGI) that are directly fused to C terminus of Cas9n, a Cas9 mutant with a D10A amino acid substitution, through a 32 amino acid linker (Fig. 1A). BE4 enables direct conversion of cytidine (C) to uridine (U) in chosen bases of DNA sequence [9] . However, feasibility and efficacy of this system has not been assessed in vivo. In the current study, we have tested and confirmed that BE4 system is able to perform a multiplexed base editing with high precision and efficiency in mice. BE4 system shows great potentials in modeling human genetic diseases and validation for gene therapy.

## 2. Results

### *2.1 Screening for efficient sgRNAs in HEK293 cells*

The fourth generation of cytosine base editor (BE4) expresses a Cas9n (D10A) fused to cytidine deaminase (rAPOBEC1) and two copies of uracil glycosylase inhibitor (UGI). To test whether it works in our hand, we first transfected BE4 and pEGFP-sgRNAs in cells that stably express EGFP (Fig. 1A) [9]. HEK293 cells stably expressing EGFP were generated by transfection of plasmid pEGFP-N1. EGFP expression was checked under a fluorescence microscope. Result showed that up to 90% of cells with fluorescence (Fig. 1D). Positive clones (HEK293-EGFP) were selected using G418 (500μg/mL). Two sgRNAs were designed to target EGFP gene (Fig. 1E, Supplementary Table 1). The specificity score of both sgRNA-1 and sgRNA-2 were 75%, and 83% respectively based on software analysis (https://benchling.com/). HEK293-EGFP cells were co-transfected with BE4 and plasmids encoding sgRNAs (pEGFP-sgRNA1 and pEGFP-sgRNA2). Two days after transfection, EGFP fluorescence intensity was analyzed by fluorescence microscopy (Fig. 1D). Majority of cells

transfected with pEGFP-sgRNA1 still express relatively high levels of EGFP (64%), while cells transfected with pEGFP-sgRNA2 exhibited weak signal with an intensity of 36%, indicating EGFP was knocked down more efficiently (Fig. 1D, F). Genomic DNA was extracted for PCR and sequencing to confirm successful base editing.

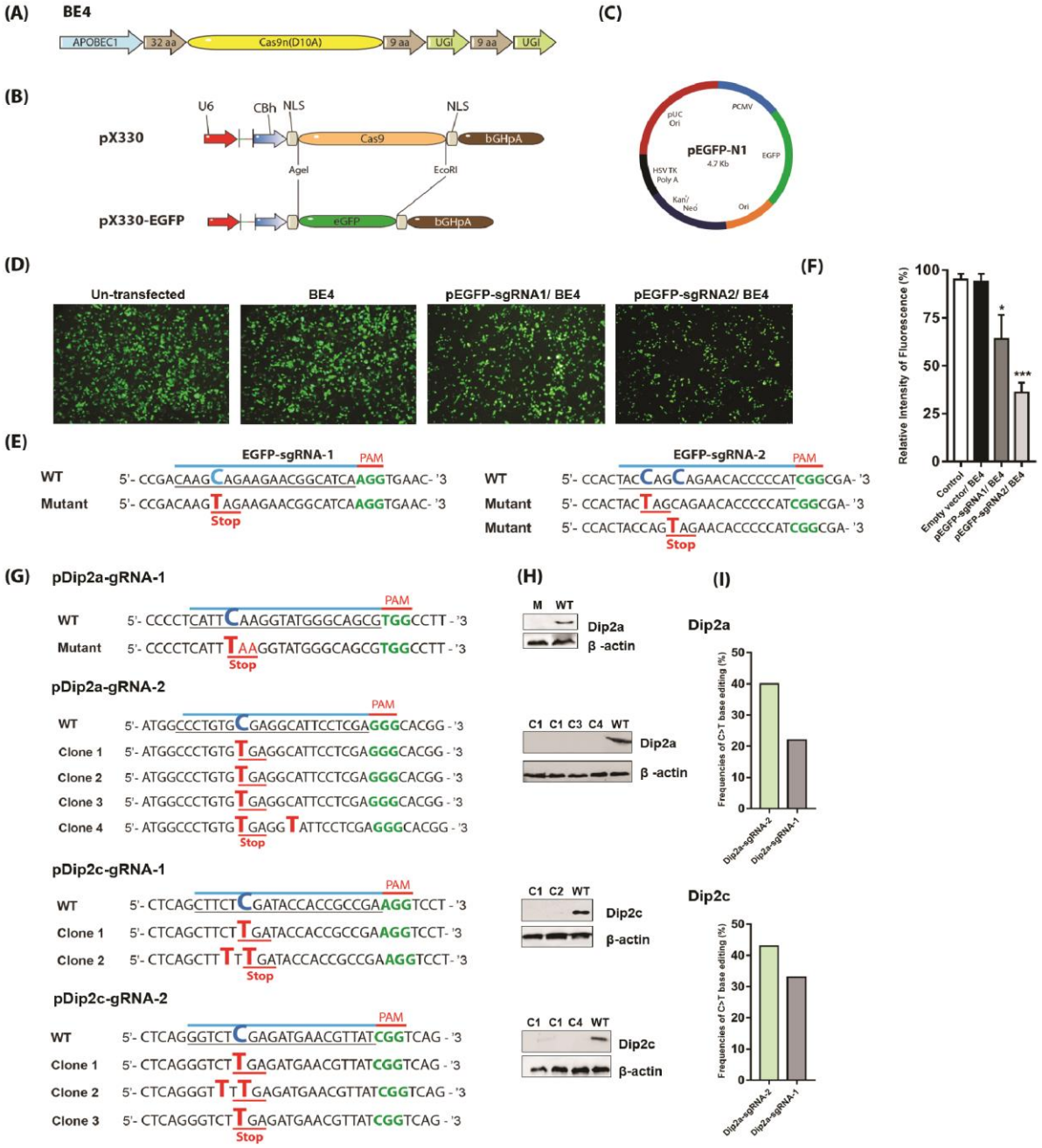


**Figure 1.** Screening of base editing in cells. (A) Architecture of cytosine base editor 4 (BE4). (B) Replacement of Cas9 in pX330 with EGFP. (C) Plasmid source of EGFP. (D) Fluorescent imaging of transfected HEK293 cell stably expressing EGFP. Scale bar, 100 μm. (E) Relative fluorescence intensity of D. (F) Base change in EGFP by sequencing. (G) Premature stop codon targeting of Dip2a and Dip2c genes in B16 cells. PAM sequences labeled in green, wild type base in blue, mutated base in red and stop codon underlined. (H) Western blot analysis. β-actin served as a loading control. (I) Efficiency of C>T base editing. *P*-value was determined by *t*-test. *P<0.05, ***P<0.001.

*2.2 Knockout of Dip2a and Dip2c genes in tumor cells using BE4*

Next, BE4 system was analyzed in murine tumor cell B16-F10. Dip2a and Dip2c genes were each targeted with two sgRNAs (Fig. 1G, Supplementary Table 1). Base substitution was screened by PCR amplification, sequencing and western blotting. pDip2a-sgRNA-1 transfection results showed Q54Z mutation with an efficiency of 22% while pDip2c-sgRNA-1 showed mutations S72F and R73Z with a total efficiency of 33%. Knockout of Dip2a and Dip2c proteins using sgRNAs-1 are shown in Fig. 1G, H, I, Supplementary Fig. 2S, and 3S. Similarly, pDip2a-sgRNA-2 and pDip2c-sgRNA-2 were transfected together with BE4 plasmid. Both pDip2a-sgRNA-2 and pDip2c-sgRNA-2 appeared to work more efficiently and induced 40% and 43% mutations at targeted sites respectively (Fig. 1G, H, I, Supplementary Fig. 4S, 5S). Expression of Dip2a and Dip2c genes from WT and mutated clones were shown in Supplementary Fig. 6S.

*2.3 BE4 can induce C>T substitution in mice*

To explore whether BE4 system can induce site-specific base conversion in mice, sgRNAs targeting exon 1 of *Tyr* locus was designed to inactivate tyrosinase gene (Fig. 2A). Target sequences were synthesized and cloned into pX330-EGFP to express both sgRNA and EGFP. pTyr-sgRNA (2.35ng/μl) and BE4 plasmid (2.65ng/μl) were co-injected into nucleus of B6/D2F1 mouse zygotes and transplanted into surrogate mothers at two-cell stage (Fig. 2B). pTyr-sgRNA/BE4 schematic depiction was shown in Fig. 2C. A total of 16 live pups were obtained (Fig. 2D). Mice were genotyped using following primer (*Tyr*): Forward: 5′-AGAAATTCGAGAACTAACTG-3′, Reverse: 5′-CAGTTAGTTCTCGAATTTCT-3′ (Fig. 2E, Fig. 3A, B). PCR products were purified and sequenced to verify targeted point mutations (Fig. 3C, D, E). A total of 10 mice (62.5%) showed point mutations with C>T and C>A base conversion (Fig. 3E, Table 1). Mutations occurred at high efficiency at 13-15bp in front of PAM (Fig. 3C). The editing frequencies of nonsense mutations in *Tyr* locus with expected amino-acid conversion (C>T) from arginine to a stop codon (R224Z) were 56.25% (9 out of 16) (Fig. 3E, Table 1). These mutations resulted in a mosaic pigmentation phenotype. Some C>T substitution happened at two-cell stage after microinjection. Several founders exhibited obvious chimeric phenomenon with a combination of non-mutant and mutant cells and a combination of homozygous and heterozygous cells. Founders 1#, 3#, and #5 were homozygous for nonsense mutation at targeted site with a conversion rate of 18.75%. Founders 2#, #4, #6, #11, #13, and #15 were showed heterozygous mutation with a frequency of 37.5%. Founders #7, #8, #9, #10, #14, and #16 were mostly wild-type alleles with a frequency of 37.5%. Meanwhile, C>A substitutions in founders #12 without amino acid change was observed with a frequency of 6.25%. No indels were detected at target site (Fig. 3E). No off-target mutations were detected at potential off-target sites (Supplementary Fig. 1S). All results suggest that BE4 system is precise and efficient in introducing single point mutations in vivo.

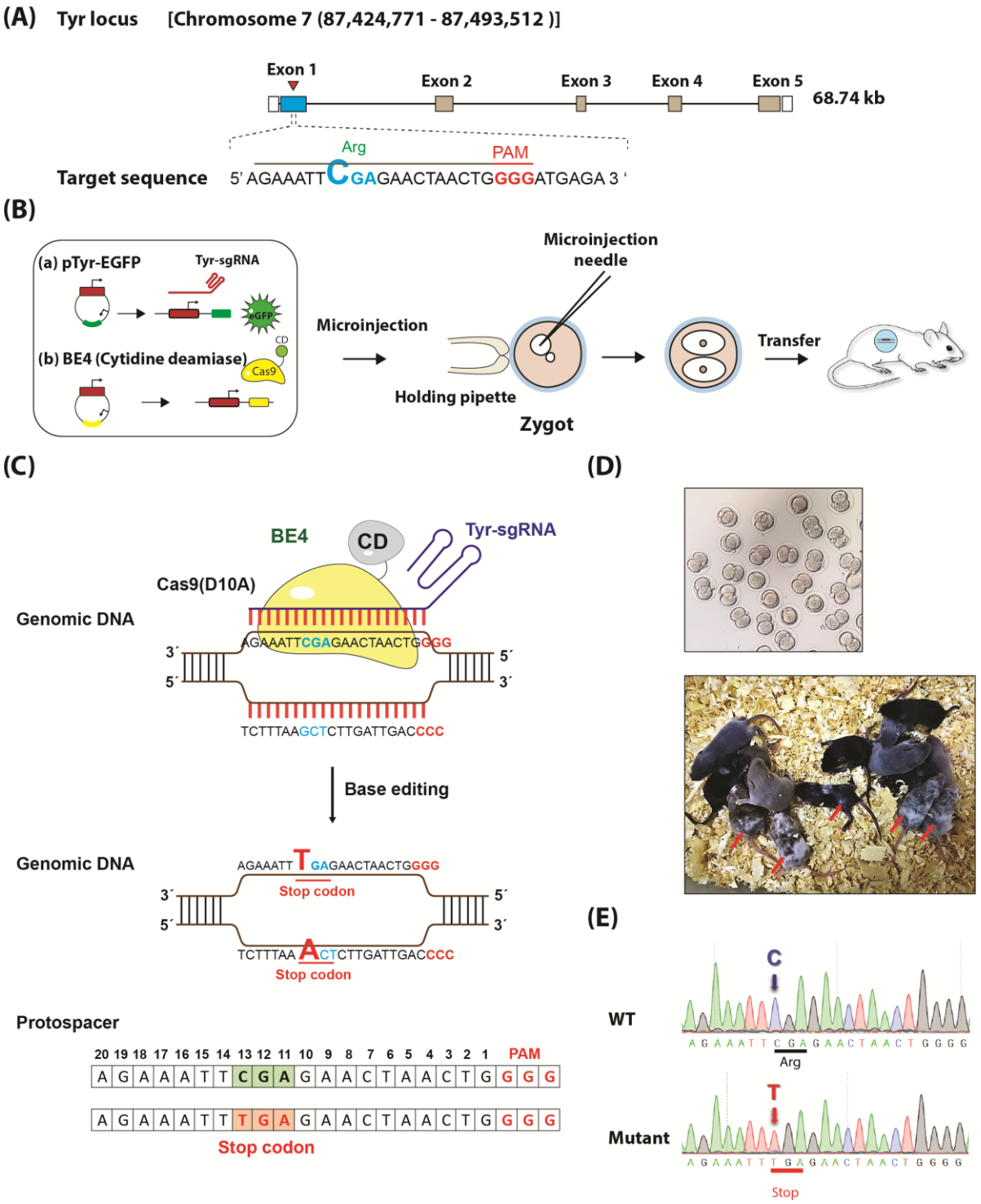


**Figure 2.** BE4 mediated C>T base editing in mice. (A) Schematic of sgRNA design at Tyr locus. (B) Working model of base editing in mice. (C) Schematic depiction of BE4 base editing. (D) Coat color of 8 day old Tyr mutant founders with mosaic pigmentation. (E) Chromatograms of WT and mutant sequences showing C>T substitution.

**Table 1. Summary of the constructs, number of zygotes, and mutant mice obtained after BE4 and Tyr sgRNA microinjection.**

| Construct | Concentration | No. of injected zygotes | No. of transferred zygotes (%)[a] | No. of newborn mice (%)[a] | No. of living mice (%)[a] | Frequency substitution (%)[b] | Frequency of desired substitution (%)[b] | Error rate (%)[b] | Percentage of success (%)[b] |
|---|---|---|---|---|---|---|---|---|---|
| BE4<br>Tyr-sgRNA | 2.64 ng/μl<br>2.35 ng/μl | 70 | 21 (30%) | 16 (76.1%) | 16 (76.1%) | 10 (62.5%) | 9 (56.25%) | 1 (6.25%) | 9 (56.25%) |

a Calculated from the number of zygotes
b Calculated from the number of living mice

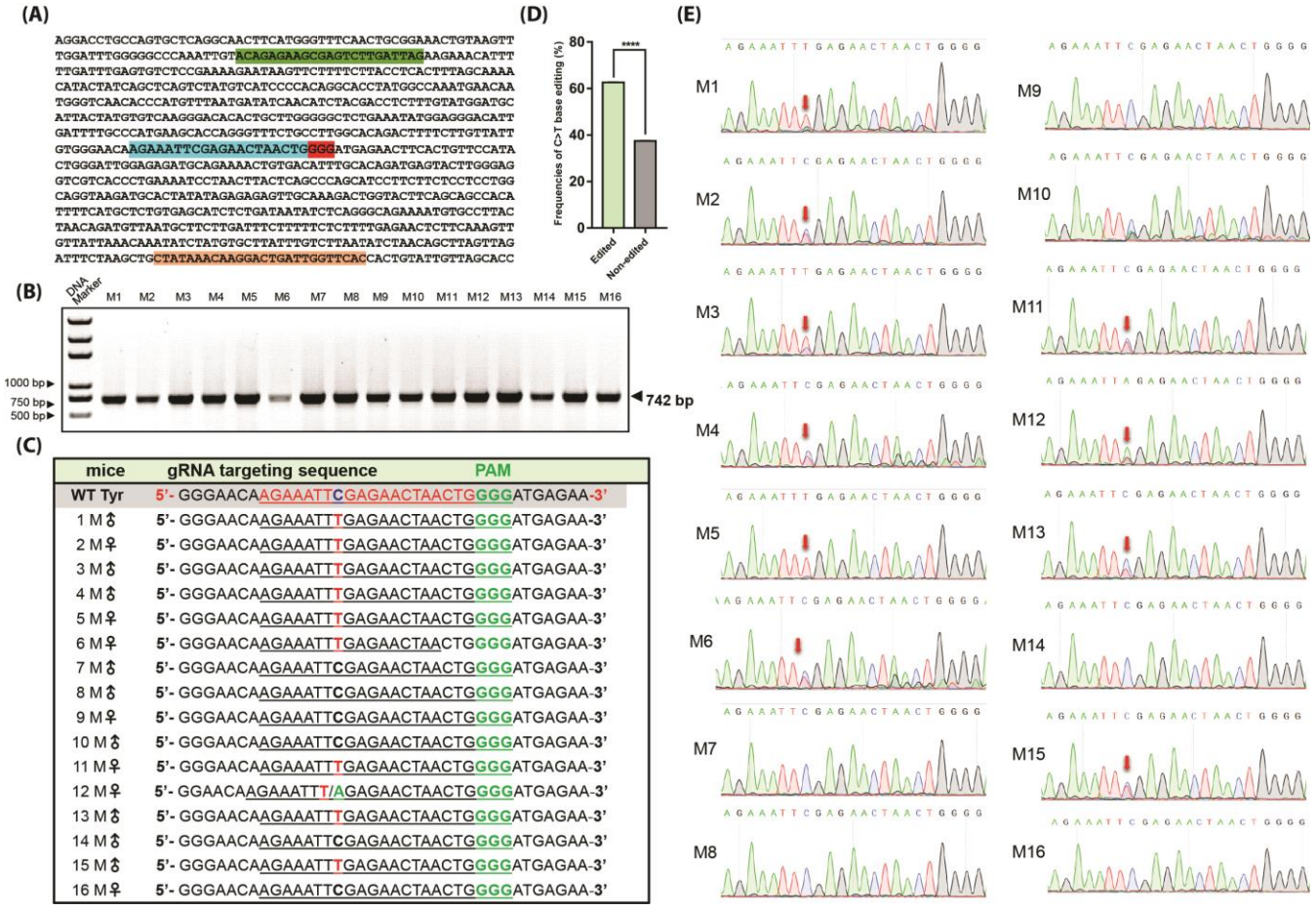


**Figure 3.** Screening of mutations in mice by genomic PCR and sequencing. (A) Sequences of Tyr gene target region in exon 1. SgRNA target sequence in blue and PCR primer sequences in green and orange. (B) Genomic PCR of target regions of founders 1-16 (F0). (C) Alignments of major genomic sequence signals from all founders. C-to-T base substitution is shown in red and green. Wild type in blue. (D) Frequencies of C>T base editing. (E) Chromatograms showing sequencing signals of PCR amplified region. Red arrow shows the base change. *P*-value was determined based on *t*-test. ****P<0.0001.

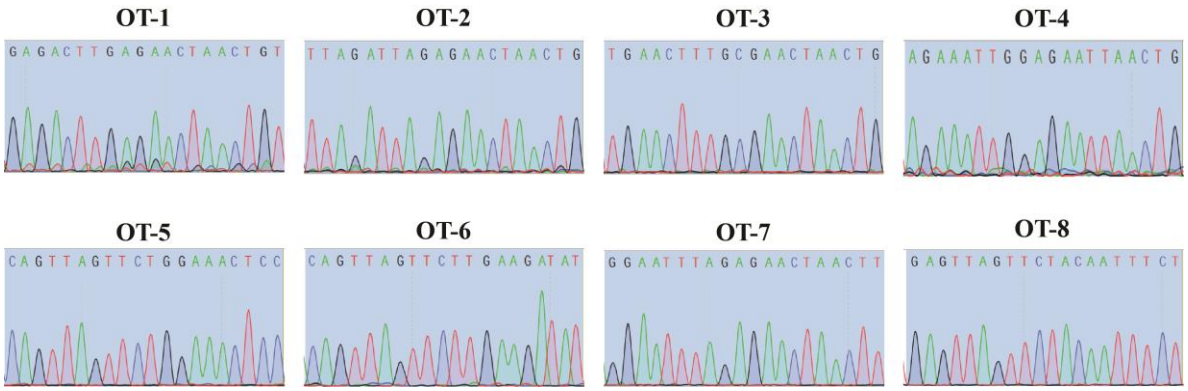


**Figure 4.** Chromatogram sequencing analysis of potential off-target sites (POTs) for sgRNAs predicted according to the online platform (https://benchling.com/)

**Table 2. List of potential off-target sites**

| Off-target | Sequence (5′-3′) | PAM | Score | Mismatch numbers | Chromosome | Strand | Position |
|---|---|---|---|---|---|---|---|
| OT-1 | GAGACTTGAGAACTAACTGT | TAG | 1.514492754 | 3 | chr2 | 1 | 116898149 |
| OT-2 | TTAGATTAGAGAACTAACTG | AGG | 1.490585774 | 4 | chr15 | 1 | 20069149 |
| OT-3 | TGAACTTTGCGAACTAACTG | AAG | 1.45825 | 4 | chr19 | 1 | 43539711 |
| OT-4 | AGAAATTGGAGAATTAACTG | GAG | 0.996830986 | 2 | chr7 | 1 | 74559357 |
| OT-5 | GGAGTTTCCAGAACTAACTG | GAG | 0.92625 | 4 | chr8 | -1 | 46869033 |
| OT-6 | ATATCTTCAAGAACTAACTG | TGG | 0.903189834 | 4 | chr10 | -1 | 103189990 |
| OT-7 | GGAATTTAGAGAACTAACTT | TAG | 0.900340909 | 4 | chr4 | -1 | 24595654 |
| OT8 | AGAAATTGTAGAACTAACTC | AAG | 0.853783 | 3 | chr9 | -1 | 100154656 |

**Table 3. Primers list of potential off-target sites.**

| Off-target | Primer Sequences (5′-3′) | Amplicon (bp) | Annealing temperature (°C) |
|---|---|---|---|
| **OT-1** | Forward: 5′- CTCTACAGCTTGGCTCCTAAAC -3′ | **243** | **57** |
| | Reverse: 5′- GGTGGATTGCTCCAGAAAGA -3′ | | |
| **OT-2** | Forward: 5′-GCTCAGCCTGCTCTCTTATAG-3′ | **420** | **57** |
| | Reverse: 5′-CAATATAGCCATGTATAGCCATGG-3′ | | |
| **OT-3** | Forward: 5′- AAACACGTTCCTAGAGGAGAAAG-3′ | **428** | **57** |
| | Reverse: 5′- GTGTGTGATCAAAGAAGAGGATATTG-3′ | | |
| **OT-4** | Forward: 5′- GCCTGACAATATCTGCCTAACA-3′ | **318** | **57** |
| | Reverse: 5′- GGGAGATCCAGAAAGCAAAGA-3′ | | |
| **OT-5** | Forward: 5′- TTCTTTGTTTGCCTGGGTTTATC-3′ | **495** | **57** |
| | Reverse: 5′- TTATGGGTGCTTGACTCCTTAC-3′ | | |
| **OT-6** | Forward: 5′- ATGCCGTCAATGCCAGTAAG-3′ | **398** | **57** |
| | Reverse: 5′- GCACTTGGGAGGTTAGAGTAGA-3′ | | |
| **OT-7** | Forward: 5′- TGGGAATGTACCTCAGTGTTAG-3′ | **501** | **57** |
| | Reverse: 5′- CGGATGTCTCATATCCCTTCTC-3′ | | |
| **OT-8** | Forward: 5′- CACAAACACCCTAGGATAGCTAAA-3′ | **349** | **57** |
| | Reverse: 5′- GGCCAAAGTCTCTGAAGGTAAA-3′ | | |

## 5. Discussion

Majority of human genetic diseases arise from point mutations. G:C>A:T or T:A>C:G point mutations represent nearly half of all pathogenic single-nucleotide polymorphisms (SNPs) [1,**17**]. While most animal models gene targeting are generated by HDR using embryonic stem cells, which

is time-consuming and costly, CRISPR/Cas9 system has revolutionized the way of generating mouse models. However, CRISPR/Cas9 system can give unpredictable deletion, insertion and off target mutation. Generation of point mutation mouse model is most time consuming with low success rate [17,18]. The point mutation mouse models are the best human disease models that can precisely mimic human pathology. Gene therapy involves mostly correction of point mutations. Previous reports have demonstrated that cytosine base editing (CBE) systems are versatile and potentially applicable in cultured cells, different animal models and plants [17-21]. Moreover, CBE is a secure system that can modify genomic DNA without DSBs with very low off-target effects [10,15]. Yet, applications of base substitutions in animal models are still limited. David Liu has developed a variety of versions of base editing systems but the newest system BE4 has not been tested in vivo.

In this study, we have tested BE4 plasmid along with sgRNA expression plasmid in cell culture and applied to animal model by transgenic microinjection. We designed a precise base editing method which knockout *tyrosinase* gene and results in loss-of-pigmentation (Albinism). High percentage of albino offspring was seen. We achieved successful C>T transition with high efficiency. C>T conversions have occurred exclusively within the approximate editing window of protospacers (positions~4–8). BE-mediated STOP-codon generation disrupts gene function by converting C to T in coding sequences (CAG, CAA, CGA). The study demonstrates that BE4 system is highly efficient and most precise in base editing. In addition, the stop codon generation provides a secure approach to generate knockout animal models with minimum change of genome structure, mimicking most of the genetic diseases [21-23]. Our results highlighted that BE4 system can introduce site-specific and single-base substitution with high precision and efficiency in mouse embryos with no off-target mutation. This study demonstrate the great values of BE4 editing in human disease modeling and future gene therapy.

## 4. Materials and Methods

### *4.1 Animals*

All mice and experimental protocols used in this project has been approved by Institutional Animal Care and Use Committee for Animal Experimental Ethics Committee of Northeast Normal University (NENU/IACUC, AP2018011) and carried out in accordance with recommendations in Guide for Care and Use of Laboratory Animals of National Institutes of Health as well. Mice were bred and maintained under specific pathogen-free condition in animal facility with controlled temperature at 21±1°C, 30%-60% humidity, 12:12 light/dark cycles and free access to food and water.

### *4.2 Reagents*

Chemicals and reagents. Fetal bovine serum (FBS) was purchased from Hangzhou Sijiqing Biological Engineering Material, Co., Ltd. (Beijing, China). Dulbecco's modified Eagle's medium (DMEM) was obtained from Gibco (Thermo Fisher Scientific, Inc., Waltham, MA, USA). Opti-Mem medium, Lipofectamine, dimethyl sulfoxide (DMSO) were purchased from Sigma-Aldrich (Merck KGaA, Darmstadt, Germany). Dip2a and Dip2c antibodies were purchased from Novus and Signalway Antibody, Inc. (USA).

*4.3 Plasmid construction and sgRNA design*

Cas9 coding region of pX330 plasmid (Gifted from Dr. Feng Zhang, Addgene accession no. 42230) was replaced with EGFP cDNA (Fig. 1B). EGFP sequence was PCR amplified from pEGFP-N1 (Clontech cat# 6059-1) (Fig. 1C) using following primers: EGFP-F: 5′-GGCCACCGGTGATCCACCGGTCGCCACCAT-3′ (20bp) and EGFP-R: 5′-GGCCGAATTCTTACTTGTACAGCTCGTCCATG-3′ (22bp) with *AgeI* site at 5′-end and *EcoRI* site at 3′-end (*AgeI* and *EcoRI* are shown by underline). PCR was performed at 94°C for 4min, 24 cycles of 94°C for 30s, 56°C for 30s, 72°C for 1min and 72°C for 10min. EGFP PCR products were digested with *AgeI* and *EcoRI* (NEB) and inserted into *AgeI* and *EcoRI* sites of pX330. Resultant pX330-EGFP plasmid (Fig. 1B) was confirmed by sequencing. Oligos coding for sgRNA targets were synthesized by Genewiz (Beijing, China), annealed at 95°C for 5min and ramped down to 25°C (-5°C/min) and then subcloned into BbsI sites of pX330-EGFP. BE4 plasmid (Fig. 1A) was gifted from David Liu lab (Addgene access no. 100802). The sgRNAs were designed using online platform https://benchling.com/ and all sgRNAs oligos are listed in Supplementary Table 1.

*4.4 Cells culture and EGFP stable expression*

Human embryonic kidney (HEK293) cells were from American Type Culture Collection (ATCC CRL-1573, Manassas, USA) and cultured in Dulbecco's modified Eagle's Medium (DMEM, Sigma) supplemented with fetal bovine serum (10%) and penicillin/streptomycin (Gibco, Life Technologies). Cells were maintained at 37◦C and 5% CO2 in a humidified incubator. To stably express EGFP in HEK293, cells were seeded in 12-well plates with 1ml of DMEM. When cells reached 60–80% confluency, medium was replaced with Opti-MEM (Gibco, Life Technologies). Then cells were transfected using Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol. One μg of pEGFP-N1 was transfected with 2μL Lipofectamine 2000. Medium was replaced with fresh DMEM medium with serum 6hrs after transfection. 48hrs later, cells were treated with G418 (500 μg/mL, Sigma) for 15 days with medium changed every 3 days. Colonies were picked into 96 wells and expanded into 6-well plates before genomic DNA extraction, PCR amplification, and sequencing.

*4.5 Plasmid transfection*

SgRNA oligos (Supplementary Table 1) were annealed and cloned into pX330-EGFP plasmid. HEK293 and B160F10 cells were transfected according to the manufacturer's protocols (Invitrogen, Cat. No. 11668-027). In brief, HEK293 and murine B16-F10 cells were seeded on 12-well plates in 1ml of DMEM. When cells reached 60–80% confluency, medium was changed to Opti-MEM. Cells were then transfected with Lipofectamine 2000 (Invitrogen) according to manufacturer's protocol. One μg pX330-sgRNAs and 2μg BE4 plasmids were mixed with 2μL Lipofectamine 2000. Six hours later, medium was replaced with fresh DMEM. Cells were then subjected to G418 treatment as described above.

*4.6 Oocyte/DNA microinjection and oviduct transfer*

Six-week old F1 female mice (B6D2F1) were obtained from mating of C57BL/6 and DBA2. Mice were superovulated with 10IU of pregnant mare's serum gonadotropin (Ningbo Hormone Products

CO., Ltd, Ningbo, Zhengjiang, China) and followed by 5IU of human chorionic gonadotropin (Ningbo Hormone Products CO., Ltd, Ningbo, Zhengjiang, China) 48hrs later. Superovulated B6D2F1 females were crossed with B6D2F1 males. Fertilized eggs at pronucleus stage were collected in M2 medium. Mixtures of pTyr-gRNAs (2.35ng/ul) and BE4 plasmids (2.64ng/ul) were injected into nucleus in a droplet of M2 medium using inverted microscope equipped with a pair of micromanipulators (Olympus, Tokyo, Japan). Then the injected embryos were incubated in M16 culture medium at 37 °C, 6% $CO_2$ overnight, followed by transfer into the oviduct of a recipient mother at two-cell stage.

*4.7 Genomic DNA extraction and genotyping*

Genomic DNA was extracted from mouse tail tips using G-NTK lysis buffer [24] and proteinase K (1mg/ml) (Beijing Solarbio Science & Technology Co., Ltd., Beijing, China) at 55ºC overnight. Proteinase K was deactivated at 95°C for 15min and PCR was performed in 25µl reaction volume with diluted tail DNA and genotyping primers (supplementary table 2). PCR master mix was as follow: 1.2µl of each primer (10µM), 16.4µl of $ddH_2O$, 1.5µl of 25mM $MgCl_2$, 2.5µl of 10X PCR buffer, 0.5µl of 10mM dNTP Mix and 0.25µl of Taq DNA Polymerase. The PCR conditions were as follows: 95ºC for 5 min, 32 cycles of 95ºC 30sec, 58ºC 30sec and 72ºC 30sec, and 72ºC 10min using PCR machine by Bio-Rad, Hercules, CA, USA.

*4.8 RNA extraction*

One ml RNAiso plus reagent (Takara, Dalian, China) was added to cells on 100mm Petri dish. Cell lysates were collected and incubated at room temperature for 5min. Cells were then centrifuged at 13500 ×g for 5min at 4°C. A 200µL of $CHCl_3$ was added, followed by 30sec mixing and 5min incubation at RT, samples were centrifuged at 13500 ×g for 15min at 4°C to separate RNA into aqueous phase. Aqueous phase (about 600ul) was transferred to a new tube and RNA was precipitated with 750 µL of absolute isopropanol at RT for 10min and then centrifuged at 13500 ×g for 10min at 4°C. Precipitate was washed with 1mL 70% ethanol, followed by centrifugation at 13500 ×g for 5min at 4°C. RNA pellet was resuspended in 50µL of DEPC-treated water. RNA concentrations were determined using NanoDrop 2000 (Thermo Fisher Scientific, USA). RNA integrity was checked on 0.8% agarose gel.

*4.9 RT-qPCR*

One µg of total RNA was reverse-transcribed into first-strand complementary DNA (cDNA) with Prime Script RT Reagent Kit (Perfect Real Time, TaKaRa, Dalian, China) according to the manufacturer's instructions. Real-time PCR was performed with 50ng of cDNA using One-Step SYBR PrimeScriptTM RT-PCR kit (Takara, Dalian, China). All reactions were performed in triplicate. All primers were initially evaluated for efficiency using relative standard curve and electrophoresis on gel. Primer sequences are listed in Supplementary Table 3.

*4.10 Mutation screening by sequencing*

Purified PCR products were extracted using gel extraction kit (Qiagen, Germany) and cloned into pMD18-T plasmid (TaKaRa, Dalian, China). Positive clones were sequenced in two directions utilizing M13 forward and reverse primers. Mutations were identified by alignment to wild-type sequences.

*4.11 Western blot*

Total proteins were extracted using RIPA buffer (0.5% Nonidet P-40, 0.1% sodium deoxycholate, 150mM NaCl, 50mM Tris-Cl, pH7.5 and 1x protease inhibitor cocktail). Cell lysates were subjected to high-speed centrifugation at 12000x*g* for 15min at 4°C. Protein concentrations were measured using Coomassie (Bradford) protein assay kit. Total soluble proteins were then separated on 10% SDS-PAGE and transferred into polyvinylidene difluoride membrane (Millipore, Billerica, MA). Membrane was blocked with 5% nonfat dry milk for 1h followed by incubation with diluted the primary antibodies (β-actin, 1:2000, Signalway antibody; Dip2a, 1:500, Novus; Dip2c, 1:1000, Abcam) for overnight at 4°C. Then the membrane was washed in TBST for three times, 5 min each and then incubated with secondary antibody (anti-rabbit horseradish peroxidase conjugate, 1:5,000; anti-mouse horseradish peroxidase conjugate, 1:5,000; Transgene) for 30min, followed by washing three times with TBST. Signals were detected using enhanced chemiluminescence Amersham™ ECL™ (GE Healthcare, USA) reagents. β-actin protein served as a loading control.

*4.12 Off-target detection*

Eight potential off-target sites (POTs) were identified according to an online design tool (https://benchling.com/). Selected POTs (Table 2) were amplified by PCR and sequenced. Sequences were compared with wild type. All primers used for off-target assay were listed in Table 3.

*4.13 Statistical Analysis*

Statistical analyses and graphics were performed with GraphPad Prism 5.01 (GraphPad Software Inc.) and SPSS software version 25.0 (IBM Inc., New York, USA). Parametric unpaired Student's *t* test was used to assess difference between the groups. *P*-values were two-sided; a *P*-value < 0.05 was considered statistically significant

**Supplementary Materials: Supplementary materials can be found online.**

**Author Contributions:** "Conceptualization, S.A., L.Z., Y.Z. and X.F.; methodology S.A., P.Y., Y.C. R.K.S, and M.A.; software S.A., Z.O., M.Z.Z.M, R.K.S., M.A. and F.B.B.; validation L.Z., Y.Z., and X.F.; formal analysis, S.A., F.H, N.B. and M.A.; investigation, S.A., P.Y., Y.C.; resources, Y.Z. and X.F.; writing—original draft preparation, S.A., L.Z., Y.Z. and X.F.; writing—review and editing, S.A., L.Z., Y.Z. and X.F.; visualization, S.A.; supervision, L.Z., Y.Z. and X.F.; project administration, L.Z., Y.Z. and X.F.; funding acquisition, L.Z., Y.Z. and X.F. All authors have read and agreed to the published version of the manuscript."

**Funding:** This work was funded supported in part or in whole by National Natural Science Foundation of China (81270953), the Natural Science Foundation of Jilin Province (20160101344JC) and Science and Technology Project of Jilin Provincial Education Department (JJKH20180023KJ). The funders had no roles in this study, including study design, data collection, data analysis or decision to publish or preparation of the manuscript.

**Acknowledgments:** We are grateful to Ms. Huiyan Wu for microinjection and managing mouse colony.

**Conflicts of Interest:** The authors declare that they have no competing interests.

## Abbreviations

| | |
|---|---|
| CRISPR | Clustered Regularly Interspaced Short Palindromic Repeats |
| Cas9 | CRISPR-associated protein 9 |
| SNPs | Single-Nucleotide Polymorphisms |
| BE4 | Base Editor4 |
| DSBs | DNA double-strand breaks |
| HDR | Homology-directed repair |
| NHEJ | Nonhomologous end joining |
| CBEs | Cytosine Base Editors |
| SgRNA | Single guided RNA |
| EGFP | Enhanced Green Fluorescence Protein |
| POTs | Potential Off-Target sites |
| UGI | Uracil Glycosylase Inhibitor |
| HEK293 | Human Embryonic Kidney 293 |
| Dip2a | Disco Interacting Protein 2 Homolog a |
| Dip2c | Disco Interacting Protein 2 Homolog c |

## References


1. Gaudelli, N.M.; Komor, A.C.; Rees, H.A.; Packer, M.S.; Badran, A.H.; Bryson, D.I.; Liu, D.R. Programmable base editing of A• T to G• C in genomic DNA without DNA cleavage. *Nature* **2017**, *551*, 464.
2. Rath, D.; Amlinger, L.; Rath, A.; Lundgren, M. The CRISPR-Cas immune system: biology, mechanisms and applications. *Biochimie* **2015**, *117*, 119-128.
3. Sontheimer, E.J.; Barrangou, R. The bacterial origins of the CRISPR genome-editing revolution. *Hum Gene Ther* **2015**, *26*, 413-424.
4. Wu, H.; Liu, Q.; Shi, H.; Xie, J.; Zhang, Q.; Ouyang, Z.; Li, N.; Yang, Y.; Liu, Z.; Zhao, Y. Engineering CRISPR/Cpf1 with tRNA promotes genome editing capability in mammalian systems. *Cell Mol Life Sci* **2018**, *75*, 3593-3607.
5. Yang, Y.; Wang, K.; Wu, H.; Jin, Q.; Ruan, D.; Ouyang, Z.; Zhao, B.; Liu, Z.; Zhao, Y.; Zhang, Q. Genetically humanized pigs exclusively expressing human insulin are generated through custom endonuclease-mediated seamless engineering. *J Mol Cell Biol* **2016**, *8*, 174-177.
6. Yang, L.; Guell, M.; Byrne, S.; Yang, J.L.; De Los Angeles, A.; Mali, P.; Aach, J.; Kim-Kiselak, C.; Briggs, A.W.; Rios, X. Optimization of scarless human stem cell genome editing. *Nucleic Acids Res* **2013**, *41*, 9049-9061.
7. Yin, H.; Xue, W.; Chen, S.; Bogorad, R.L.; Benedetti, E.; Grompe, M.; Koteliansky, V.; Sharp, P.A.; Jacks, T.; Anderson, D.G. Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nat Biotechnol* **2014**, *32*, 551.
8. Davis, A.J.; Chen, D.J. DNA double strand break repair via non-homologous end-joining. *Translational cancer research* **2013**, *2*, 130.
9. Komor, A.C.; Zhao, K.T.; Packer, M.S.; Gaudelli, N.M.; Waterbury, A.L.; Koblan, L.W.; Kim, Y.B.; Badran, A.H.; Liu, D.R. Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C: G-to-T: A base editors with higher efficiency and product purity. *Sci Adv* **2017**, *3*, eaao4774.
10. Komor, A.C.; Kim, Y.B.; Packer, M.S.; Zuris, J.A.; Liu, D.R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **2016**, *533*, 420.

11. Nishida, K.; Arazoe, T.; Yachie, N.; Banno, S.; Kakimoto, M.; Tabata, M.; Mochizuki, M.; Miyabe, A.; Araki, M.; Hara, K.Y. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* **2016**, *353*, aaf8729.
12. Ma, Y.; Zhang, J.; Yin, W.; Zhang, Z.; Song, Y.; Chang, X. Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells. *Nat Methods* **2016**, *13*, 1029.
13. Xie, J.; Ge, W.; Li, N.; Liu, Q.; Chen, F.; Yang, X.; Huang, X.; Ouyang, Z.; Zhang, Q.; Zhao, Y. Efficient base editing for multiple genes and loci in pigs using base editors. *Nat Commun* **2019**, *10*, 2852.
14. Kim, K.; Ryu, S.-M.; Kim, S.-T.; Baek, G.; Kim, D.; Lim, K.; Chung, E.; Kim, S.; Kim, J.-S. Highly efficient RNA-guided base editing in mouse embryos. *Nat Biotechnol* **2017**, *35*, 435.
15. Kim, D.; Lim, K.; Kim, S.-T.; Yoon, S.-h.; Kim, K.; Ryu, S.-M.; Kim, J.-S. Genome-wide target specificities of CRISPR RNA-guided programmable deaminases. *Nat Biotechnol* **2017**, *35*, 475.
16. Liu, Z.; Lu, Z.; Yang, G.; Huang, S.; Li, G.; Feng, S.; Liu, Y.; Li, J.; Yu, W.; Zhang, Y. Efficient generation of mouse models of human diseases via ABE-and BE-mediated base editing. *Nat Commun* **2018**, *9*, 2338.
17. Liu, Z.; Chen, M.; Chen, S.; Deng, J.; Song, Y.; Lai, L.; Li, Z. Highly efficient RNA-guided base editing in rabbit. *Nat Commun* **2018**, *9*, 2717.
18. Zhang, Y.; Qin, W.; Lu, X.; Xu, J.; Huang, H.; Bai, H.; Li, S.; Lin, S. Programmable base editing of zebrafish genome using a modified CRISPR-Cas9 system. *Nat Commun* **2017**, *8*, 118.
19. Zong, Y.; Wang, Y.; Li, C.; Zhang, R.; Chen, K.; Ran, Y.; Qiu, J.-L.; Wang, D.; Gao, C. Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. *Nat Biotechnol* **2017**, *35*, 438.
20. Li, G.; Liu, Y.; Zeng, Y.; Li, J.; Wang, L.; Yang, G.; Chen, D.; Shang, X.; Chen, J.; Huang, X. Highly efficient and precise base editing in discarded human tripronuclear embryos. *Protein Cell* **2017**, *8*, 776-779.
21. Yang, G.; Zhu, T.; Lu, Z.; Li, G.; Zhang, H.; Feng, S.; Liu, Y.; Li, J.; Zhang, Y.; Chen, J. Generation of isogenic single and multiplex gene knockout mice by base editing-induced STOP. *Sci Bull* **2018**, *63*, 1101-1107.
22. Kuscu, C.; Parlak, M.; Tufan, T.; Yang, J.; Szlachta, K.; Wei, X.; Mammadov, R.; Adli, M. CRISPR-STOP: gene silencing through base-editing-induced nonsense mutations. *Nat Methods* **2017**, *14*, 710.
23. Billon, P.; Bryant, E.E.; Joseph, S.A.; Nambiar, T.S.; Hayward, S.B.; Rothstein, R.; Ciccia, A. CRISPR-mediated base editing enables efficient disruption of eukaryotic genes through induction of STOP codons. *Mol Cell* **2017**, *67*, 1068-1079. e1064.
24. Malumbres, M.; Mangues, R.; Ferrer, N.; Lu, S.; Pellicer, A. Isolation of high molecular weight DNA for reliable genotyping of transgenic mice. *Biotechniques* **1997**, *22*, 1114-1119.