

# Engineering an adenine base editor in human embryonic stem cells with minimal DNA and RNA off-target activities

Zhenwu Zhang,<sup>1,5</sup> Wanyu Tao,<sup>1,2,5</sup> Shisheng Huang,<sup>1,2,3,5</sup> Wenjun Sun,<sup>1,2</sup> Yue Wang,<sup>1,2</sup> Wen Jiang,<sup>4</sup> Xingxu Huang,<sup>1,3</sup> and Chao-Po Lin<sup>1</sup>

<sup>1</sup>School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China; <sup>2</sup>University of Chinese Academy of Sciences, Beijing 100049, China; <sup>3</sup>Zhejiang Lab, Hangzhou, Zhejiang 311121, China; <sup>4</sup>Department of Hematology, Southwest Hospital, Third Military Medical University (Army Medical University), Chongqing 400038, China

**Genome editing in pluripotent stem cells (PSCs) using CRISPR technology holds great promise for therapeutic applications. Yet, it has been reported that Cas9-mediated cleavage could cause large deletions or rearrangements of DNA, and the selection of edited PSCs could acquire p53 mutations. Adenine base editors (ABEs) do not introduce DNA double-strand breaks and thus have been proposed as alternatives to circumvent those problems, but their off-target effects still limit their applications. Here, we tested different combinations of off-target reduction methods to further diminish off-target effects of ABEs without compromising their on-target editing efficiencies. We subsequently chose the best editor, CE-8e-dV, which contains V106W substitution, R153 deletion, and Cas-embedding strategy, to establish a single-cell-derived human embryonic stem cell (hESC) line expressing tetracycline-inducible CE-8e-dV. By performing RNA and whole-genome sequencing, we demonstrated that the expression of CE-8e-dV did not produce nearly any DNA or RNA off-target effects in hESCs. Our results provide stringent proof of the safety of ABEs in PSCs and suggest that CE-8e-dV could be suitable for related therapeutic strategies, such as generation of engineered stem cells *in vitro* and gene therapy *in vivo*.**

## INTRODUCTION

Pluripotent stem cells (PSCs) are able to self-renew indefinitely and differentiate into nearly all cell types of three germ layers, the property holding great promises for regenerative medicine. In the past 10 years, the efficiency and the ease of CRISPR-based gene-editing technologies have been applied in PSCs for basic researches, such as knockouts/knockins, building disease models, and correcting genetic mutations.<sup>1</sup> Recent studies suggest that the efficiency of Cas9-nuclease-mediated gene editing in human induced PSCs (hiPSCs) is 3- to 8-fold lower than that in transformed cells (293T and K562).<sup>2</sup> The low efficiency is, at least in part, due to the activation of p53 in response to Cas9-mediated DNA cleavage, which leads to large deletions or rearrangements in DNA.<sup>3,4</sup> Meanwhile, the selection of successfully edited clones could favor the accumulation of p53 mutations,

hampering further utilization of this powerful technology in human embryonic stem cells (hESCs).<sup>4</sup>

Adenine base editors (ABEs), as one of the CRISPR-based gene-editing tools, can achieve A-to-G base substitutions in the genome. Notably, ABEs greatly avoid the generation of DNA double-strand breaks that could otherwise lead to genomic aberrations or p53 activation,<sup>5</sup> making ABEs suitable candidates for correcting disease mutations in human stem cells.<sup>6</sup> For example, Lin et al. illustrated a new therapeutic strategy for treating spinal muscular atrophy (SMA) by ABE-mediated correction of splicing in patient-derived iPSCs.<sup>7</sup> Osborn et al. applied ABE to correct COL7A1 mutations in recessive dystrophic epidermolysis bullosa (RDEB) patient-derived iPSCs.<sup>8</sup> Huang et al. showed that correction of C625T mutation in the RS1 gene using ABE restored the phenotype of retinal organoids generated from patients with X-linked juvenile retinoschisis (XLRS).<sup>9</sup> These applications of ABEs in human pluripotent and adult stem cells provide important indications for therapeutic purposes. Nevertheless, the adenosine deaminase of ABEs, such as TadA-7.10, is evolved from an *Escherichia coli* tRNA adenosine deaminase (TadA),<sup>6</sup> which possesses RNA off-target effects in a single guide RNA (sgRNA)-independent manner, as demonstrated by Rees et al. and Grunewald et al.<sup>10,11</sup> For hESCs, there is little characterization or information regarding the off-target effects of ABEs. Therefore, the safety of ABEs in hESCs still remains to be carefully solidified for future clinical applications.

Recently, the Liu group constructed ABE8e derived by phage-assisted evolution, exhibiting a large improvement in A-to-G

Received 20 December 2021; accepted 22 July 2022;  
<https://doi.org/10.1016/j.omtn.2022.07.026>.

<sup>5</sup>These authors contributed equally

**Correspondence:** Xingxu Huang, School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China.

**E-mail:** [huangxxx@shanghaitech.edu.cn](mailto:huangxxx@shanghaitech.edu.cn)

**Correspondence:** Chao-Po Lin, School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China.

**E-mail:** [linzhh@shanghaitech.edu.cn](mailto:linzhh@shanghaitech.edu.cn)



base-editing efficiency with up to 9.4-fold increase compared with ABE7.10.<sup>12</sup> However, corresponding to its high DNA-editing efficiency, ABE8e showed higher off-target editing effects on DNA and RNA in both sgRNA-dependent and -independent manners.<sup>12</sup> Several methods were applied to further reduce off-target mutations induced by adenosine deaminases, including point mutation of ABE8e,<sup>12</sup> deletion of the key residue in ABE8e<sup>13</sup> and a Cas-embedding strategy.<sup>14</sup> In this study, we combined these methods to create a new editor: CE-8e-dV. We observed that this combination further reduced RNA off-targeting effects without affecting its on-target editing efficiency. Since hESCs in culture could contain sporadic mutations, we constructed a single-cell-derived hESC colony with inducible CE-8e-dV to evaluate DNA and RNA off-target effects by whole-genome sequencing and RNA sequencing in an isogenic background. We demonstrated that long-term expression of CE-8e-dV produces barely any DNA or RNA off-targets in hESCs, providing the proof of the safety of CE-8e-dV in PSCs and suggesting its further applications for related therapeutic strategies *in vitro* and *in vivo*.

## RESULTS

### Combination of off-target reduction methods did not alter on-target editing efficiency

Previously, we showed that both engineered ABEs (eABE) and Cas-embedding ABEs (CE-ABEs) can reduce off-target editing without affecting on-target editing efficiencies.<sup>13,14</sup> eABEs delete 153 arginine (R153), which locates in the activity area for RNA deamination, from adenosine deaminase to reduce RNA off-target effects, while CE-ABEs reduce off-targeting editing by inserting the adenosine deaminase into the middle of Cas9 nickase (nCas9). It was also reported that V106W mutation in TadA8e reduces both DNA and RNA off-target effects.<sup>12</sup> We assume that the performance of ABEs could be further improved by combining these off-target reduction methods in TadA8e. To this end, we generated SpRY<sup>15</sup> (a SpCas9 variant with high PAM compatibility)-based ABE constructs with individual or combinations of V106W substitution, R153 deletion, and CE strategy (Figures 1A and S1).

To investigate whether those modifications alter on-target editing activities of ABEs, we compared on-target editing efficiencies of these eight constructs across 48 endogenous sites. Each sgRNA, together with eight individual constructs, were co-transfected into HEK293T cells. As demonstrated by the results, there were no significant changes of on-target editing efficiencies in those eight constructs ( $48\% \pm 5.7\%$ ) (Figure S2). We classified these editors using the same method to reduce off-target efficiency into V106W, delta 153, and CE categories and found that superimposing one variant on another did not have much impact on editing efficiencies (Figure 1B). Interestingly, these base editors with superimposed V106W substitution seem to exhibit slightly increased editing efficiencies (50.12% versus 44.05%) (Figure S3). In addition, the active editing windows of those four combinations were still within positions 3–10, similar to the known active editing window of ABE8e-RY, suggesting that combinations of those

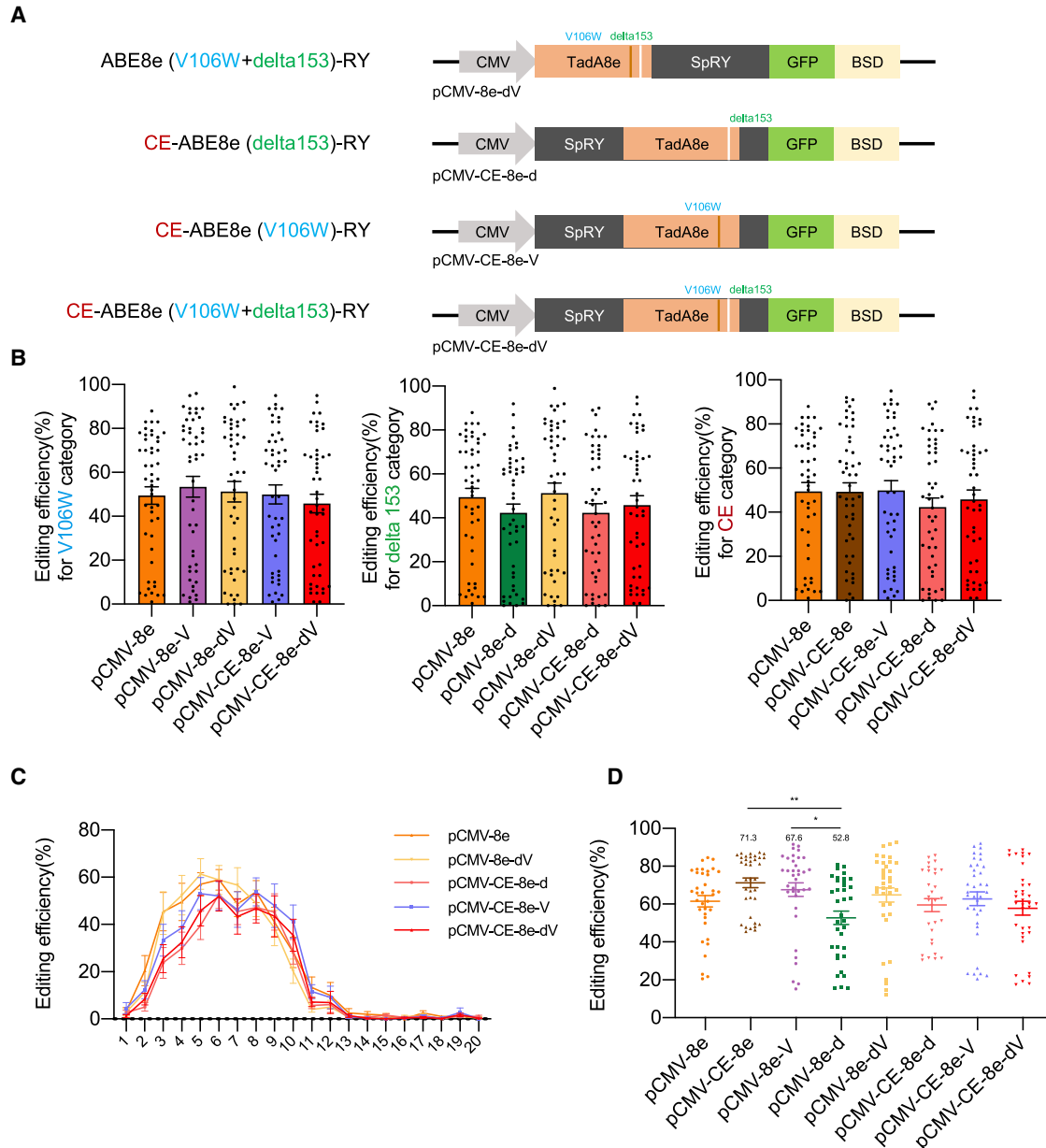
modifications on ABEs did not alter the widths of editing windows (Figure 1C). Interestingly, the editors with CE strategy showed improved editing efficiency at positions close to PAMs (28.2% for pCMV-8e versus 41.6% for pCMV-CE-8e for the base at position 10). Since SpRY exhibits relieved PAM restriction,<sup>15</sup> we deep sequenced amplicons of those 12 target sites with different PAMs (NAN, NGN, NCN, and NTN) and found that SpRY still maintained its expanded editing scope with combinations of these methods (Figures 1D and S4).

### CE-8e-dV showed the lowest off-target editing effect on RNA in HEK293T cells

We next investigated whether combinations of off-target reduction methods could achieve lower RNA off-target effects. To detect transcriptome-wide RNA off-target effects, we performed RNA sequencing on HEK293T cells expressing the eight constructs or GFP (negative control) with an sgRNA targeting NCA SITE3 that shows high DNA on-target editing efficiency (Table S1; Figure 2A). ABE constructs with individual or combinations of these methods can all reduce RNA off-target effects (Figures 2B and S5). Among these, combinations of two off-target reduction methods further reduced off-target effects (reduced by 51.9%–64.1% compared with ABE8e-RY), especially the combination of V106W substitution and CE (Figure 2B). The CE-8e-dV combination, which includes V106W substitution, R153 deletion, and CE strategy, exhibited the lowest off-target effect, nearly 3.3 times fewer RNA edits than ABE8e-RY (5,971.5 versus 19,863.5). Nevertheless, we didn't detect linear additive effects on the reduction of off targets in combinations of these modification categories (V106W, delta 153, and CE) (Figure 2B). One possible explanation could be the partially overlapped mechanisms of these methods. We further analyzed the off-target effects (A-to-I RNA editing) of those constructs. CE-8e-dV again showed the lowest A-to-I RNA edits among the other editors (Figure 2C). We also found the CE-8e-dV has lower levels of insertions or deletions (indels) compared with other editors (Figure S6). Since there are rare well-characterized off-target sites for ABE8e, we examined the off-target editing activities of ten off-target sites using RNA sequencing (RNA-seq) datasets used in this and another study.<sup>16</sup> We found that combinations of those modifications on ABEs significantly reduced off-target editing activities compared with ABE8e, especially for OT1, OT5, OT6, OT8, and OT10 sites (Figure S7). Taken together, our results indicate that the combination of different off-target reduction methods indeed further reduces off-target effects, among which CE-8e-dV exhibited the lowest off-target effect.

### Construction of the hESC colony with inducible CE-8e-dV expression

We next want to examine the editing efficiency and safety of CE-8e-dV in cell types more potentially relevant to therapeutics, such as PSCs. To this end, we first tested base-editing efficiencies on several sites by transiently transfecting sgRNAs and ABE constructs with different combinations of off-target reduction methods in SHhES8 hESC line,<sup>17</sup> and no significant difference was observed between the

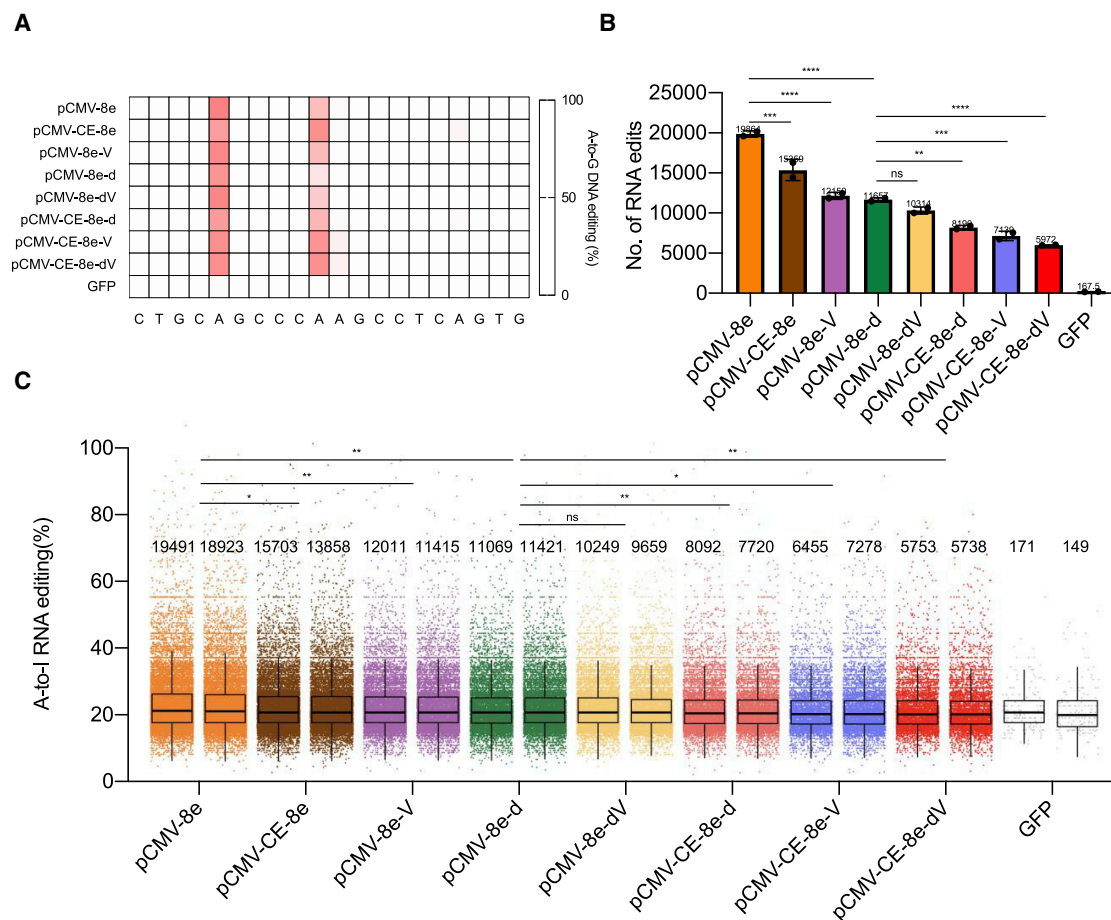


**Figure 1. Combination of off-target reduction methods did not alter on-target editing efficiencies**

(A) Schematic diagrams for four combinations' editor of two or three pairs of off-target reduction methods (V106W substitution [V106W], R153 deletion [del 153], and Cas-embedding strategy [CE]). (B) Mean A-to-G editing efficiency plots classified by the method about off-target reduction. The data are presented for the edited base with the highest editing efficiency among the 48 different target sites. No significant differences were observed between any of two pairs as calculated by one-way ANOVA. (C) Aggregate distribution of A-to-G edits made across the editing window. The data are presented for each edited base in the editing window at 48 endogenous sites. (D) Comparison of the editing efficiency of the eight editors. The horizontal lines represent the mean editing efficiencies at 12 endogenous sites for each editor. Each dot represents the editing efficiency of each edited base with the highest editing efficiency ( $n = 3$  independent replicates). No significant differences were observed between any of two pairs, except for the comparisons marked by one asterisk ( $*p < 0.05$ ) and two asterisk ( $**p < 0.01$ ), Kruskal-Wallis test.

different editors (Figure 3A). The editing outcomes of CE-8e-dV in SHhES8 cells were comparable to that in HEK293T cells (Figure 3B) (42.5% [SHhES8] versus 54.6% [HEK293T] for CE-8e-dV editor), consistent with previous reports.<sup>18</sup>

It is known that hESCs in culture exhibit certain heterogeneity and could acquire sporadic mutations.<sup>19</sup> To further analyze the off-target situations of ABEs in a more controlled, isogenic background, we employed the piggyBac system to establish some single-cell-derived



**Figure 2. CE-8e-dV showed the lowest off-target editing effect on RNA in HEK293T cells**

(A) Heatmaps showing the on-target editing frequencies in NCA-SITE3 site. (B) Total number of RNA off-target edits detected in RNA sequencing experiments. The data shown are from two independent replicates. ns, not significant,  $p > 0.05$ ,  $**p < 0.01$ ,  $***p < 0.001$ ,  $****p < 0.0001$ , one-way ANOVA. (C) Jitter plots showing the efficiencies of A-to-I RNA off-target editing. The data shown are from two independent replicates. ns, not significant,  $p > 0.05$ ,  $**p < 0.01$ , one-way ANOVA. The numbers of A-to-I RNA edits are indicated upon the plots.

stable SHhES8 clones, with the CE-8e-dV base editor being induced by doxycycline (Figure 3C). We picked three clones and assessed the mRNA and protein levels of the ABEs by western blot and qRT-PCR (Figures S8A and S8B). No obvious leakage of expression was detected without doxycycline treatment in all three clones, suggesting the robustness of this inducible system in hESCs. Since all three clones exhibit similar editing efficiencies (Figure S8C), we picked clone 4, which expressed moderate CE-8e-dV (Figure S8) and high nuclear pluripotency markers OCT4 and SOX2 (Figures 3D and S9), for the following investigations.

We next want to examine whether different overexpression methods (transient transfection and induction) result in difference editing efficiencies in PSCs. We examined editing efficiencies of ABEs at 12 sites in the same CE-8e-dV SHhES8 clone and found that the editing efficiency of inducible CE-8e-dV (+doxycycline [dox]) is similar to that of transfected CE-8e-dV (–dox with transfection) in the same

hESC clone (Figure 3E). All positions were covered at least twice. In fact, we found that the editing efficiency of the inducible system is as robust as the transient transfection system (Figure 3E), with similar active editing windows (Figure S10). Altogether, this stable SHhES8 cell line showed effective base editing efficiency and exhibited negligible background expression of ABE, which is suitable for further studies.

#### Assessing RNA and DNA off-target effects of ectopically expressed CE-8e-dV in hESCs

After proving the editing ability of CE-8e-dV in hESCs, we aimed to analyze its off-target effects on RNA. We first wanted to examine whether the overexpression methods influence off-targeting activities of CE-8e-dV. We transfected CE-8e-dV plasmid (–dox with transfection) or treated dox (+dox) to the same CE-8e-dV stable SHhES8 clone, together with an sgRNA-targeting NCA SITE3 that shows high DNA on-target editing efficiency (Figure S11A). We found

that RNA off-target effects are similar compared with the transient transfection system (Figures S11B and 11C).

We next wanted to examine transcriptome-wide RNA off-target effects in PSCs. Individual single-cell-derived clones from CE-8e-dV hESCs were further expanded and treated with or without dox (2  $\mu\text{g}/\text{mL}$ ) for 2 days to evaluate RNA mutation landscapes (Figure 4A). No obvious sgRNA-dependent off-targeting activities were detected (Figure S12). An unsupervised principal-component analysis (PCA) revealed tight clustering across replicates (Figure S13A). Importantly, very few differentially expressed genes were uncovered upon ABE expression in hESCs (Figure S13B). No significant differences in RNA mutations were observed between the absence or presence of dox (210 versus 203), and the A-to-G mutation type constituted of the largest proportion (37.6% for untreated hESCs and 41.6% for dox-treated hESCs) (Figure 4B). The accumulation of these mutations is likely the background under cultured conditions since it has been reported that individual human adult and PSCs accumulate  $3.5 \pm 0.5$  base substitutions per population doubling.<sup>19</sup> We further analyzed the distribution of mutations in CE-8e-dV stable SHhES8 cell line untreated or treated with dox. We found that the mutations within protein coding regions accounted for a large proportion, whereas the mutations within oncogene regions or tumor suppressor regions accounted only for a small percentage (Figure 4C).

To analyze its off-target effects on DNA, CE-8e-dV stable SHhES8 were grown for 21 days in the presence or absence of dox to induce CE-8e-dV base-editor expression. Individual single-cell-derived clones were isolated after long-term expression of CE-8d-dV, further scaled up, and harvested for whole-genome sequencing. The numbers of sequence mutations were analyzed by comparing the sequencing results of all newly derived hESC clones with parent hESC clones. The sequence variations, which were detected in the three clones derived from the non-induced group, were almost indistinguishable from those treated with dox for 21 days (668 versus 638) (Figure 4B). These mutations were mainly concentrated in intergenic regions and intron regions (Figure 4D). In summary, our study proved that CE-8e-dV produces minimal DNA or RNA off-target effects in hESCs.

## DISCUSSION

PSCs are one of the key players in regenerative medicine, as exemplified by the application of PSC-derived chimeric antigen receptor (CAR)-T cells or natural killer (NK) cells in cancer immunotherapy. However, concerns were raised, as gene-editing based on Cas9-mediated DNA cleavage in hESCs produces deletions/rearrangements of DNA.<sup>3,5</sup> As a cleavage-independent approach for gene editing, the efficiency and the extent of off-target effects on DNA and RNA of base editors in PSCs remain to be improved and assessed. In this study, we demonstrated that combinations of three off-target reduction methods in ABEs can further lower their off-target effects without influencing their on-target editing efficiency in HEK293T cells and hESCs. Interestingly, we obtained similar editing efficiencies in both cell types in transformed and non-transformed cells. Since we did not find the enrichment of genes in the p53 signaling pathway

upon ABE expression, the high editing efficiency of ABEs in hESCs support the conclusion that ABEs did not elicit DNA damage in hESCs, which are susceptible to p53-induced cell death.

To thoroughly examine potential off-target effects caused by our editor (CE-8e-dV, with V106W substitution, R153 deletion, and CE strategy), RNA-seq and whole-genome sequencing were conducted in single-cell-derived SHhES8 colonies. The results unambiguously confirmed that our editor induces minimal DNA and RNA off-target SNVs in hESCs. Interestingly, it has been reported that the editing activity with NCN PAM is lower in SpRY compared with other PAMs.<sup>15</sup> In our results, the editing activity of TadA8e-fused SpRY with NCN PAM is at the similar level compared with NAN and NGN PAMs, indicating that TadA8e may influence the PAM preference of SpRY.

The induction system employed in this study exhibit similar robustness in editing efficiency while maintaining the lowest off-targeting activities compared with transfection in hESCs. This system has several advantages: (1) the single-cell-derived inducible system is more suitable for evaluating the off-targeting effects on DNA and RNA for other base editors because of the isogenic background and homogeneous expression. (2) The engineered cells are genetically stable, making their editing efficiency less affected by delivery methods (e.g., viral vectors or transfecting reagents) or microenvironments *in vivo*. (3) The tetO promoter can be further modified and/or coupled with conditionally expressed rtTA for a spatiotemporal controlled ABE expression, which is instrumental for functional genomics, disease modeling, drug discovery, and regenerative medicine.

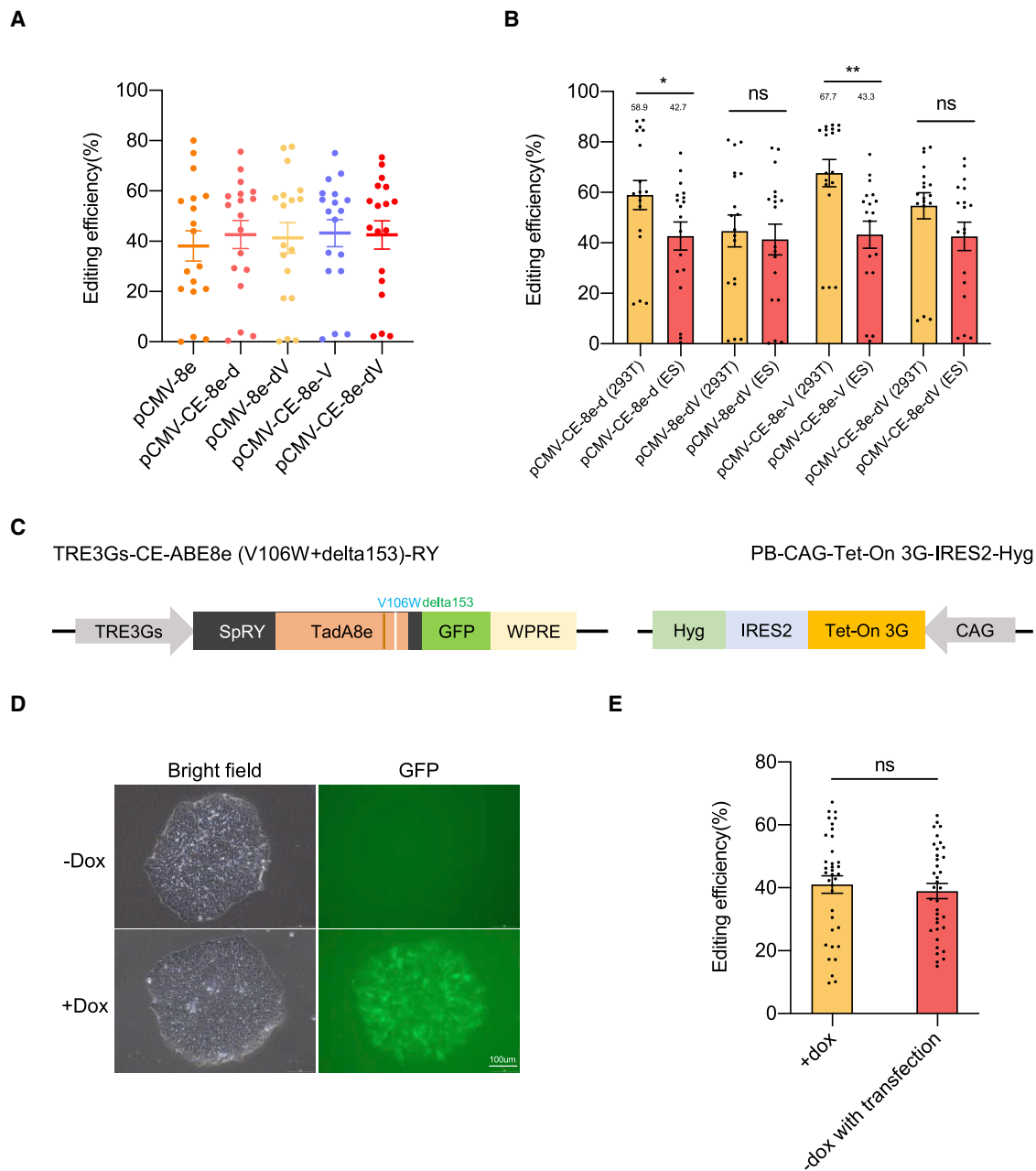
## MATERIALS AND METHODS

### Cell culture

HEK293T cells were purchased from the American Type Culture Collection (ATCC). The cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum and 1% penicillin-streptomycin. We maintained the cell line at 37°C in a 5% CO<sub>2</sub> cell culture incubator. hESC line SHhES8 was routinely cultured in mTeSR1 (STEMCELL Technologies) on Matrigel-coated (STEMCELL Technologies) plates. Cells were passaged every 4 days using Gentle Cell Dissociation Reagent (STEMCELL Technologies).

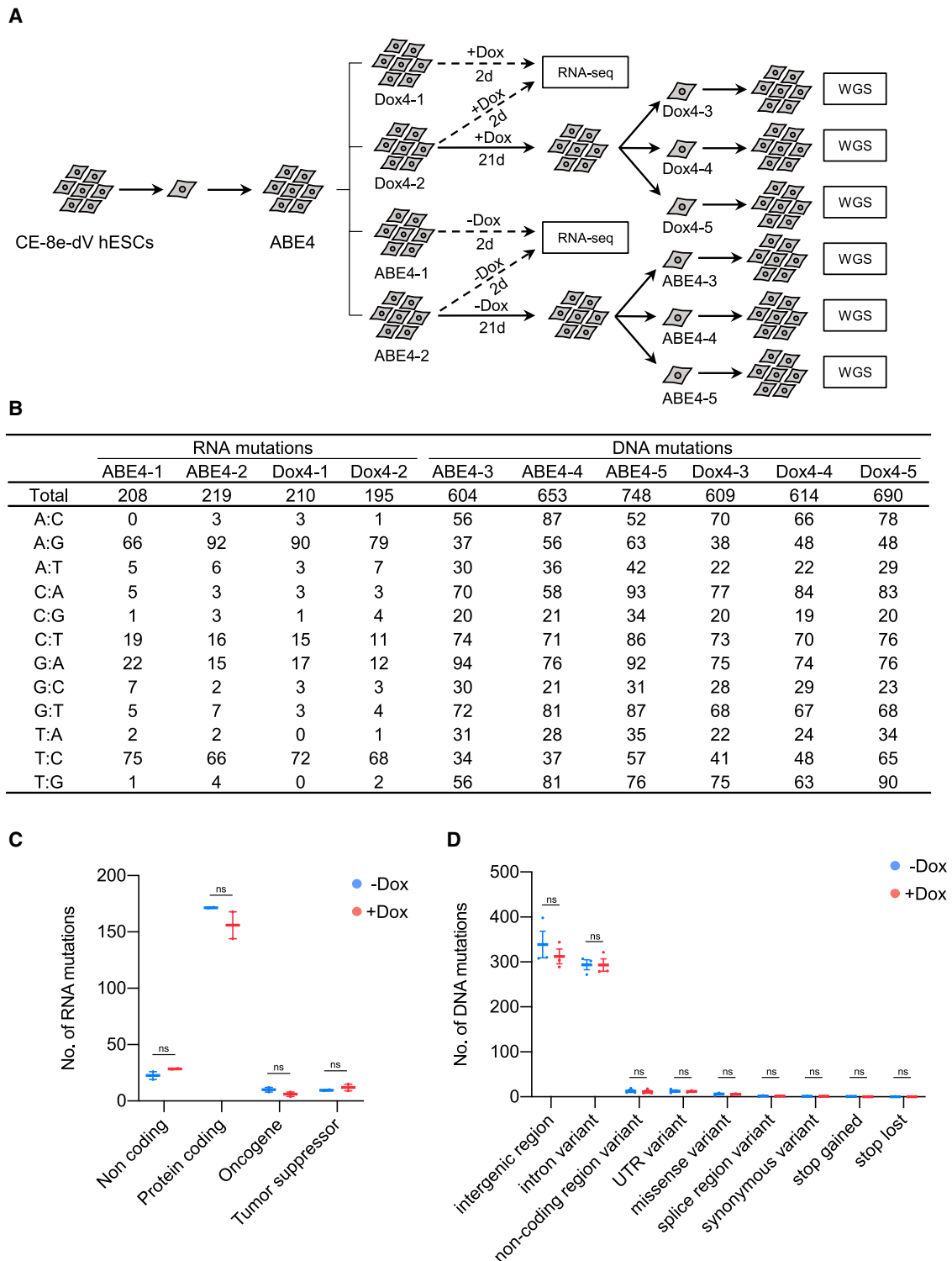
### Plasmid construction

To construct base-editor expression plasmids, TadA8e were synthesized by GenScript and cloned into the pCMV-SpRY(D10A) backbone containing C-terminal-fused EGFP and BSD. Gene recombination was performed using ClonExpress MultiS Cloning Kit (Vazyme) to generate pCMV-CE-8e, pCMV-8e-V, pCMV-8e-d, pCMV-8e-dV, pCMV-CE-8e-d, pCMV-CE-8e-V, and pCMV-CE-8e-dV. To construct sgRNA expression vectors, synthesized oligos were annealed and ligated into a BsaI-digested sgRNA expression vector (pGL3-U6-sgRNA-mCherry plasmid). Sequences of sgRNA constructs used in this work are listed in Table S1. For inducible ABE long expression in hESCs, the TRE3Gs fragment was synthesized by GenScript and cloned into a piggyBac plasmid containing the 5' and 3' PiggyBac



**Figure 3. Construction of the hESC colony with inducible CE-8e-dV expression**

(A) Relative A-to-G base-editing efficiencies of the four editors in SHhES8 cells at three sites that already tested in HEK293T cells ( $n = 3$  independent replicates). No significant differences were observed between any of two pairs as determined by one-way ANOVA. (B) Comparison of the editing efficiency of the four editors at three sites by deep sequencing in HEK293T cells and SHhES8 cells. Each dot represents the editing efficiency of each edited base ( $n = 3$  independent replicates). The error bars represent the standard error of the mean (SEM) values. ns, not significant,  $p > 0.05$ , Student's  $t$  test. (C) Schematic diagram for constructions of TRE3Gs-CE-ABE8e (V106W + delta 153)-RY and PB-CAG-Tet-On 3G-IRES2-Hyg, which are inserted into the SHhES8 cell line. (D) Representative images of SHhES8 cells with long-term stable CE-8e-dV expression in the absence or presence of doxycycline (dox). Scale bars represent 100  $\mu\text{m}$ . (E) Comparison of the A-to-G base-editing efficiencies between the induction system (+dox) and transfection system (–dox with transfection) at 12 sites in the CE-8e-dV stable SHhES8 cell line. Each dot represents the highest editing efficiency ( $n = 3$  independent replicates).



**Figure 4. Assessing RNA and DNA off-target effects of ectopic expressed CE-8e-dV in hESCs**

(A) Diagrammatic representation of experimental design for RNA sequencing and whole-genome sequencing analysis. For RNA sequencing, CE-8e-dV stable SHhES8 clones were treated with or without dox (2  $\mu$ g/mL) for 2 days and harvested. For whole-genome sequencing, the clones were grown for 21 days in the presence or absence of dox. Individual single-cell-derived clones were isolated after long-term expression of CE-8d-dV, further scaled up, and harvested. (B) Table about each type of mutation in

(legend continued on next page)

homology arms. CE-8e-dV was constructed into piggyBac-TRE3Gs vector using ClonExpress MultiS Cloning Kit (Vazyme).

### Inducible base-editor expression in hESCs

SHhES8 cells were dissociated with Accutase (Thermo Fisher Scientific) and plated 50,000–75,000 cells/well in the precoated 24-well plate supplement with 10  $\mu$ M Y-27632 (Tocris) before transfection. The PB-TRE3Gs-CE-ABE8e (V106W + delta 153)-RY (500 ng), PB-CAG-Tet-On 3G-IRES2-Hyg (500 ng), and piggyBac transposase plasmid (300 ng) were co-transfected using Lipofectamine Stem reagent (Thermo Fisher Scientific) according to the manufacturer's protocol. Two days later, 5  $\mu$ g/mL blasticidin and 20  $\mu$ g/mL hygromycin were added to the medium for 1 week to select stably transfected cells. Single-cell suspension was then plated into Matrigel-coated 96-well plates with 0.5 cell/well to obtain subclones of the stably transfected cells. Cells from one subclone were split into two culture conditions. To one of the subculture conditions, dox (2  $\mu$ g/mL) was added. Two days later, cells were harvested for RNA isolation using TRIzol reagent (Invitrogen) and RNA-seq. Remaining cells were subcultured for 19 days either with or without daily dox addition. During this period of expansion, cells from the two conditions were always passaged at the same time at each passage. After 21 days, single-cell-derived clones were again expanded, without dox, from each condition for about 14 more days before genomic DNA isolation and sequencing.

### Analysis of on-target editing

To evaluate editing efficiency, HEK293T cells were seeded into 24-well plates 1 day before transfection. The base-editor expression plasmids (1,000 ng) and corresponding sgRNA plasmids (500 ng) were co-transfected using EZ *trans* (Shanghai Life iLab Bio Technology) as the manufacturer's protocol recommended. For SHhES8 and H1 cells, cells were dissociated with Accutase (Thermo) and plated 50,000–75,000 cells/well in the precoated 24-well plate with 10  $\mu$ M Y-27632 (Tocris) before transfection. The base-editor expression plasmids (1,000 ng) and corresponding sgRNA plasmids (500 ng) were co-transfected using Lipofectamine Stem reagent (Thermo) according to the manufacturer's protocol. For CE-8e-dV stable SHhES8 cells, sgRNA plasmids (500 ng) were transfected using Lipofectamine Stem reagent (Thermo) according to the manufacturer's protocol. Seventy-two hours after transfection,  $\sim$ 10,000 cells with dual fluorescence signals (GFP and mCherry) were collected by fluorescence-activated cell sorting (FACS) to improve efficiency, and these cells were harvested for genomic DNA extraction using QuickExtract DNA Extraction Solution (Lucigen) according to the manufacturer's protocols. The genomic regions encompassing the target sites were amplified from the genomic DNA with Phanta Max Super-Fidelity DNA polymerase (Vazyme, P505-03). The primers used are listed in [Table S2](#). The PCR products were analyzed by Sanger sequencing

or high-throughput sequencing as indicated. For Sanger sequencing, the chromatograms were quantified using EditR. For high-throughput sequencing, the PCR products were sequenced on an Illumina HiSeq X Ten (2  $\times$  150 PE) at the Novogene Bioinformatics Institute (Beijing, China). The sequencing data were analyzed using CRISPResso2.

### Analysis of RNA off-target editing

HEK293T cells were seeded into 6 cm dishes and transfected with 4  $\mu$ g of each editor or GFP (control) plasmids and 2  $\mu$ g of sgRNA expression vector using Lipofectamine 2000 at  $\sim$ 70% confluency. Two days after transfection, the top 30% of the GFP-signal-positive cells were harvested by FACS. RNA was immediately extracted using TRIzol reagent (Invitrogen) according to the manufacturer's instructions. The RNA samples were subjected to deep sequencing ( $\sim$ 20 million reads per sample) on an Illumina HiSeq X Ten platform (2  $\times$  150 PE) at the Novogene Bioinformatics Institute (Beijing, China). The clean data were first mapped to the human reference genome (v.hg38) with annotations from GENCODE version v.30 by STAR software (v.2.5.1). After removing duplicates, GATK HaplotypeCaller (v.4.1.2) was used to identify and filter the edits. All edits were verified, and the efficiency was calculated using the bam-readcount program with the parameters -q 20 -b 30. Importantly, for the reference allele in the wild-type sample, the depth of a given edit had to be at least 10 $\times$ , and all edits had to be present in at least 99% of reads.

### Whole-genome sequencing

DNA extracted from harvested cells was sequenced using Illumina NovaSeq (PE150) at the Annoroad Gene Technology, Beijing, China. All cleaned reads were mapped to the human reference genome (GRCh38/hg38) using BWA v.0.7.17 with default parameters. Sequence reads were removed for duplicates using Sambamba v.0.6.7. Variants were identified by GATK (v.4.1.8.1) HaplotypeCaller, and the following criteria were applied to all variants: (1) sequencing depth (for each individual) between one-third to 3-fold of corresponding mean depth; (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)  $>$  -12.5; and (6) Z score from the Wilcoxon rank-sum test of Alt versus Ref read position bias (ReadPosRankSum)  $>$  -8. These variants were required to have at least 3 reads supporting the variant.

### Statistics and reproducibility

The data shown in this research were statistically analyzed by unpaired two-tailed Student's t test using GraphPad Software

---

uninduced (ABE) and induced (dox) clones. The RNA mutations in each clone (ABE4-1, ABE4-2, Dox4-1, Dox4-2) were identified by RNA sequencing, while the DNA mutations in each clone (ABE4-3, ABE4-4, ABE4-5, Dox4-3, Dox4-4, Dox4-5) were identified by whole-genome sequencing. (C) The situation of mutation distribution in RNA sequencing experiments. Blue/red represents the number of RNA off-target edits in uninduced/induced CE-8e-dV stable SHhES8 cell line. ns, not significant,  $p > 0.05$ , Student's t test. (D) The situation of mutation distribution in whole-genome sequencing experiments. Blue/red represents the number of DNA off-target edits in uninduced/induced CE-8e-dV stable SHhES8 cell line. ns, not significant,  $p > 0.05$ , Student's t test.



(GraphPad Prism 8). A *p* value smaller than 0.05 was considered to indicate statistical significance. The error bars represent the standard error of the mean (SEM) values.

#### DATA AVAILABILITY

The high-throughput sequencing datasets are available through NCBI BioProject: PRJNA790034 (<http://www.ncbi.nlm.nih.gov/bioproject>). All other data are available upon reasonable request.

#### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.omtn.2022.07.026>.

#### ACKNOWLEDGMENTS

We thank Dr. Ying Jin for the SHhES8 cell line and Dr. Hanzhi Zhao for the technical assistance. We thank members of the Huang lab and the Lin lab for helpful discussions. We also thank the Molecular and Cell Biology Core Facility (MCBCF) at the School of Life Science and Technology, ShanghaiTech University, for providing technical support. This research was supported by National Key R&D Program of China (no. 2020YFA0710800), National Natural Science Foundation of China (nos. 31901049 and 31871487), Shanghai Science and Technology Commission (no. 19JC1414200), the ShanghaiTech University start-up fund, and the Leading Talents of Guangdong Province Program (2016LJ06S386).

#### AUTHOR CONTRIBUTIONS

X.H. and C.-P.L. conceived, designed, and supervised the project. Z.Z., W.T., and S.H. performed most experiments with the help of other authors. S.H. analyzed and interpreted the data. Z.Z., W.T., and X.H. wrote the paper with inputs from all authors.

#### DECLARATION OF INTERESTS

The authors declare that they have no competing interests.

#### REFERENCES

- Hendriks, D., Clevers, H., and Artegiani, B. (2020). CRISPR-cas tools and their application in genetic engineering of human stem cells and organoids. *Cell Stem Cell* 27, 705–731.
- Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. *Science* 339, 823–826.
- 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.
- Newby, G.A., Yen, J.S., Woodard, K.J., Mayuranathan, T., Lazzarotto, C.R., Li, Y., Sheppard-Tillman, H., Porter, S.N., Yao, Y., Mayberry, K., et al. (2021). Base editing of haematopoietic stem cells rescues sickle cell disease in mice. *Nature* 595, 295–302.
- 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.
- Lin, X., Chen, H., Lu, Y.Q., Hong, S., Hu, X., Gao, Y., Lai, L.L., Li, J.J., Wang, Z., Ying, W., et al. (2020). Base editing-mediated splicing correction therapy for spinal muscular atrophy. *Cell Res.* 30, 548–550.
- Osborn, M.J., Newby, G.A., McElroy, A.N., Knipping, F., Nielsen, S.C., Riddle, M.J., Xia, L., Chen, W., Eide, C.R., Webber, B.R., et al. (2020). Base editor correction of COL7A1 in recessive dystrophic epidermolysis bullosa patient-derived fibroblasts and iPSCs. *J. Invest. Dermatol.* 140, 338–347.e335.
- Huang, K.C., Wang, M.L., Chen, S.J., Kuo, J.C., Wang, W.J., Nhi Nguyen, P.N., Wahlin, K.J., Lu, J.F., Tran, A.A., Shi, M., et al. (2019). Morphological and molecular defects in human three-dimensional retinal organoid model of X-linked juvenile retinoschisis. *Stem Cell Rep.* 13, 906–923.
- Grunewald, J., Zhou, R., Garcia, S.P., Iyer, S., Lareau, C.A., Aryee, M.J., and Joung, J.K. (2019). Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. *Nature* 569, 433–437.
- Rees, H.A., Wilson, C., Doman, J.L., and Liu, D.R. (2019). Analysis and minimization of cellular RNA editing by DNA adenine base editors. *Sci. Adv.* 5, eaax5717.
- Richter, M.F., Zhao, K.T., Eton, E., Lapinaite, A., Newby, G.A., Thuronyi, B.W., Wilson, C., Koblan, L.W., Zeng, J., Bauer, D.E., et al. (2020). Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nat. Biotechnol.* 38, 883–891.
- Li, J., Yu, W., Huang, S., Wu, S., Li, L., Zhou, J., Cao, Y., Huang, X., and Qiao, Y. (2021). Structure-guided engineering of adenine base editor with minimized RNA off-targeting activity. *Nat. Commun.* 12, 2287.
- Liu, Y., Zhou, C., Huang, S., Dang, L., Wei, Y., He, J., Zhou, Y., Mao, S., Tao, W., Zhang, Y., et al. (2020). A Cas-embedding strategy for minimizing off-target effects of DNA base editors. *Nat. Commun.* 11, 6073.
- Walton, R.T., Christie, K.A., Whittaker, M.N., and Kleinstiver, B.P. (2020). Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* 368, 290–296.
- Cao, X., Guo, J., Huang, S., Yu, W., Li, G., An, L., Li, X., Tao, W., Liu, Q., Huang, X., et al. (2022). Engineering of near-PAMless adenine base editor with enhanced editing activity and reduced off-target. *Mol. Ther. Nucleic Acids* 28, 732–742.
- Zhu, Z., Li, C., Zeng, Y., Ding, J., Qu, Z., Gu, J., Ge, L., Tang, F., Huang, X., Zhou, C., et al. (2017). PHB associates with the HIRA complex to control an epigenetic-metabolic circuit in human ESCs. *Cell Stem Cell* 20, 274–289.e277.
- Park, J.-C., Jang, H.-K., Kim, J., Han, J.H., Jung, Y., Kim, K., Bae, S., and Cha, H.-J. (2022). High expression of uracil DNA glycosylase determines C to T substitution in human pluripotent stem cells. *Mol. Ther. Nucleic Acids* 27, 175–183.
- Kuijk, E., Jager, M., van der Roest, B., Locati, M.D., Van Hoeck, A., Korzelius, J., Janssen, R., Besselink, N., Boymans, S., van Boxtel, R., et al. (2020). The mutational impact of culturing human pluripotent and adult stem cells. *Nat. Commun.* 11, 2493.