

# Efficient Gene Silencing by Adenine Base Editor-Mediated Start Codon Mutation

Xinjie Wang,<sup>1,7</sup> Zhiwei Liu,<sup>2,3,7</sup> GuangLei Li,<sup>4,7</sup> Lu Dang,<sup>5</sup> Shisheng Huang,<sup>4</sup> Lei He,<sup>2,3</sup> Yu'e Ma,<sup>2,3</sup> Cong Li,<sup>2,3</sup> Ming Liu,<sup>6</sup> Guang Yang,<sup>4</sup> Xingxu Huang,<sup>4</sup> Fei Zhou,<sup>2,3</sup> and Xiaodong Ma<sup>1</sup>

<sup>1</sup>Institute for Brain Research and Rehabilitation, Guangdong Key Laboratory of Mental Health and Cognitive Science, Center for Studies of Psychological Application, South China Normal University, Guangzhou 510631, China; <sup>2</sup>Cambridge-Suda Genomic Resource Center, Soochow University, Suzhou 215123, China; <sup>3</sup>Jiangsu Key Laboratory of Neuropsychiatric Diseases Research, Soochow University, Suzhou 215123, China; <sup>4</sup>School of Life Science and Technology, ShanghaiTech University, 100 Haike Road, Pudong New Area, Shanghai 201210, China; <sup>5</sup>Affiliated Cancer Hospital & Institute of Guangzhou Medical University, 78 Hengzhigang Road, Guangzhou 510095, China; <sup>6</sup>State Key Laboratory of Respiratory Disease, Guangzhou Institute of Respiratory Disease, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou 510120, China

**Traditional CRISPR/Cas9-based gene knockouts are generated by introducing DNA double-strand breaks (DSBs), but this may cause excessive DNA damage or cell death. CRISPR-based cytosine base editors (CBEs) and adenine base editors (ABEs) can facilitate C-to-T or A-to-G exchanges, respectively, without DSBs. CBEs have been employed in a gene knockout strategy: CRISPR-STOP or i-STOP changes single nucleotides to induce in-frame stop codons. However, this strategy is not applicable for some genes, and the unwanted mutations in CBE systems have recently been reported to be surprisingly significant. As a variant, the ABE systems mediate precise editing and have only rare unwanted mutations. In this study, we implemented a new strategy to induce gene silencing (i-Silence) with an ABE-mediated start codon mutation from ATG to GTG or ACG. Using both *in vitro* and *in vivo* model systems, we showed that the i-Silence approach is efficient and precise for producing a gene knockout. In addition, the i-Silence strategy can be employed to analyze ~17,804 human genes and can be used to mimic 147 kinds of pathogenic diseases caused by start codon mutations. Altogether, compared to other methods, the ABE-based i-Silence method is a safer gene knockout strategy, and it has promising application potential.**

## INTRODUCTION

Standard CRISPR-mediated gene knockout strategies rely on the repair of DNA double-strand breaks (DSBs) by non-homologous end joining (NHEJ) or homology-directed repair (HDR), which may lead to uncontrolled and unexpected DSB-induced toxicity.<sup>1–3</sup> By fusing a cytidine or adenosine deaminase enzyme with a Cas9 nickase, the newly discovered cytosine and adenine base editor (CBE and ABE) systems can introduce precise single base conversion of C-to-T or A-to-G to modify the genome without a double-strand break.<sup>4,5</sup> The CBE and ABE systems are now being used in different ways to generate genetically modified models that mimic human diseases.<sup>6–9</sup> Recently, two groups reported an alternative approach to generate gene knockouts by introducing early stop codons using CBEs without DSBs, named CRISPR-STOP<sup>10</sup> and i-STOP.<sup>11</sup> It was

reported that pretermination of specific genes by CBE-introduced stop codons could result in the efficient knocking out of a gene in cells or model animals.<sup>10,11</sup> However, this strategy also has limitations. Early termination of gene expression usually targets the first 20% of the coding region to avoid the production of truncated protein variants.<sup>12</sup> However, as revealed by high-throughput analysis, only approximately 80% of human gene open reading frames (ORFs) are targetable, so there are still some genes that cannot be knocked out by i-STOP.<sup>12</sup> Furthermore, recent publications also reported concerns about unwanted mutations being introduced at the DNA level by CBEs but not by ABEs.<sup>13,14</sup> Until now, the ABE system has been shown to facilitate efficient and precise A-to-G conversion with rare unwanted mutations, which makes it especially useful for generating desired genetically modified model organisms. Considering the methodological limitations of the CBE-mediated gene knockout, we tried to explore another alternative method.

In this study, we developed an efficient strategy that is safer than CBE-mediated editing; the developed strategy for gene disruption is based on ABE-mediated precise editing. Induced gene silencing (i-Silence) was achieved by introducing a mutation to the start codon (ATG to GTG or ACG) using the ABE system. To make a comprehensive demonstration of the method, we tested this strategy both *in vitro* and *in vivo*. Our results show that the ABE system can induce efficient modification at the start codon of targeted genes, leading to successful gene silencing in both cultured cells and mice, which indicates that this strategy is applicable to generate gene knockout models. Meanwhile, whole-genome

Received 29 May 2019; accepted 23 November 2019;  
<https://doi.org/10.1016/j.jmthe.2019.11.022>.

<sup>7</sup>These authors contributed equally to this work.

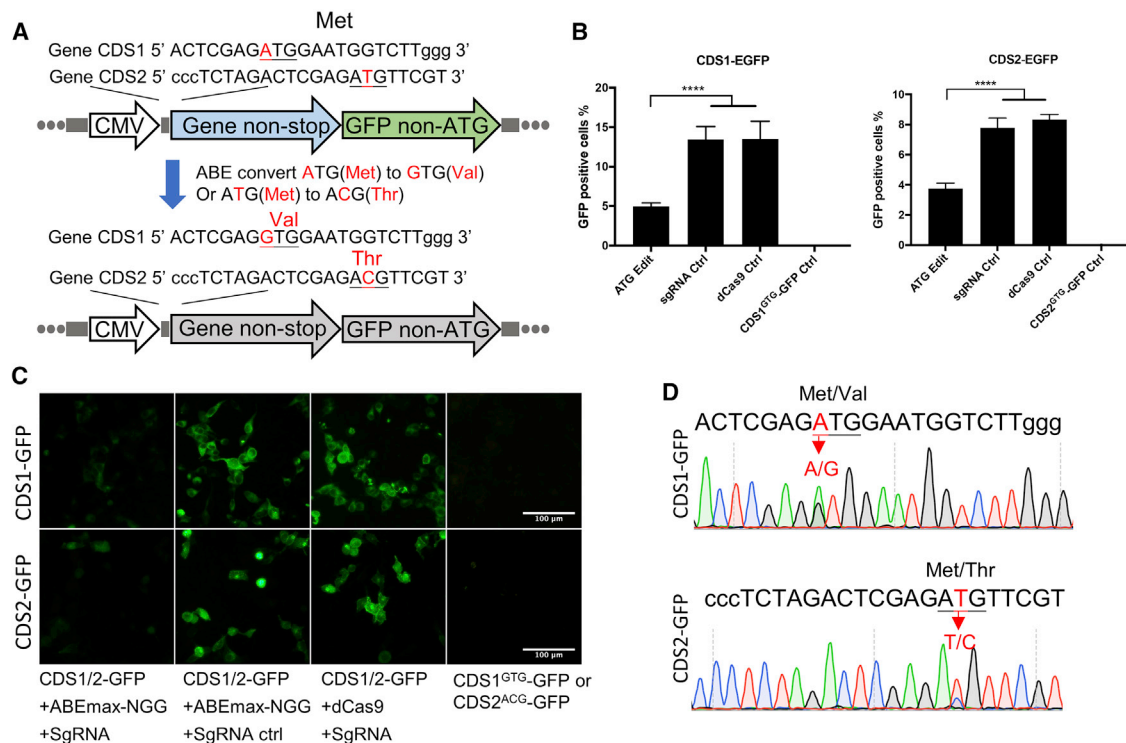
**Correspondence:** Xiaodong Ma, Institute for Brain Research and Rehabilitation, Guangdong Key Laboratory of Mental Health and Cognitive Science, Center for Studies of Psychological Application, South China Normal University, Guangzhou 510631, China.

**E-mail:** [sciencema@hotmail.com](mailto:sciencema@hotmail.com)

**Correspondence:** Fei Zhou, Cambridge-Suda Genomic Resource Center, Soochow University, Suzhou 215123, China.

**E-mail:** [fzhou@suda.edu.cn](mailto:fzhou@suda.edu.cn)





**Figure 1. Verification of Gene Disruption by ABE-Mediated i-Silence Using an *In Vitro* EGFP Reporter System**

(A) Schematic representation of the EGFP reporter system and targeting strategy. The gene coding sequence (CDS) without a stop codon was cloned and fused with the EGFP gene without a start codon. The ABE-mediated i-Silence technique was designed to target the start codon (ATG) of the fusion protein, and the targeting efficacy was monitored by assessing the EGFP signal. (B and C) Analysis of the gene silencing results were determined by FACS analysis (B) and fluorescence microscopy (C). Cells transfected with scrambled sgRNA, dCas9, and the mutant CDS1<sup>GTG</sup>-GFP or CDS2<sup>ACG</sup>-GFP plasmids were included as control groups. Data were obtained from three independent experiments ( $n = 3$ ) and were analyzed by Student's *t* tests (shown as the mean  $\pm$  SEM). Scale bars, 100  $\mu$ m. (D) Sanger sequencing chromatographs show the targeted CDS from DNAs that were isolated from the transfected cells. The PAM sequence is shown in lowercase, and the targeted substituted sequence is shown in red and is underlined. The red arrow indicates the substituted base.

sequencing (WGS) indicates that the ABE-mediated i-Silence method generated no off-target effects in the edited mouse genome. In summary, the i-Silence technique is mediated by the ABE system and can be used as an efficient and precise gene disruption strategy for functional genetics studies and disease model generation.

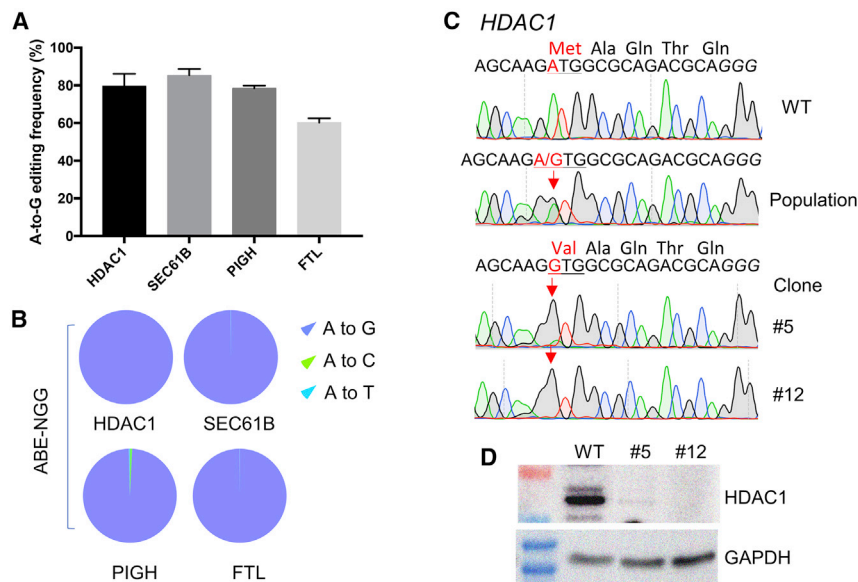
## RESULTS

### i-Silence Induces Successful Gene Disruption in a Reporter System

To test the methodological feasibility of ABE-mediated i-Silence, we first tested the strategy in an enhanced green fluorescent protein (EGFP)-based reporter system. Basically, the coding sequence of EGFP without the start codon ATG was PCR amplified and fused to the C-terminal end of the target gene coding sequence (CDS) without the stop codon. Fusion protein expression could be monitored by EGFP signals, which would be decreased upon i-Silence-mediated start codon alteration. To validate the strategy, two transcript variants, CDS1-GFP and CDS2-GFP, were included in the test. For ABE-mediated targeting, the single-guide RNA (sgRNA) was designed to target the ATG start codon (Figure 1A). Then, the

reporter plasmid was transfected into HEK293T cells together with plasmids expressing ABEmax, sgRNAs, or control vectors. Meanwhile, start codon mutated reporter plasmids (CDS1<sup>GTG</sup>-GFP or CDS2<sup>ACG</sup>-GFP) were used as negative controls.

EGFP fluorescence was analyzed 48 h after transfection. When cotransfected with ABEmax and the targeting sgRNA, the EGFP reporter expression was significantly decreased compared with what was observed in cells transfected with the control sgRNA or dead Cas9 (dCas9), as revealed by flow cytometry (Figure 1B; Figure S1) and fluorescence microscopic analyses (Figure 1C). Then, transfected cells were harvested, and DNA samples were extracted. As further confirmed by Sanger sequencing, the ATG start codon in some of the cells was successfully modified to GTG or ACG (Figure 1D). To further verify that the decrease of protein expression was not caused by ABEmax binding to the start codon, we performed the analysis on both day 2 and day 4 post-transfection. As shown in Figure S2, the GFP-positive rate of the i-Silence-edited group was significantly decreased compared with the ABEmax-mut (with dCas9) group, the scramble sgRNA group, and the reporter-only group. These results



**Figure 2. ABE-Mediated i-Silence Induces Efficient and Precise Endogenous Gene Knockout in HEK293T Cells**

(A) Quantification of ABEmax-induced A-to-G editing frequency of target genes in HEK293T cells. Data are shown as the mean  $\pm$  SD from three independent experiments ( $n = 3$ ). (B) Characterization of the targeted base editing product purity. (C) Sanger sequencing analysis of the start codon locus of HDAC1-targeted cells. WT, WT HEK293T cells; population, mixed cell population after ABE-mediated i-Silence; clone #5 and #12, single clones isolated from i-Silence-edited cells. The on-target start codon ATG is underlined, and the A-to-G conversion is indicated with an arrow. (D) Analysis of HDAC1 expression in i-Silence-edited cell clone #5 and #12, as assessed by western blot. GAPDH was used as a protein loading control.

demonstrated successful gene silencing by i-Silence-mediated start codon editing *in vitro*.

#### Efficient Endogenous Gene Knockout in HEK293T Cells by i-Silence

To further verify the efficiency of ABE-mediated gene i-Silence, four endogenous genes (*HDAC1*, *SEC61B*, *PIGH*, and *FTL*) were selected for validation in HEK293T cells. Both *HDAC1* and *SEC61B* cannot be knocked out by the i-STOP strategy, and *PIGH* and *FTL* were reported to have start codon ATG mutations in clinical case reports.<sup>15,16</sup> The corresponding sgRNAs were cotransfected into cultured cells together with the ABEmax construct. According to the deep sequencing data, the editing frequencies for all genes ranged from 60% to 80% (Figure 2A). The A-to-G substitution occupied more than 99% of all ABE-mediated base editing products, and the ratio of unwanted byproducts A-to-C or A-to-T was less than 0.5% (Figure 2B). The Sanger sequencing results also concur with the conclusion above (Figure S4). For *HDAC1* gene targeting, we subcloned the edited cells for further verification. The i-Silence editing in clone nos. 5 and 12 was confirmed by Sanger sequencing, and the protein expression levels were analyzed by western blot, which indicated successful sequence modification and loss of expression (Figures 2C and 2D). The *SEC61B* gene was also confirmed to be knocked out in HEK293T cells by Sanger sequencing and western blotting (Figure S5). In conclusion, validation with multiple target genes showed that the ABE-mediated i-Silence strategy is applicable for efficient silencing of endogenous genes, and unwanted editing is rare.

#### Successful Gene Knockout in Mice by i-Silence

We then tried to determine whether i-Silence is an efficient approach to knock out a gene *in vivo*. The mouse gene programmed cell death protein 1 (*PD-1*) was chosen as the target (Figure 3A). First, the designed targeting system, especially the

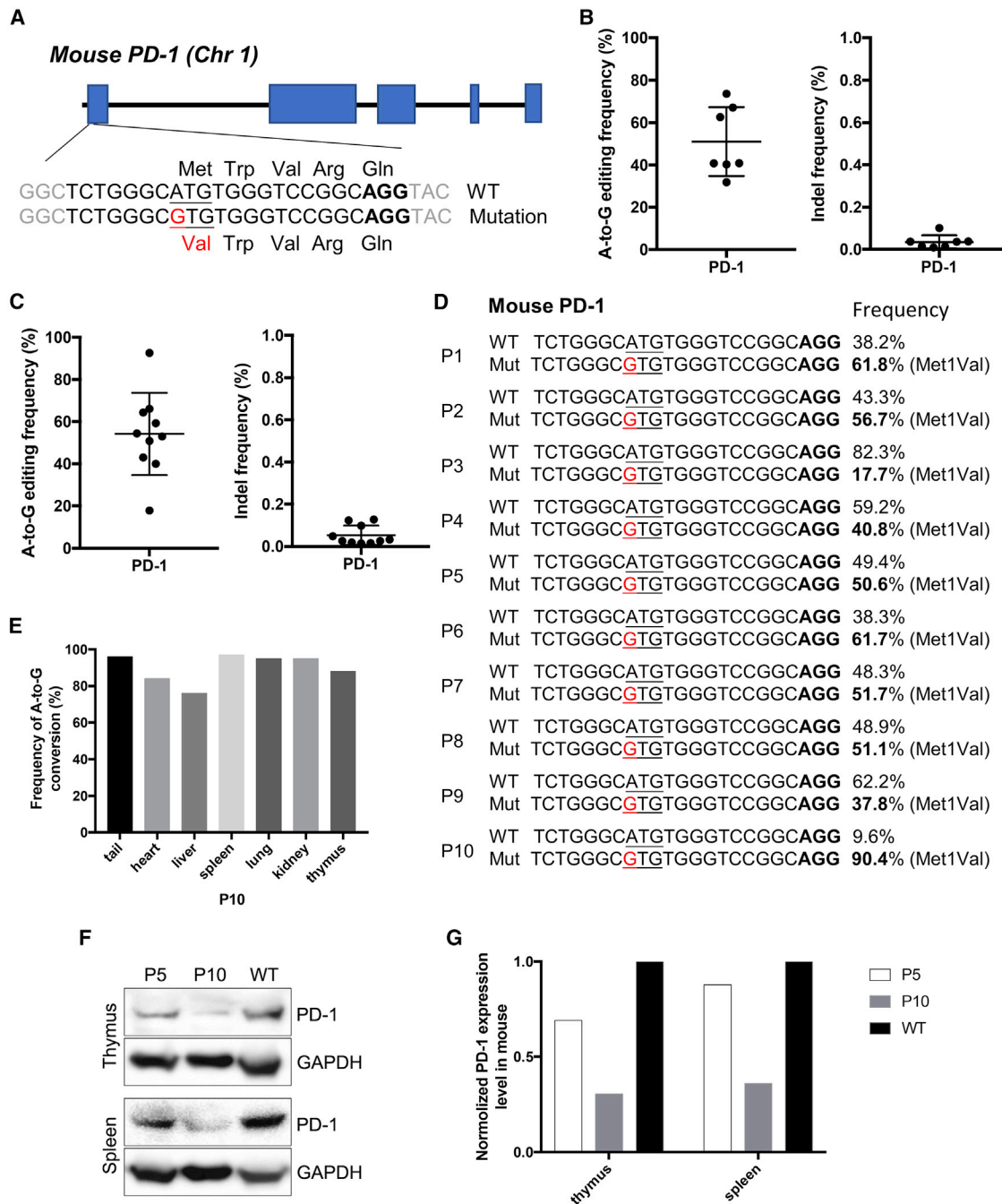
sgRNA, was validated in mouse Neuro-2a (N2a) cells. The statistical results of deep sequencing indicated that the ATG start codon editing efficiency was approximately 33%, and the indel frequency was low (Figure S6). Therefore, the mRNA of ABEmax combined with the sgRNA was injected into 10 mouse zygotes. Seven injected zygotes were cultured *in vivo* to the blastocyst stage, and the genomic DNA was extracted for further analysis (Table S1). Notably, the average editing efficiency ranged from 31.8% to 73.6% in different embryos, and the indel frequency was less than 0.1% (Figure 3B).

Based on the successful start codon mutation of mouse zygotes, we generated a *PD-1* knockout mouse using the i-Silence strategy. A total of 49 embryos were injected with the same materials and then transferred into surrogate mice. As a result, 10 pups (named P1 to P10) were born, and genomic DNA was extracted from the tails of each and was analyzed by deep sequencing (Table 1). As confirmed by Sanger sequencing (Figure S7) and deep sequencing (Figures 3C and 3D), all the pups were successfully edited and had few indels. The highest editing efficiency was 90.4%, which was seen in mouse no. 10 (P10). Furthermore, the editing efficiency was confirmed in six more organs, and the results indicated that the *PD-1* start codon editing was the same in these tissues as it was in the tail (Figure 3E; Figure S8D).

We further validated the *PD-1* protein expression level by western blot using samples from the edited pups (P5 and P10). The *PD-1* protein level of P10 was significantly decreased in both the thymus and spleen compared to the levels observed in the wild-type (WT) control; P5 showed an intermediate expression level (Figures 3F and 3G). The protein level confirmed successful editing of the *PD-1* gene by ABE-mediated i-Silence.

#### Off-Target Analysis by WGS and Deep Sequencing

To further verify the precision of ABE-mediated i-Silence, the off-target effects were investigated by WGS and amplicon-based deep sequencing. WGS was performed on the *PD-1* gene edited mouse



**Figure 3. Efficient PD-1 Knockout in Mice by i-Silence**

(A) Scheme of the ABE-mediated i-Silence strategy for the mouse *PD-1* gene. The upper sequence indicates the reference sequence, while the lower sequence indicates the expected mutant sequence, and the adjacent sequence is shown in gray. The PAM sequence and the targeted start codon are bold and underlined, respectively. The targeted A-to-G conversion is shown in red. The corresponding amino acid is shown up or under the sequence. (B) Editing efficiency and indel frequency of PD-1 on mouse embryos are shown. (C) Editing efficiency and indel frequency of PD-1 on mouse pups are shown. (D) Representative alignments of modified sequences from 10 i-Silence-edited PD-1 mutant newborn pups. The PAM sequences and substitutions are bold and underlined, respectively. The A-to-G mutation is highlighted in red, and the corresponding amino acid that is changed is shown on the right. (E) The conversion frequency of the PD-1 start codon in different organs of P10 mice is shown. (F) Western blot and (G) band quantification analysis were performed to detect PD-1 protein levels in the thymus and spleen of P5 and P10 mice.

**Table 1. Summary of the Manipulation and Genotyping of Newborn Pups**

Target Gene	Methods	No. of Examined Embryos	No. of Transferred Embryos	No. of Offspring (%)	Mutant Ratio (%)	
					Mutants/Tested Offspring	No. of Indels Mutants/Total Mutants
PD-1	microinjection <sup>a</sup>	50	49 <sup>b</sup>	10 (20) <sup>c</sup>	10/10 (100)	0/10 (0)

<sup>a</sup>sgRNA of PD-1 + ABE<sub>max</sub>-NGG mRNA.<sup>b</sup>One embryo died during the microinjection.<sup>c</sup>Calculated from the total number of transferred embryos.

(P10) and a WT mouse at mean depths of 34.7× and 24.7×, respectively (Figure 4E). A total of 48,440 and 49,568 single-nucleotide polymorphisms (SNPs) were detected for mouse P10 and the WT mouse (Figure 4A). Only 3,076 unique SNPs were found in the mouse P10 genome after filtering out the database of SNPs (dbSNPs) and WT SNPs. There was not a significant increase in the number of SNPs in the ABE-edited mouse P10 compared to what was observed in the WT mouse (Figure 4A). To analyze the potential for off-target effects, 0-, 1-, and 2-base mismatches in the seed region and 0–7 bp in the non-seed region were analyzed. As a result, a total of 10,911 sites were analyzed, and no off-target changes were identified (Figures 4C and 4D). To further explore the potential off-target effects generated by the i-Silence strategy, deep sequencing was used on three PD-1 mutant mice with the highest editing efficiency (P10, P5, and P6) and a WT mouse. The potential off-target sites were predicted with up to 4-base mismatches in the mouse genome, and the results showed no 0-, 1- or 2-base mismatches (Table S6). Eighteen top off-target sites on the list were selected and subjected to deep sequencing with an average over 900,000 in depth. The results showed no base substitution at any of these sites compared with the WT control (Figure 4F). In summary, the WGS and off-target deep-sequencing results demonstrated the precision of ABE-mediated i-Silence editing *in vivo*.

#### ABE-Mediated i-Silence Strategy Enables the Knockout of Human Genes on a Genome Scale

To determine whether i-Silence could be utilized for genome-wide studies of human genes, we analyzed all of the coding genes reported in the human reference genome (NCBI). A gene that is targetable by i-Silence means that all of its transcripts can be targeted. An i-STOP/CRISPR-STOP gene is targetable if the first 20% of its coding sequence is targetable, which avoids the production of truncated protein.<sup>12</sup> Remarkably, a total of 17,804 genes are targetable with i-Silence using six different protospacer-adjacent motifs (PAMs), which accounts for approximately 92.4% of human genes (Figure 5A). Among all i-Silence targetable genes, 20% of genes (5,025/19,247) have potential downstream ATG codons within the first 3% of their sequence (Figure 5B). We further analyzed the coverage of the i-Silence and i-STOP strategies on the human genome. As visualized in a Venn diagram (Figure 5C), a total of 16,730 genes can be targeted by both strategies, whereas 118 genes are not targeted by either strategy. Meanwhile, 1,074 genes can only be targeted by i-Silence and 1,345 genes can only be targeted by i-STOP. These results indicate

that the i-Silence strategy is a good complement to i-STOP for achieving genome-wide targeting.

More importantly, start codon mutations are frequently associated with human disease. To determine whether i-Silence could be utilized to model ATG mutations, we analyzed the human variation database (ClinVar) and found 247 human diseases with reported start codon mutations in specific genes. Among these, 147 genes can be modeled by the i-Silence strategy (Figure 5D).

In summary, these analyses demonstrate that the ABE-mediated i-Silence strategy is a promising approach for genome-wide functional studies, and it also has potential clinical value for modeling human diseases caused by start codon mutations in specific genes.

#### DISCUSSION

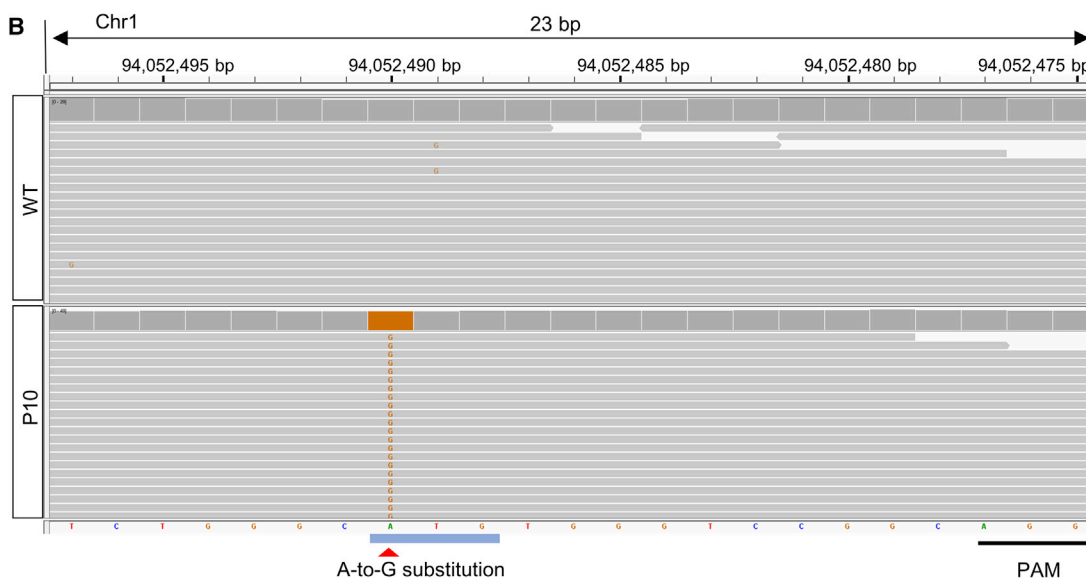
In this study, we designed an alternative approach to induce gene silencing by mutating the start codon with ABE-mediated editing. Using a series of validation tests performed in both cultured cell lines and a mouse model, we demonstrated that the i-Silence strategy can facilitate efficient mutation of a start codon, leading to successful disruption of gene expression. WGS analysis revealed that ABE-mediated i-Silence is also a safer strategy to precisely induce targeted base editing without generating unwanted edits. Therefore, i-Silence induced by ABE is a good complement to CBE-mediated gene knockout strategies.

The i-Silence strategy provides possible approaches for addressing some challenging biological questions. For example, i-Silence could allow investigation of the function of different isoforms of a gene by mutating the start codon of the specific isoform. The protein translation may start from a downstream ATG codon when the original start codon is mutated. According to our analysis, 20% of i-Silence targetable genes (5,025/19,247) have downstream ATG codons within the first 3% of their sequence (Figure 5B). To validate whether there is unwanted protein produced when the original start codon ATG is mutated, one or two nucleotide insertions between the CDS2 and GFP genes (Figure S3A) were constructed to validate the method by assessing GFP expression. Both incorporated GFP reporters (CDS2<sup>ATG</sup>-G-GFP and CDS2<sup>ATG</sup>-GT-GFP) showed no GFP expression (Figure S3B). To further investigate the function of alternative ATG codons, we performed an experiment by constructing three kinds of ATG codons in the CDS2<sup>ACG</sup>-GFP reporter (Figure S3C).

**A**

	WT	P10	Uniquely assigned to P10
total SNPs	49,568	48,440	
SNPs after excluding dbSNPs	16,518	15,078	3,076
A>G SNPs	1,785	1,589	220
A>C SNPs	1,198	1,171	343
A>T SNPs	975	895	187
T>C SNPs	1,725	1,555	229
T>G SNPs	882	846	197
T>A SNPs	1,397	1,264	360

WT: Wild type



**C**

Mismatch in seed region	Mismatch in non-seed region	No. of potential on-/off-target sites
0	0	1
0	0-7	8
1	0-7	395
2	0-7	10508
Total		1/10911

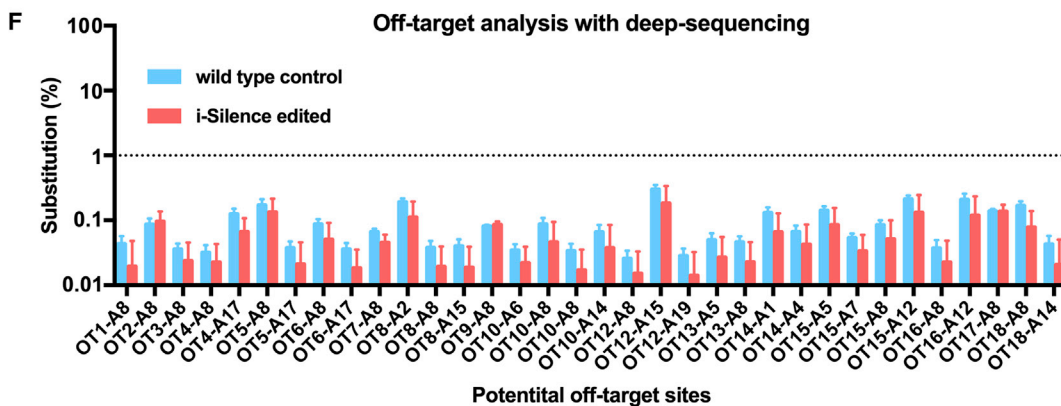
**D**

Sample	On-target	Off-target
P10	1/1	0/10911

**E**

Sample	No. of reads	Mean coverage
WT	463,940,226	24.7X
P10	646,472,235	34.7X

WT: Wild type



(legend on next page)

No EGFP signals were detected in any of these test groups compared with the CDS2<sup>ATG</sup>-GFP control group (Figure S3D), which indicated that no unwanted protein variants were produced in the reporter system. This simplified test excluded the possibility of splicing initiation by alternative ATG codons, but detailed targeting risk analysis needs to be verified case by case during practical use.

As revealed by a human genome scale analysis, most of the genes that are not targetable by i-STOP can be targeted by the i-Silence strategy, and vice versa (Figure 5C). A combination of i-Silence and i-STOP targeting strategies generates almost full coverage of the human genome. Moreover, an investigation of the human disease database also revealed the potential application of i-Silence in mimicking 147 human diseases caused by start codon mutations (Figure 5D). Some particularly complicated situations were not tested in this study; for example, we did not simultaneously target multiple transcript variants of the same gene with distinct start codons. However, we think that distinct variants would also be targetable with the i-Silence method through the use of multiple sgRNAs, as has been done in the i-STOP strategy.

In summary, by using a series of comprehensive analyses, the ABE-mediated i-Silence strategy is confirmed to be an efficient and safer approach for generating gene knockout and has potential clinical applications. Besides the i-Silence strategy, the ABEs-mediated nucleic acid mutation can also be used for correction of the nonsense mutation, which is termed as CRISPR-Pass in a previous study.<sup>17</sup> These two strategies showed the contrary applications of ABEs for studying the gene function, and both of them present a great power in basic research and clinical application.

## MATERIALS AND METHODS

### Mice

In this study, C57BL/6J and ICR mice were used as embryo donors and foster strains, respectively. Mice were purchased from the National Resource Center of Model Mice of Nanjing University and were maintained in a specific pathogen-free (SPF) barrier facility of Soochow University with normal diet and a 12-h dark/12-h light cycle. All mice-related experiments were reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) of the CAM-SU Genomic Resource Center (CAM-SU GRC), Soochow University.

### Plasmid Construction

EGFP reporter plasmids were constructed to contain the CDS1/CDS2 gene fused with the EGFP gene, and the sequence information for CDS1/2 is shown in Table S7. In brief, the EGFP gene without the ATG start codon was cloned into the pcDNA3.1 (Invitrogen) vector using BamHI and HindIII restriction sites. CDS1 or CDS2 genes

without stop codons were amplified and cloned into the pcDNA3.1-EGFP vector using the XhoI and BamHI restriction sites. For the sgRNA plasmid construction, oligonucleotides were annealed and ligated into BsaI-linearized pGL3-U6-sgRNA-PGK-puromycin (Addgene, 51133), pGL3-U6-sgRNA-EGFP, or pUC57-sgRNA expression vectors (Addgene, 51132). The ABEmax-mut (dCas9) plasmid was constructed by site mutation of H840A in nickase Cas9 (D10A). The sequences of the oligonucleotides used for plasmid construction are listed in Table S2.

### Cell Culture and Transfection

HEK293T cells and N2a cells were purchased from ATCC and cultured in DMEM (HyClone, SH30243.01) supplemented with 10% fetal bovine serum (Gemini, 900-108) and 1% penicillin-streptomycin (Gibco, 15140122). HEK293T and N2a cells were transfected using Lipofectamine 2000 (Life Technologies, Invitrogen) according to the manufacturer's protocol. For the validation test in the reporter system, HEK293T cells were seeded and cotransfected with 10 ng of CDS1-EGFP or CDS1-EGFP, 300 ng of sgRNA, and 600 ng of ABEmax plasmids and 2  $\mu$ L of Lipofectamine 2000. Scrambled sgRNA, dCas9, ABEmax-mut, CDS1<sup>GTG</sup>-GFP, or CDS2<sup>ACG</sup>-GFP was transfected as controls. For endogenous gene targeting, cells were seeded on poly-L-lysine (Sigma, P4707)-coated 24-well plates (JET BIOFIL); then, 300 ng of sgRNA and 600 ng of ABEmax plasmids were transfected with 2  $\mu$ L of Lipofectamine 2000. The cells were cultured at 37°C with 5% CO<sub>2</sub>, and the GFP-positive cells were analyzed by fluorescence-activated cell sorting (FACS) 48 and 96 h post-transfection.

### Flow Cytometry

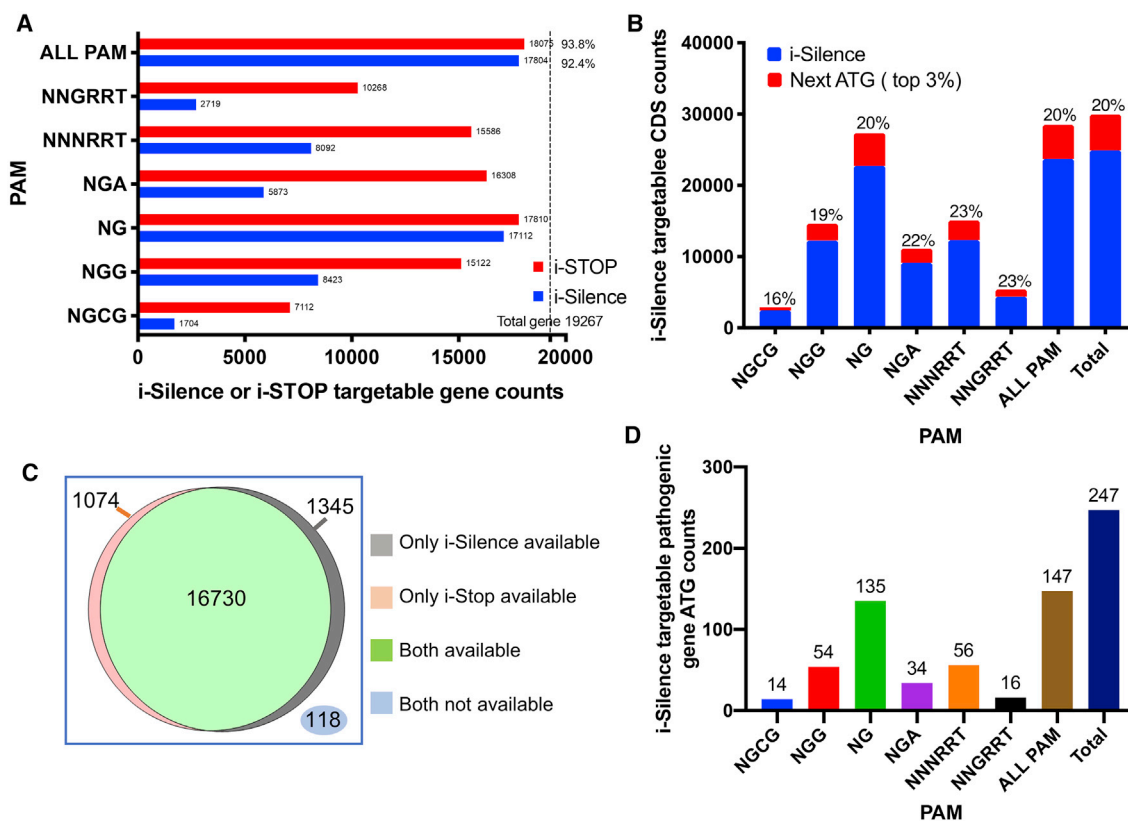
HEK293T or N2a cells were harvested and subjected to flow cytometry 48 h post-transfection. For the validation test in the reporter system, a total of 10,000 cell events were recorded and analyzed using FlowJo v10. For the EGFP reporter system, transfected cells were harvested for analysis and Sanger sequencing. For endogenous gene targeting efficiency analysis, 10,000 GFP-positive cells were harvested for genomic DNA isolation. To obtain a mutation cell colony, single-cell sorting was performed. Sorted cells were cultured in a 96-well plate and validated by sequencing.

### In Vitro Transcription

The ABEmax plasmid was linearized by treatment with BbsI enzyme (NEB), and it was transcribed using the T7 ultra kit (Ambion) according to the manufacturer's protocols. Then, ABEmax mRNA was purified using a mini kit (QIAGEN). The sgRNA was amplified from a pUC57-sgRNA expression vector with a T7 promoter (Addgene, 51132) and was transcribed using a MEGAshortscript T7 transcription kit (Ambion). Then, the sgRNAs were purified

## Figure 4. Off-Target Analysis with WGS and Deep Sequencing on PD-1 i-Silence-Edited and WT Mice

(A) A summary of SNP analysis by WGS is shown. After filtering out naturally occurring variants in the mouse SNP database, 15,078 SNPs were obtained from the P10 genome. The number of A/T conversions is shown. (B) Confirmation of the on-target base editing by analyzing the WGS results of P10 mice. (C) Summary of on-target and off-target site information. A total of 10,912 sites, including 1 on-target site and 10,911 off-target sites, were analyzed. (D) A summary of the WGS is shown. (E) A summary of off-target analysis with WGS is shown. (F) Targeted deep sequencing was performed for 18 potential off-target sites, and the mutation rates in i-Silence-edited mice are shown (n = 3); the WT mice served as controls.



**Figure 5. ABE-based i-Silence Is Capable of Genome-Scale Application**

(A) The number of i-Silence and i-STOP targetable human coding genes (NCBI reference sequence) are shown for six distinct PAM specificities. A gene that is targetable by i-Silence means that all of its transcripts can be targeted. An i-STOP/CRISPR-STOP gene is targetable if the first 20% of its coding sequence is targetable, which avoids the production of truncated protein.<sup>12</sup> (B) Percentage of i-Silence targetable genes are shown where the next ATG codon occurs in the beginning 3% of the sequence. (C) A Venn diagram shows the summary of human coding genes that are (not) targetable by the i-STOP or i-Silence strategy. (D) The number of clinical human start codon mutation cases is shown. The total and distinct PAM can be precisely mimicked with the i-Silence strategy.

with a MEGAclear kit (Ambion) and recovered by alcohol precipitation according to the manufacturer's protocol. The RNAs were aliquoted and stored at  $-80^{\circ}\text{C}$  until use.

#### Mouse Zygote Microinjection

Female C57BL/6J mice (approximately 3 weeks old) were superovulated by intraperitoneal (i.p.) injection with pregnant mare serum (PMS) on day 1 and human chorionic gonadotropin (HCG) on day 3. Immediately after HCG injection, female mice were mated to male C57BL/6J mice. Female mice were sacrificed the next day, and zygotes were collected from the oviducts (E0.5). ABEmax mRNA (50 ng/ $\mu\text{L}$ ) and sgRNA (25 ng/ $\mu\text{L}$ ) were mixed and injected into the cytoplasm of zygotes. Injected zygotes were cultured in M16 medium (Sigma) for *in vitro* analysis, or they were transferred into the oviducts of pseudopregnant ICR female mice for offspring generation.

#### Genomic DNA Extraction

The genomic DNA of HEK293T or N2a cells was isolated using QuickExtract DNA extraction solution (Lucigen) according to the

manufacturer's protocols. Embryos were lysed, and the genomic DNA was isolated and preamplified using commercial kits (N601, Vazyme). Genomic DNA from the mouse tail and other organs was extracted by a phenol-chloroform method.

#### Western Blot Analysis

HEK293T cells and homogenized mouse thymus and spleen were lysed with radioimmunoprecipitation assay (RIPA) buffer supplemented with protein inhibitor. The protein was then separated by SDS-PAGE and transferred onto a membrane. The membrane was then blotted with corresponding primary and secondary antibodies, and the images were captured with an Amersham Imager 600 (GE Healthcare). The antibodies were as follows: rabbit anti-HDAC1 (sc-7872; Santa Cruz Biotechnology), rabbit anti-SEC61B (15087-1-AP; Proteintech), rabbit anti-mouse PD-1 (18106-1-AP; Proteintech), mouse anti-GAPDH (G8795; Sigma), horseradish peroxidase (HRP)-conjugated goat anti-rabbit IgG (AS014; ABclonal), and HRP-conjugated goat anti-mouse IgG (AS003; ABclonal).



### Targeted Deep Sequencing

The on-target sites were amplified from genomic DNA using Super-Fidelity DNA polymerase (P505-d3, Vazyme) and the primers listed in Table S4. The PCR products were purified and subjected to high-throughput sequencing using an Illumina HiSeq X Ten (2 × 150) platform to obtain at least 1 million reads for each DNA sample. The adaptor pair from the pair-end reads was removed using AdapterRemoval version 2.2, and pair-end read alignments of 11 or more bases were combined into a single consensus read. All processed reads were then mapped to the target sequences using the BWA-MEM algorithm (BWA v0.7.16). For each target site, the mutation rate was calculated using bam-readcount with the following parameters:  $-q = 20$ ,  $-b = 30$ . Indels were calculated based on reads containing at least one inserted or deleted nucleotide in the protospacer. Indel frequency was calculated as the number of indel-containing reads/total mapped reads.

The potential off-target sites of i-Silence edited mice were predicted by Cas-OFFinder (<http://www.rgenome.net/cas-offinder/>).<sup>18</sup> The top 18 predicted potential off-target sites were analyzed with deep sequencing. The primers used for deep sequencing are listed in Table S5, and the potential off-target site information is provided in Table S6.

### WGS Analysis

For the WGS, mouse tail genome DNA was isolated using a phenol-chloroform method. WGS libraries were prepared using standard protocols for the Illumina HiSeq X Ten platform at the Novogene Bioinformatics Institute (Beijing, China). The resulting clean reads were aligned to the reference mouse GRCm38 assembly with BWA v0.7.16 using default parameters. After duplicated reads were removed with Sambamba v0.6.7, the sequence was realigned by using the Genome Analysis Toolkit (GATK v3.7) IndelRealigner. To identify variants, the GATK HaplotypeCaller was used with the following parameters: (1) sequencing depth  $>1/3\times$  and  $<3\times$ ; (2) variant confidence/quality by depth  $>2$ ; (3) root mean square (RMS) mapping quality (MQ)  $>40.0$ ; (4) Phred-scaled p value using Fisher's exact test to detect strand bias  $<60$ ; (5) Z score from the Wilcoxon rank sum test of Alt versus Ref read MQs (MQRankSum) greater than  $-12.5$ ; and (6) Z score from the Wilcoxon rank sum test of Alt versus Ref read position bias (ReadPosRankSum) greater than  $-8$ . After filtering out variants found in the dbSNP and in the wild-type genome, potential off-target sites were predicted by CasOT-1.0, considering up to a 2-bp mismatch in the seed region and a 5-bp mismatch in the non-seed region with NGG PAM.

### Statistical Analysis

The results were obtained from three independent experiments and are presented as the mean  $\pm$  SD. Statistical analyses and graphing were carried out using GraphPad Prism 7.0. The p values were calculated with two-tailed Student's t tests.

### Data Availability

All of the deep sequencing data were deposited in National Omics Data Encyclopedia (NODE) under the project accession no. NODE: OEP000323. All other data are available upon reasonable request.

### SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.ymthe.2019.11.022>.

### AUTHOR CONTRIBUTIONS

X.M., F.Z., and X.H. conceived of and designed the project. X.W. and Z.L. performed the majority of experiments with technical assistance from G.L., L.D., L.H., Y.M., C.L., G.Y., and M.L. S.H. and L.D. analyzed the data. X.W., F.Z. and G.L. wrote the manuscript with review comments from all authors. X.M. and X.H. edited the paper. X.M. and F.Z. supervised and managed the project.

### CONFLICTS OF INTEREST

The authors declare no competing interests.

### ACKNOWLEDGMENTS

We thank the members of the Ma lab, the Zhou lab, and the Huang lab for their scientific advice. This study was supported by grants from the National Natural Science Foundation of China (81600380, 31600958, 81572399, and 81773249), the Natural Science Foundation of Jiangsu Province (BK20160313 and BK20160317), the National Center for International Research (2017B01012), the Young Teachers Research and Development Fund from South China Normal University (18KJ06), and Key Realm R&D Program of Guangdong Province (2019B030335001).

### REFERENCES

- Haapaniemi, E., Botla, S., Persson, J., Schmierer, B., and Taipale, J. (2018). CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. *Nat. Med.* 24, 927–930.
- Ihry, R.J., Worringer, K.A., Salick, M.R., Frias, E., Ho, D., Theriault, K., Kommineni, S., Chen, J., Sondey, M., Ye, C., et al. (2018). p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. *Nat. Med.* 24, 939–946.
- Cullot, G., Boutin, J., Toutain, J., Prat, F., Pennamen, P., Rooryck, C., Teichmann, M., Rousseau, E., Lamrissi-Garcia, I., Guyonnet-Duperat, V., et al. (2019). CRISPR-Cas9 genome editing induces megabase-scale chromosomal truncations. *Nat. Commun.* 10, 1136.
- Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424.
- Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.L., and Liu, D.R. (2017). Programmable base editing of A·T to G·C in genomic DNA without DNA cleavage. *Nature* 551, 464–471.
- Kim, K., Ryu, S.M., Kim, S.T., Baek, G., Kim, D., Lim, K., Chung, E., Kim, S., and Kim, J.S. (2017). Highly efficient RNA-guided base editing in mouse embryos. *Nat. Biotechnol.* 35, 435–437.
- Lee, H.K., Willi, M., Miller, S.M., Kim, S., Liu, C., Liu, D.R., and Hennighausen, L. (2018). Targeting fidelity of adenine and cytosine base editors in mouse embryos. *Nat. Commun.* 9, 4804.
- Kim, Y.B., Komor, A.C., Levy, J.M., Packer, M.S., Zhao, K.T., and Liu, D.R. (2017). Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nat. Biotechnol.* 35, 371–376.
- Kim, D., Lim, K., Kim, S.T., Yoon, S.H., Kim, K., Ryu, S.M., and Kim, J.S. (2017). Genome-wide target specificities of CRISPR RNA-guided programmable deaminases. *Nat. Biotechnol.* 35, 475–480.
- Zhang, H., Pan, H., Zhou, C., Wei, Y., Ying, W., Li, S., Wang, G., Li, C., Ren, Y., Li, G., et al. (2018). Simultaneous zygotic inactivation of multiple genes in mouse through CRISPR/Cas9-mediated base editing. *Development* 145, dev168906.

11. Liu, Z., Lu, Z., Yang, G., Huang, S., Li, G., Feng, S., Liu, Y., Li, J., Yu, W., Zhang, Y., et al. (2018). Efficient generation of mouse models of human diseases via ABE- and BE-mediated base editing. *Nat. Commun.* 9, 2338.
12. Billon, P., Bryant, E.E., Joseph, S.A., Nambiar, T.S., Hayward, S.B., Rothstein, R., and Ciccia, A. (2017). CRISPR-mediated base editing enables efficient disruption of eukaryotic genes through induction of STOP codons. *Mol. Cell* 67, 1068–1079.e4.
13. Zuo, E., Sun, Y., Wei, W., Yuan, T., Ying, W., Sun, H., Yuan, L., Steinmetz, L.M., Li, Y., and Yang, H. (2019). Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science* 364, 289–292.
14. Jin, S., Zong, Y., Gao, Q., Zhu, Z., Wang, Y., Qin, P., Liang, C., Wang, D., Qiu, J.L., Zhang, F., and Gao, C. (2019). Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. *Science* 364, 292–295.
15. Pagnamenta, A.T., Murakami, Y., Anzilotti, C., Titheradge, H., Oates, A.J., Morton, J., Kinoshita, T., Kini, U., and Taylor, J.C.; DDD Study (2018). A homozygous variant disrupting the PIGH start-codon is associated with developmental delay, epilepsy, and microcephaly. *Hum. Mutat.* 39, 822–826.
16. Cremonesi, L., Cozzi, A., Girelli, D., Ferrari, F., Fermo, I., Foglieni, B., Levi, S., Bozzini, C., Camparini, M., Ferrari, M., and Arosio, P. (2004). Case report: a subject with a mutation in the ATG start codon of L-ferritin has no haematological or neurological symptoms. *J. Med. Genet.* 41, e81.
17. Lee, C., Hyun Jo, D., Hwang, G.H., Yu, J., Kim, J.H., Park, S.E., Kim, J.S., Kim, J.H., and Bae, S. (2019). CRISPR-pass: gene rescue of nonsense mutations using adenine base editors. *Mol. Ther.* 27, 1364–1371.
18. Bae, S., Park, J., and Kim, J.-S. (2014). Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* 30, 1473–1475.