Comprehensive analysis of prime editing outcomes in human embryonic stem cells

Omer Habib†, Gizem Habib†, Gue-Ho Hwang and Sangsu Bae [ORCID]*

Department of Chemistry and Research Institute for Convergence of Basic Sciences, Hanyang University, Seoul 08826, South Korea

ABSTRACT

Prime editing is a versatile and precise genome editing technique that can directly copy desired genetic modifications into target DNA sites without the need for donor DNA. This technique holds great promise for the analysis of gene function, disease modeling, and the correction of pathogenic mutations in clinically relevant cells such as human pluripotent stem cells (hPSCs). Here, we comprehensively tested prime editing in hPSCs by generating a doxycycline-inducible prime editing platform. Prime editing successfully induced all types of nucleotide substitutions and small insertions and deletions, similar to observations in other human cell types. Moreover, we compared prime editing and base editing for correcting a disease-related mutation in induced pluripotent stem cells derived form a patient with α 1-antitrypsin (A1AT) deficiency. Finally, whole-genome sequencing showed that, unlike the cytidine deaminase domain of cytosine base editors, the reverse transcriptase domain of a prime editor does not lead to guide RNA-independent off-target mutations in the genome. Our results demonstrate that prime editing in hPSCs has great potential for complementing previously developed CRISPR genome editing tools.

INTRODUCTION

Human pluripotent stem cells (hPSCs), including embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs), have the capacity to differentiate into cells derived from the three embryonic germ layers. The ability to engineer the genome of hPSCs has led to the generation of powerful *in vitro* systems in the fields of developmental biology, disease modeling, drug discovery and stem cell-based therapeutics (1–4).

Recently, the CRISPR–Cas9 system has been developed as a major tool for genome editing. This system relies on customizable RNAs that guide Cas9 nucleases to target DNA sequences, resulting in the generation of a double strand break (DSB) (5–8); the response of the cell's repair systems, such as nonhomologous end-joining (NHEJ) or homology-directed repair (HDR), leads to targeted genome editing. However, it has been reported that even a single DSB can induce chromosomal rearrangements, deletions or large deletions (9,10), and, especially in hPSCs, can lead to the activation of the p53 pathway, resulting in apoptosis (11,12). Moreover, the efficiency of genome editing via CRISPR–Cas9-induced DSBs is generally low in hPSCs (8,13). Alternatively, DNA base editors (BEs), including cytosine base editors (CBEs) and adenine base editors (ABEs) can generate single-nucleotide substitutions without introducing a DSB (14). But these tools have potential limitations, such as a restricted editing window, bystander nucleotide editing, and applications confined to transition mutations (C●G to T●A conversions for CBEs and A●T to G●C conversions for ABEs) (15). Furthermore, several studies have revealed genome-wide CBE-mediated off-target DNA editing, transcriptome-wide CBE- and ABE-mediated off-target RNA editing, and unwanted, incorrect editing such as ABE-mediated cytosine deamination (16–18). Recently, engineered base editors have been developed for reducing genome- and transcriptome-wide off-target effects and for increasing editing activities, but they still have limitations in terms of editing ability (19–24).

Prime editors (PEs) are newly developed genome editing tools that can introduce all types of base-to-base conversions (i.e. transition and transversion mutations) as well as small insertions and deletions (indels) without inducing DSBs (25). PE2 consists of a Cas9 nickase (nCas9) containing a H840A mutation and an engineered reverse transcriptase. PE2 employs a prime editing guide RNA (pegRNA) to find its target site and installs the desired edit at the site. The pegRNA, an engineered form of standard single-guide RNA (sgRNA), has an extension sequence at the 3′ end that includes a primer binding site (PBS) that primes the RT reaction and an RT template that encodes the desired edit. When the PE-pegRNA complex binds to the target DNA,

---

*To whom correspondence should be addressed. Tel: +82 2 2220 2618; Fax: +82 2 2298 0319; Email: sangsubae@hanyang.ac.kr
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.
Present address: Omer Habib, CTCELLS Inc., 206-C, R7, DGIST, Daegu 42988, South Korea.

the nCas9 domain nicks the strand containing the protospacer adjacent motif (PAM) and generates a 3′ end at the target site. Next, the liberated 3′ end hybridizes to the PBS of the pegRNA to prime the RT reaction, after which the RT incorporates the desired sequences into the target using the pegRNA as a template. Annealing of the newly generated 3′ flap with the non-edited strand exposes a 5′ flap that is subsequently removed by the DNA repair machinery (Supplementary Figure S1). Introduction of a second nick into the opposite strand with an additional sgRNA (a nicking sgRNA) can further improve prime editing efficiency; systems that incorporate a nicking sgRNA are designated as PE3. Thus, both PE2 and PE3 are capable of copying information directly from the pegRNA into the target DNA sequence. Recently, a study harnessed PEs to edit a reporter gene in hiPSCs (26), but PEs have not been extensively tested or characterized in hPSCs.

Here, we constructed a drug-inducible, PE2-expressing hESC line. We demonstrated the value of this PE platform for introducing small indels and all types of nucleotide substitutions at target sites by testing 63 pegRNAs. In addition, we compared PEs with BEs for the correction of a disease-causing mutation in patient-derived iPSCs. Finally, through whole genome sequencing (WGS), we revealed that prolonged expression of PE2 in hPSCs does not result in the generation of unintended mutations in the genome.

MATERIALS AND METHODS

hPSC culture

Cells were maintained in Essential 8 (E8) medium (Gibco A1517001) on iMatrix-511 (Matrixome, 892 021) and fed daily. ReLeSR (Stemcell Tech., 05873) was used to dissociate hPSCs into clumps for passaging, after which cells were replated in E8 containing p160-Rho-associated coiled-coil kinase (ROCK) inhibitor Y-27632 (2 μM; Stemcell Tech, 72307).

Generation of cell lines with inducible Cas9, BEs (CBE/ABE) and PE2 expression

The plasmid for inducible PE2 expression used in this study was derived from pAAVS1-NDi-CRISPRi (Addgene, 73497) by replacing the sequences encoding KRAB-dCas9-mCherry with that encoding PE2 (amplified from pCMV-PE2; Addgene, 132775). CBEs and ABEs were amplified form pCMV_AncBE4max (Addgene, 112094) and ABE8e (Addgene, 138489) respectively and cloned into pAAVS1-PDi-CRISPRn (Addgene, 75300) by replacing with the sequence encoding Cas9. IRES2 EGFP was introduced downstream of AncBE4max and T2A mCherry was fused with ABE8e to monitor transgene expression. To generate iPE2, iCas9, iCBE and iABE cell lines, two million H9 hESCs were co-electroporated with the appropriate knockin vector (5 μg) and plasmids encoding AAVS1-targeting TALENs (2 μg; addgene, 59025 and 59026) using an Amaxa 4D Nucleofector system (Lonza). Serial cell dilutions were then seeded in six-well plates in E8 supplemented with Y-27632 (10 μM). After selection with the appropriate antibiotic, clones were picked, expanded, and screened by treating with dox and staining for Cas9. For genotyping,

genomic DNA was extracted with a DNeasy Blood & Tissue Kit. KOD -Multi & Epi (Toyobo, KME-101) was used for junction PCR according to the manufacturer's protocol.

Western blotting

Harvested cells were lysed in Ripa buffer (GenDEPOT, R4100-010) supplemented with a 1:100 dilution of protease inhibitor cocktail (GenDEPOT, P3300-001) on ice. Cell lysates were centrifuged at 13 000g for 15 min at 4°C, and the total protein concentration in the supernatant was quantified using a BCA Protein Assay Kit (Pierce, 23225) and mixed with loading buffer before boiled. Samples were resolved with SDS-PAGE and antibodies were used for capturing GAPDH (Cell Signaling, 3683) and Cas9 (Biolegend, 844301).

Electroporation of pegRNA or sgRNA

iPE2 knockin hPSCs (H9-iPE2 cells) were treated with dox (1 μg/ml) for 1 or 2 days before and during electroporation. Before electroporation, accutase (Stemcell Technologies, 07920) was used to generate a suspension of single cells. 1×10^5 cells were electropotated with 250 ng of pegRNA-encoding plasmid and 83 ng of sgRNA-encoding plasmid (for PE3) using a NEON system (ThermoFisher) at 1050 V 30 ms (two pulses). Cells were then seeded in 48-well plates in E8 supplemented with Y-27632 (10 μM for 24 h). Electroporation of sgRNA-encoding plasmid into the H9-iCas9, H9-iABE and H9-iCBE cells was performed under the same conditions. The same condition was used for base editing to correct the PiZZ mutation in a patient derived iPSC (CReM Boston University, PiZZ1). For PE mediated correction of the PiZZ mutation in the PiZZ1 cells, the cells were electroporated with 750 ng of pCMV-PE2, 250 ng of pegRNA, and 83 ng of nicking sgRNA-encoding plasmid.

High-throughput DNA sequencing and analysis

Cells were cultured for 3 days following nucleofection, after which the medium was removed and the cells were washed with 1× phosphate-buffered saline (Wellgene, LB 001-01). For genomic DNA extraction, 50 μl lysis buffer (10 mM Tris–HCl, 0.05% SDS, 25 μg/ml proteinase K) was added directly into each well of the tissue culture plate, which were then incubating at 37°C for 1 h. Proteinase K was then inactivated by incubation at 85°C for 15 min. For sequencing library generation, target sites were amplified directly from genomic DNA in the lysate using a KAPA HiFi HotStart PCR kit (KAPA Biosystems, KK2501) as previously described (20). Briefly, each target was amplified using primers containing Illumina forward and reverse adapters followed by a second amplification with Illumina barcoding primers. The second-round PCR products were pooled and purified using a PCR purification kit (GeneAll, 103-150). Libraries were quantified and subjected to paired-end read sequencing using Illumina MiniSeq platform. For sequence analysis, alignment of raw reads to a reference sequence was performed using PE analyzer (http://www.rgenome.net/pe-analyzer/) (27). For whole genome sequencing, NGS data was generated by the Hiseq and was

mapped to the human genome (GRCh38) using iSAAC aligner (iSAAC-04.18.11.09). SNV and Indels was analyzed by Strelka (2.9.10) and was annotated by SnpEff (v4.3t).

Immunofluorescence

Cells were fixed with ice cold methanol and permeabilized with 0.1% Triton X-100 in phosphate-buffered saline for 30 min at room temperature. The cells were then washed thrice with wash buffer (0.03% Triton X-100 in phosphate-buffered saline). The fixed samples were blocked for 1 h with 5% bovine serum albumin (Gendepot, A0100) in wash buffer, followed by incubation for 24 h at 4°C with anti-Cas9 (BioLegend, 844301) primary antibody diluted in wash buffer. The samples were washed with wash buffer, and then incubated with secondary antibodies at room temperature for 1.5 h. To visualize cell nuclei, the cells were counterstained with 4′,6-diamidino-2-phenylindole (DAPI) (Thermo Fisher Scientific, R37606). Cells were then viewed and photographed using a fluorescence microscope (Zeiss, axio observer z1).

Quantitative real-time PCR

Total RNA was extracted from cells with a Direct-zol RNA Miniprep Kit (Zymo, R2050) according to the manufacturer's instructions. The RNA was eluted from the columns with RNase-free water and quantified by Nanodrop. cDNA synthesis was performed with the PrimeScript™ RT Master Mix (Takara, RR036A) using 2 ug total RNA according to the manufacturer's recommendations. The quantitative real-time PCR (qRT-PCR) was performed in triplicate with the Power SYBR Green Master Mix (Applied Biosystems, 4367659). For real-time amplification, 5 ng total RNA was used as the template. GAPDH was used for qRT-PCR normalization. The primer sequences are listed in Supplementary datasheet.

In vitro differentiation

Eight thousand cells were seeded in 96-well ultra-low attachment round-bottom plates with Essential 6 media (Gibco, A1516401). Half-media changes were performed every 3 days. After day 10 of differentiation, spin embryoid bodies (EBs) were then collected and subjected to qRT-PCR.

Statistics

The Student's *t*-test was used to test for the significance. Mean values ± standard deviation (SD) are shown. *P*-values <0.05, two-tailed test, were considered statistically significant. Statistical analysis is performed with a Graph-Pad Prism 9 program.

RESULTS

Construction of a doxycycline-inducible, PE2-expressing hPSC line named H9-iPE2

To systematically investigate the editing features of PEs in hPSCs, we constructed an hPSC line in which PE2 expression can be induced by doxycycline (dox). First, we prepared a construct that contains the PE2 sequence under the control of an inducible TetO promoter and a neomycin selection cassette (Figure 1A). The PE2 construct was integrated into the first intron of the *PPP1R12C* gene (also known as the AAVS1 site) in human H9 ESCs by using transcription activator-like effector nuclease (TALEN)-mediated gene targeting. After neomycin selection, individual colonies were picked and expanded for genotyping. We examined whether the PE2 construct was properly integrated into the target site using PCR with specific primer sets spanning the junctions between the endogenous sequences and the integrated donor DNA (Figure 1B) and confirmed the junctions by Sanger sequencing (Supplementary Figure S2). For a selected cell line named H9-iPE2, we found using immunostaining that PE2 was specifically expressed in the presence of dox (Figure 1C). To examine the editing capacity of the H9-iPE2 cells, we first determined prime editing efficiencies with three *HEK3*-targeting pegRNAs (Supplementary datasheet), which were previously used in HEK293T cells (25). Due to its higher editing efficiency compared with that of PE2, we used the PE3 system throughout the study. The H9-iPE2 cells were electroporated with each pegRNA and a nicking sgRNA (targeting a site +63 nt away from the 5′ end of the pegRNA-induced nick) and treated with dox for 3 days in a 48-well plate. Then, high-throughput sequencing was used to quantify the editing efficiency in bulk populations. The results revealed that PE3 introduced the intended mutations with editing efficiencies between 4.3–24.8%, supporting the editing capacity of the H9-iPE2 cells (Figure 1D). We also observed that PE3 generated unwanted indels at a frequency of 2.7–8.1%.

PE-mediated editing at endogenous loci in hPSCs

We next investigated various features of PE3-mediated editing in H9-iPE2 cells. It is known that the PE-mediated editing efficiency is dependent on the lengths of both the PBS and RT template in the pegRNA (25). To examine this point in H9-iPE2 cells, we selected target sites at two different loci, *HEK3* and *RNF2*, and tested pegRNAs with varying PBS lengths together with corresponding nicking sgRNAs. In a previous study using HEK293T cells, the Liu group reported that efficient prime editing at the *RNF2* site, which has a relatively low G/C content (35%), required a longer PBS sequence than did efficient prime editing at the *HEK3* site, which has a G/C content of 71%. Similarly, we found that efficient prime editing at the *RNF2* site required a longer PBS sequence than at the *HEK3* site (Figure 2A and B), suggesting that the optimum PBS length is sequence dependent and might be conserved among cell types (Supplementary Figure S3). In addition, similar to results from the previous study in HEK293T cells, we found that PEs can introduce both transition and transversion substitutions across the +1 to +33 positions of the *HEK3* target site with varying efficiencies in hPSCs (Figure 2C and D). We also tested the ability of PE3 to introduce small indels (1–3 bp) at the *HEK3* and *RNF2* sites and found that PE3 successfully generated the intended edits to +10 position, but not for +17 and +21 positions (Figure 2E and F). Finally, we demonstrated that large deletions (up to 80 bp) and insertions (up to 42 bp) were precisely installed by PE3 (Figure 2G and H). From these results, we concluded that
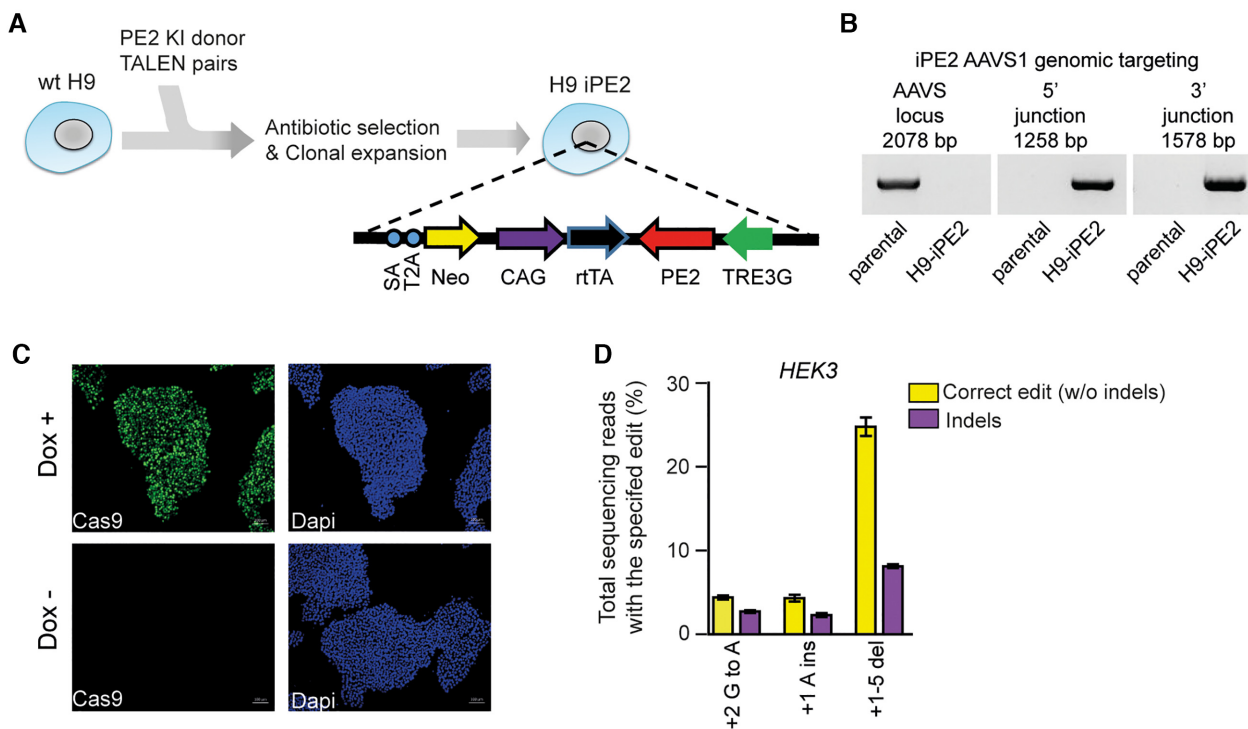
Figure 1. Generation and characterization of H9-iPE2. (**A**) Schematic diagram of the strategy for TALEN-mediated targeting of the AAVS1 locus to generate H9-iPE2 cells, in which PE2 expression is induced by dox. The AAVS1 donor vector contains a cassette in which PE2 expression is under the control of the dox-inducible TRE3G promoter. SA, splice acceptor; 2A, self-cleaving 2A peptide; Neo, neomycin resistance gene; rtTA, dox-controlled reverse transcriptional activator; CAG, cytomegalovirus early enhancer/chicken β actin promoter. (**B**) PCR-based confirmation that the iPE2 construct was targeted to the AAVS1 locus. A primer pair that flanks the AAVS1 knock-in site only amplifies a product in wild-type cells; lack of a product indicates that the H9-iPE2 clone used in this study is homozygous (left panel). Junction PCR confirmed on-target integration of the cassette into the AAVS1 locus (center and right panels). (**C**) Immunostaining of H9-iPE2 cells before and after 48 h of dox treatment with an anti-Cas9 antibody (green). Nuclei were stained with DAPI (blue). (**D**) Editing efficiency in H9-iPE2 cells. H9-iPE2 cells maintained in the presence of dox were electroporated with plasmids encoding a pegRNA and a nicking sgRNA (for PE3 system). The editing efficiency is indicated as the percentage of total sequencing reads that contain the intended edit and do not contain indels. Mean ± s.d. of $n = 3$ independent biological replicates.

PE3 can be a practical tool for inducing a broad range of mutations in hPSCs.

Double nicking strategy of PE3 is not the primary reason for unwanted indel generation

We next sought to understand the reason why indel mutations are generated by the PE3 system. To investigate which factor has a dominant role for generating such indel patterns, we examined five different scenarios: (i) a single DSB induced by wild-type (wt) Cas9 and a sgRNA, (ii) a single nick induced by nCas9-RT and a sgRNA, (iii) double DSBs induced by wt Cas9 and two sgRNAs, (iv) double nicks induced by nCas9-RT and two sgRNAs and (v) double nicks induced by nCas9-RT, one pegRNA and one sgRNA (i.e. the PE3 system) (Figure 3A). We tested all of these possibilities at both the *HEK3* and *RNF2* sites. The distances between the first and second nicks were 63 and 41 nt for the *HEK3* and *RNF2* sites, respectively. To maintain consistency with the PE3 system, we additionally established an H9 line in which wt Cas9 was expressed in a dox-inducible manner, named H9-iCas9 (Supplementary Figure S4). We also set the ratio of the first sgRNA to the second sgRNA at 3:1.

We first assessed indel frequencies after the introduction of identical sgRNAs into H9-iCas9 and H9-iPE2 cells. High-throughput sequencing data showed that the resulting single DSBs (scenario i) led to high frequencies of indels (68% and 79%) at the *HEK3* and *RNF2* sites, respectively, whereas no significant indel frequencies were observed after the introduction of single nicks (ii) (Figure 3B and C). We also electroporated a pair of sgRNAs to examine the effects of double DSBs or nicks. The sequencing data showed that double DSBs (iii) resulted in high indel frequencies, but that double nicks (iv) did not (indel frequencies of <2%), suggesting that the generation of double nicks in and of itself is not a major reason why indels are generated and that the RT enzyme in PE3 cannot generate indel mutations alone. However, when double nicks were produced following the introduction of a pegRNA and a sgRNA (v), the frequency of indels substantially increased (up to 12%), suggesting that the combination of the RT enzyme and a pegRNA could cause indel mutations. We further investigated the indel patterns generated following a single DSB (i), double DSBs (iii) and the double nicks introduced by PE3 (v). Five recurrent deletion patterns were obtained in each situation for the comparison. PE3 generated long deletions that included the region upstream of the first DSB site, whereas a single DSB led to the generation of small deletions near the
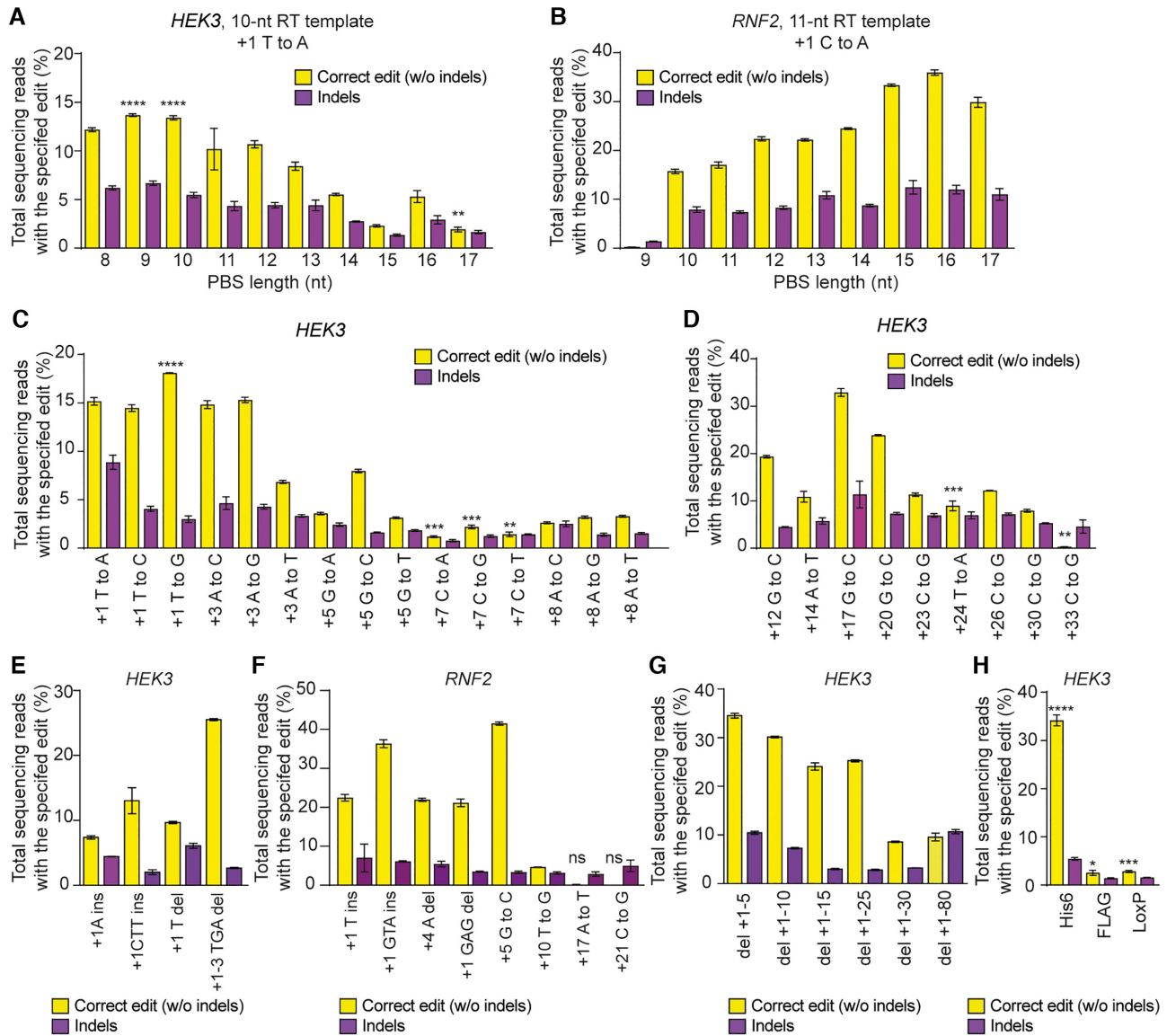
Figure 2. Prime editing of genomic DNA in H9-iPE2 cells by PE3. (**A**, **B**) PE3 editing efficiencies at *HEK3* and *RNF2* genomic sites with pegRNAs containing varying PBS lengths. (**C**) Efficiencies of all 12 types of transition and transversion edits at the indicated positions in the *HEK3* site. (**D**) Efficiencies of long distance edits at the *HEK3* site using a 34-nt RT template. (**E**, **F**) Editing efficiencies for the generation of intended indels at the *HEK3* and *RNF2* genomic sites. Yellow, desired indel; purple, undesired indel. (**G**) Editing efficiencies for the generation of targeted deletions of 5–80 bp at the *HEK3* site. (**H**) Editing efficiencies for the generation of targeted insertions of a His6 tag, Flag epitope tag, and loxP site at the *HEK3* site. The editing efficiency is indicated as the percentage of total sequencing reads that contain the intended edit and do not contain unintended indels. Mean ± s.d. of *n* = 3 independent biological replicates.

DSB site and double DSBs typically induced precise deletion of the region between the two DSB sites (Figure 3D and E). These results also supported our conclusion that unintended indels are not directly caused by double nicking per se but by the combinatory activity of the RT and the pegRNA.

Comparison of prime editing and base editing in hPSCs

We next compared the activity of PEs and BEs at two genomic sites (*HEK3* and *FANCF*) in hPSCs. To make experiment conditions equal, we established H9-iCBE and H9-iABE lines expressing AncBE4max and ABE8e, respectively, under the control of dox (Supplementary Figure S5). For H9-iPE2, H9-iCBE and H9-iABE lines, we conducted western blot experiments to measure the expression level of PE2, CBE and ABE proteins respectively in the presence of dox and found that PE and CBE showed similar levels and that the level of ABE was relatively low but comparable with them (Supplementary Figure S5).

We first measured the activity of AncBE4max, an optimized CBE variant (28), at both the *HEK3* and *FANCF* sites. The high-throughput sequencing data showed that several cytosines within the editing window were edited at a frequency of up to 75% at the *HEK3* site (Figure 4A) and 81% at the *FANCF* site (Figure 4B), indicating CBE by-
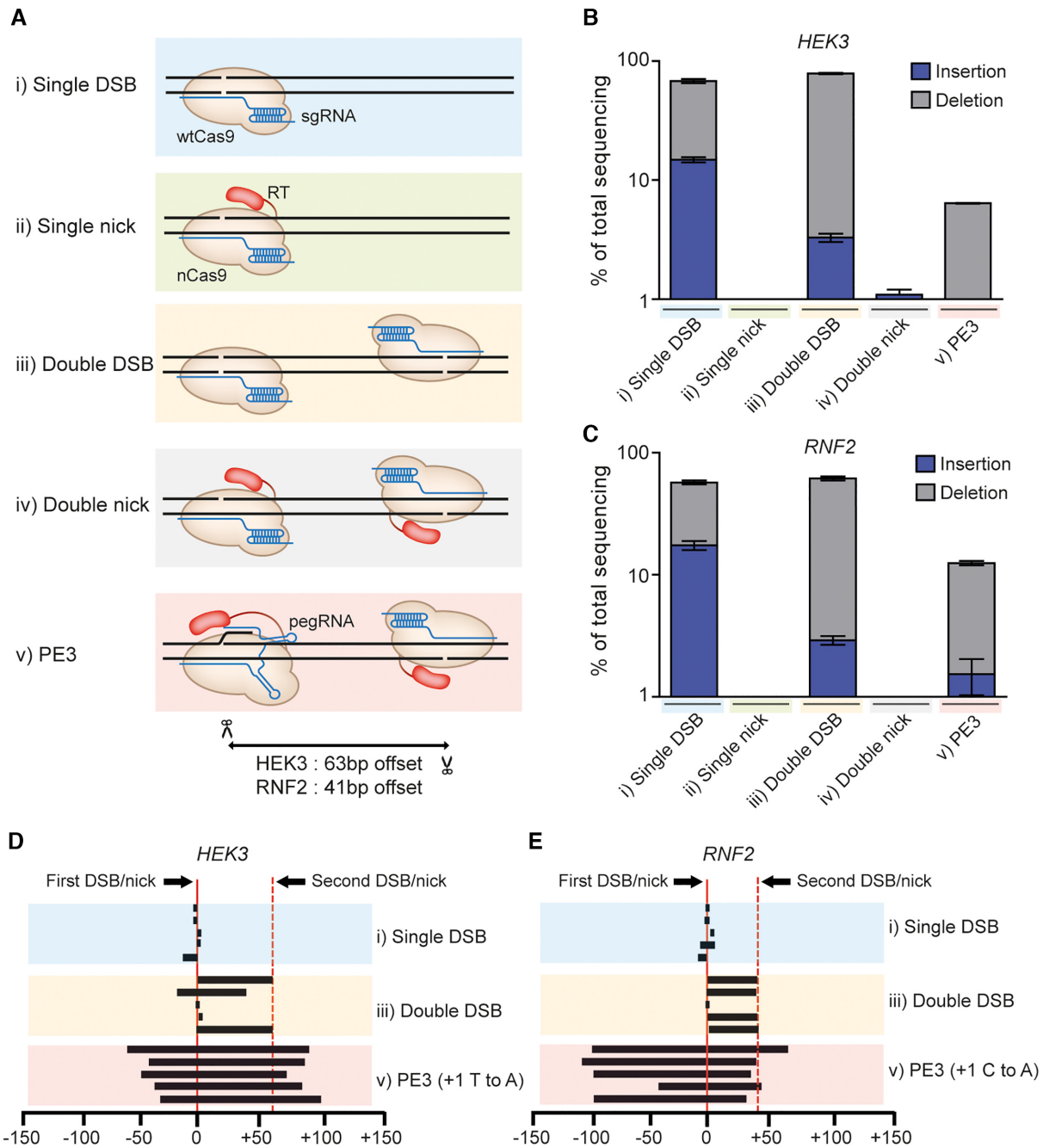
Figure 3. Comparison of indel frequencies and endpoints after different DNA cleavage and nicking scenarios. (**A**) Schematic of enzymes and guide RNAs tested to compare indel frequencies. (**B, C**) Indel frequencies generated by the experimental set ups diagrammed in (A) at the *HEK3* and and *RNF2* genomic sites. (**D, E**) Deletion patterns obtained after single cuts, double cuts, and PE3-mediated base substitutions at the *HEK3* and *RNF2* sites. Each line designates the extent of a deletion. The five most frequent types of deletions are shown for each case. Mean ± s.d. of *n* = 3 independent biological replicates.

stander effects. In contrast, for the PE experiment, we tested individual pegRNAs designed to convert each cytosine at both the *HEK3* and *FANCF* sites. The high-throughput sequencing data showed that each pegRNA led to precise editing of the targeted cytosine, although the conversion rates of the PEs were typically lower than that of the CBEs (Figure 4A and B). We also tested ABE8e, an optimized ABE

variant (29), at the same target sites, and found that they exhibited similar tendencies as the CBEs (Figure 4C and D). In these experiments, we found that both CBEs and ABEs generally showed higher activities than PEs. However, we also found that bystander nucleotide conversion by BEs was inevitable when the target contained multiple cytosines or adenines in the editing window, whereas PEs precisely con-
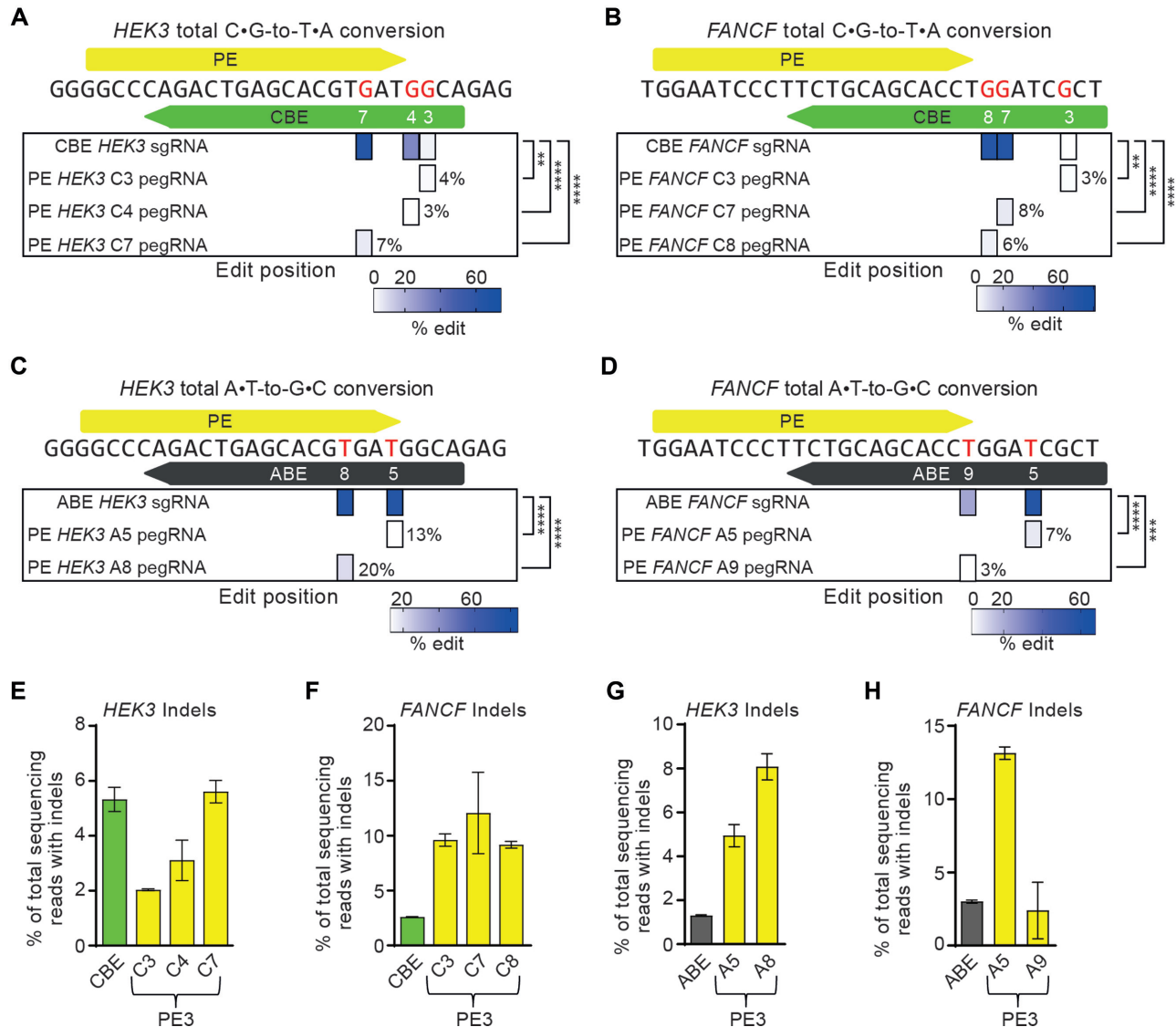
Figure 4. Comparison of prime editing and base editing outcomes at the same target sites. (**A**, **B**) CBE- and PE-generated frequencies of C•G-to-T•A edits at target nucleotides, highlighted in red, in the endogenous *HEK3* and *FANCF* sites. Base positions are numbered relative to the PAM of the CBE sgRNA. The PAM nucleotides are numbered 21–23. (**C**, **D**) ABE- and PE-generated frequencies of A•T-to-G•C edits at endogenous *HEK3* and *FANCF* sites. *******P* < 0.01, ********P* < 0.001, *********P* < 0.0001. for (A)–(D). (**E**, **F**) Indel frequencies from the experiments in (A) and (B). (**G**, **H**) Indel frequencies from the experiments in (C) and (D). Mean ± s.d. of *n* = 3 independent biological replicates.

verted target nucleotides. We also measured indel frequencies in all experiments and found that PE3s were associated with apparent indels (at frequencies ranging from 2% to 13%) (Figure 4E–H). Although BEs induced detectible indels at frequencies ranging from 1% to 5%, they showed better edit:indel ratios than PEs. Finally, we performed spontaneous differentiation in PE and BE edited cells and showed that both cell lines expressed markers of all three germ layers (Supplementary Figure S6). Taken together, our results indicate that because of their high activity, BEs are preferred over PEs for targets that lack bystander nucleotides or when bystander edits are acceptable. However, PEs edit the single base of interest by using a customized pegRNA for each base, although indel generation can occur in the case of the PE3 system.

Correction of a pathogenic mutation in iPSCs with PEs and BEs

We next applied PEs and BEs to correct a disease-causing mutation in a patient-derived iPSC line carrying a homozygous PiZZ mutation (1024 G > A, E342K) in the serpin family A member 1 (*SERPINA1*) gene, the most common mutation causing α 1-antitrypsin (A1AT) deficiency (30,31). We first tested the ability of an ABE to correct this mutation. In this experiment, we used an NG PAM-targetable ABEmax (named NG-ABEmax) and selected a sgRNA in which the target nucleotide is the seventh adenine from the 5′ end (A7). High-throughput sequencing data from NG-ABEmax-treated iPSCs showed that this ABE yielded A-to-G conversions at a frequency of 2.5%, with a 3.7% conversion rate for a bystander adenine at the fifth
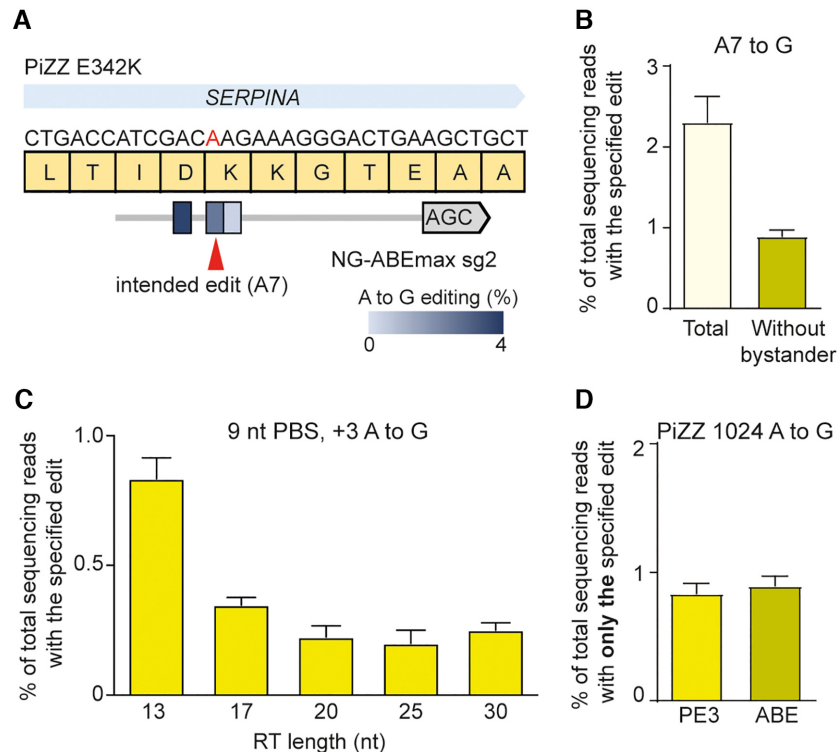
Figure 5. Repair of the PiZZ 1024 G > A mutation in induced pluripotent stem cells (iPSCs) derived from a patient with alpha-1-antitrypsin deficiency. (**A**) Schematic of the disease-associated 1024 G > A mutation in the *SERPINA1* gene, the altered protein sequence (E342K), and the sgRNA used to correct the mutation with NG-ABEmax. (**B**) Frequencies of A-to-G conversions in the editing window induced by NG-ABEmax. (**C**) Frequences of A-to-G conversions at the G > A mutation site induced by NG-PE3 system. pegRNAs with varying RT lengths were tested. (**D**) Comparison of the frequencies of precise correction of the PiZZ mutation in patient-derived iPSCs by NG-PE3 and NG-ABEmax. Mean ± s.d. of $n = 3$ independent biological replicates.

position that ultimately results in an Asp-to-Gly mutation (Figure 5A). When we considered the frequency of the intended edit alone (i.e., the A7 correction without bystander editing), it was only 0.85% (Figure 5B).

We next tested a version of PE2 that recognizes the canonical PAM (NGG) for correction of the PiZZ mutation. We designed pegRNAs with varying PBS lengths, with the targeted adenine at the +24 position. We then electroporated the pegRNAs with an additional nicking sgRNA (for PE3 system) into the patient-derived iPSCs; however, we did not observe any detectable editing (Supplementary Figure S7), similar to a previous study in liver organoids (32). Therefore, we also tested an NG PAM-targetable PE2 (NG-PE2) with an additional nicking sgRNA, which changed the edit position from +24 to +3. This PE3 system led to a low but detectable frequency of gene correction (0.83%) (Figure 5C). We assume that the low editing efficiency might be cause by a low transfection efficiency and/or a low editing activity of the used pegRNA. Thus, in this experiment, PE and ABE demonstrated comparable efficiencies of generating only the desired edit (Figure 5D).

PE is a safe editing tool in hPSCs

BEs and PEs are similar in that they both contain nCas9 domains for binding target sequences, but differ in their functional components: BEs include a cytidine/adenosine deaminase and PEs a RT. Previously, several groups reported that the cytidine deaminases in CBEs could generate sgRNA-independent DNA/RNA conversions, whereas the adenosine deaminases in ABEs could generate sgRNA-independent RNA conversions (16–19). Likewise, we wondered whether PEs would exhibit pegRNA-independent RT enzyme effects. To the best of our knowledge, the effects of prolonged PE expression in hPSCs have not been reported so far. Therefore, we set out to investigate whether long term PE expression would cause undesired mutations (Figure 6A). The H9-iPE2 cells were maintained in the absence or presence of dox for 3 weeks, after which individual clones were selected and cultured in the absence of doxycycline. After expansion of the cells, genomic DNA was isolated and subjected to whole genome sequencing (WGS). We did not find any increase in the number of single nucleotide variations (SNVs) or indels in clones in which PE2 was overexpressed compared to control clones (Figure 6B and C). We also confirmed that the most SNVs (92.23%) were common for all clones (Figure 6D), suggesting that pegRNA-independent off-target effects in DNA might be negligible in hPSCs.

DISCUSSION

Since its first demonstration in cancer cells and mouse cortical neurons, prime editing has been applied to various organisms and model systems including plants, human intestinal and liver organoids, Drosophila, mouse embryos and
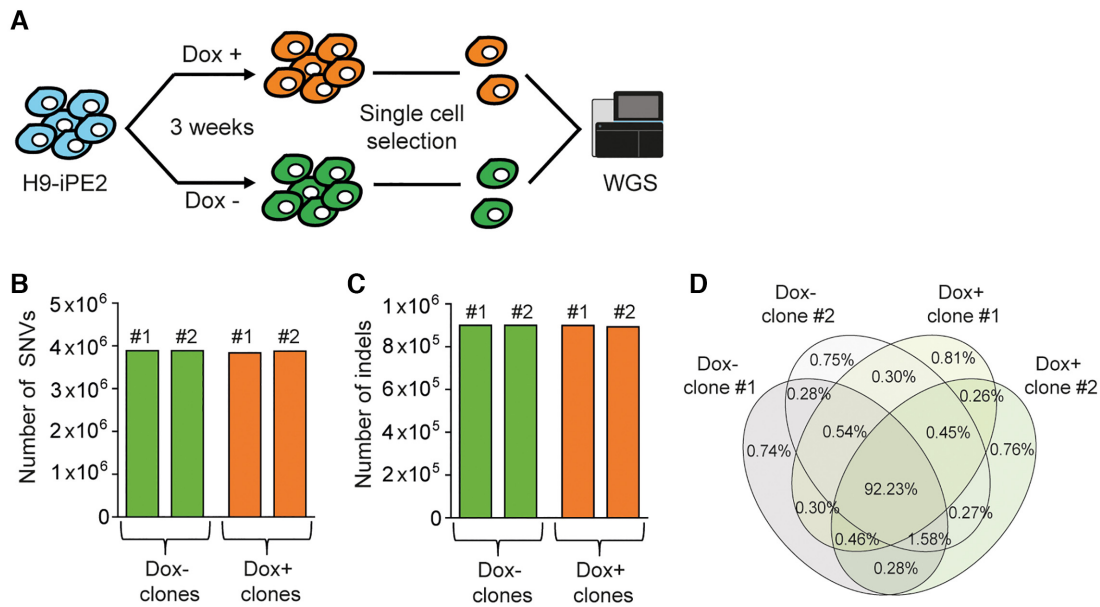
Figure 6. PE does not lead to genome-wide off-target effects. (**A**) Schematic diagram of the experimental design for clonal expansion and whole genome sequencing analysis of H9-iPE2 cells with or without dox-induced PE2 expression for 3 weeks. (**B**, **C**) Numbers of mutations in uninduced and induced H9-iPE2 clones. The numbers represent all sequence variations, including indels and single nucleotide variations. (**D**) Venn diagrams showing the ratios of common SNVs in H9-iPE2 clones.

adult mice (32–41). Here, we expanded this list to include hPSCs. In this study, we demonstrated prime editing in hESCs using a dox-inducible, PE2-expressing line. Through comprehensive tests and analysis of PE3 activity at endogenous targets, we observed several strengths and weaknesses of this system in hESCs, as follows: (i) PEs are capable of introducing all types of base substitutions at positions as far as 30 bps downstream of the nicking site in hPSCs and showed a trend of activity similar to that seen in other human cell lines such as HEK293T, but in general, prime editing efficiencies are lower in hPSCs than in other cell lines, (ii) Compared to PEs, BEs showed higher base conversion activities but generated bystander edits, whereas PEs provided precise base conversion without bystander edits, (iii) PE3-mediated base conversion is generally accompanied by the generation of indels, which are mainly caused by the combinatory activity of RT and the pegRNA, rather than by the double nicking strategy per se and (iv) WGS results revealed that long term overexpression of PE2 in hESCs did not generate significant frequencies of genome-wide SNVs or indels. Recent studies in organoid model systems showed that prime editing produces higher desired edit/indel ratios than HDR and that base editing is superior to prime editing in terms of editing efficiency (32,37), which is consistent with our conclusions.

A potential limitation of using canonical CRISPR nucleases in clinical applications is the generation of off-target mutations at sites with high homology to the intended on-target site. However, Anzalone *et al.* reported that PEs produce fewer off-target mutations than Cas9 nucleases and suggested that the three hybridization steps required for prime editing [(i) between the pegRNA spacer and the target DNA for Cas9 binding, (ii) between the pegRNA PBS and the target DNA to initiate reverse tran-

scription and (iii) between the RT product and the target DNA during flap resolution] may reduce off-target mutagenesis. These observations were further corroborated by two studies: nDigenome-seq revealed that PEs are highly precise (42) and WGS analysis of prime-edited clonal intestinal organoids showed that PEs exhibited high fidelity (32). Consistent with previous studies, we revealed here with WGS analysis that long term overexpression of PE2 produced negligible off-target effects in hESCs. These results strongly support further applications of PEs for therapeutic editing in clinical studies (41,43,44).

Although PEs have much lower off-target editing potential than Cas9 nucleases, undesired on-target edits such as indels have been reported for PEs, especially for the PE3 system (25). In a previous study, PE3 led to more undesired outcomes than intended edits at the target site in mouse embryos (40); double nicking associated with the pegRNA and sgRNA in PE3 was proposed as the primary reason for the high indel rates. In contrast, based on our observations summarized in Figure 3, we suggest that the major reason is the combinatory activity of the RT and pegRNA, rather than simply the double nicking strategy. However, the detailed mechanism responsible for indel generation was not elucidated in this study. Further characterization of the repair mechanisms that resolve the flap during prime editing would potentially provide insight into the reason for the high frequencies of PE3-mediated indel formation and allow us to manipulate the repair pathway to improve the accuracy of prime editing.

In addition, despite the unique capabilities of PEs, the complexity of pegRNA design could be a potential drawback for their use. The optimal lengths of the PBS and RT template must be determined because these sequences can have substantial effects on the efficiency of prime editing.

Recently, the Kim group developed computational models for predicting pegRNA efficiency and determined parameters that influence pegRNA activity (45). However, several pegRNAs should still be tested for each specific editing task.

In summary, prime editing adds a new dimension to CRISPR-based approaches and broadens the scope of genome editing in hPSCs by complementing other methods. It is likely that an understanding of the repair mechanisms involved in prime editing and further improvement of PEs will accelerate translation of this technology into therapeutic applications.

DATA AVAILABILITY

High-throughput sequencing and WGS data have been deposited in the NCBI Sequence Read Archive database (SRA; https://www.ncbi.nlm.nih.gov/sra) under accession number PRJNA721791. Any other additional relevant data and materials are available from the authors upon reasonable request.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Hotta,A. and Yamanaka,S. (2015) From genomics to gene therapy: induced pluripotent stem cells meet genome editing. *Annu. Rev. Genet.*, **49**, 47–70.
2. Hockemeyer,D. and Jaenisch,R. (2016) Induced pluripotent stem cells meet genome editing. *Cell Stem Cell*, **18**, 573–586.
3. Komor,A.C., Badran,A.H. and Liu,D.R. (2017) CRISPR-based technologies for the manipulation of eukaryotic genomes. *Cell*, **168**, 20–36.
4. Soldner,F. and Jaenisch,R. (2018) Stem cells, genome editing, and the path to translational medicine. *Cell*, **175**, 615–632.
5. Jinek,M., Chylinski,K., Fonfara,I., Hauer,M., Doudna,J.A. and Charpentier,E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, **337**, 816–821.
6. Cho,S.W., Kim,S., Kim,J.M. and Kim,J.S. (2013) Targeted genome engineering in human cells with the cas9 RNA-guided endonuclease. *Nat. Biotechnol.*, **31**, 230–232.
7. Jinek,M., East,A., Cheng,A., Lin,S., Ma,E. and Doudna,J. (2013) RNA-programmed genome editing in human cells. *Elife*, **2**, e00471.
8. 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.
9. Kosicki,M., Tomberg,K. and Bradley,A. (2018) Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.*, **36**, 765–771.
10. Liu,M., Zhang,W., Xin,C., Yin,J., Shang,Y., Ai,C., Li,J., Meng,F.L. and Hu,J. (2021) Global detection of DNA repair outcomes induced by CRISPR–Cas9. *Nucleic Acids Res.*, **49**, 8732–8742.
11. 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.
12. 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.
13. Kim,K.T., Park,J.C., Jang,H.K., Lee,H., Park,S., Kim,J., Kwon,O.S., Go,Y.H., Jin,Y., Kim,W. *et al.* (2020) Safe scarless cassette-free selection of genome-edited human pluripotent stem cells using temporary drug resistance. *Biomaterials*, **262**, 120295.
14. Rees,H.A. and Liu,D.R. (2018) Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat. Rev. Genet.*, **19**, 770–788.
15. Jeong,Y.K., Song,B. and Bae,S. (2020) Current status and challenges of DNA base editing tools. *Mol. Ther.*, **28**, 1938–1952.
16. 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.
17. Jin,S., Zong,Y., Gao,Q., Zhu,Z.X., Wang,Y.P., Qin,P., Liang,C.Z., Wang,D.W., Qiu,J.L., Zhang,F. *et al.* (2019) Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. *Science*, **364**, 292–295.
18. 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.
19. Zhou,C., Sun,Y., Yan,R., Liu,Y., Zuo,E., Gu,C., Han,L., Wei,Y., Hu,X., Zeng,R. *et al.* (2019) Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. *Nature*, **571**, 275–278.
20. Kim,H.S., Jeong,Y.K., Hur,J.K., Kim,J.S. and Bae,S. (2019) Adenine base editors catalyze cytosine conversions in human cells. *Nat. Biotechnol.*, **37**, 1145–1148.
21. Jeong,Y.K., Lee,S., Hwang,G.H., Hong,S.A., Park,S.E., Kim,J.S., Woo,J.S. and Bae,S. (2021) Adenine base editor engineering reduces editing of bystander cytosines. *Nat. Biotechnol.*, **39**, 1426–1433.
22. Grunewald,J., Zhou,R., Iyer,S., Lareau,C.A., Garcia,S.P., Aryee,M.J. and Joung,J.K. (2019) CRISPR DNA base editors with reduced RNA off-target and self-editing activities. *Nat. Biotechnol.*, **37**, 1041–1048.
23. Lee,S., Ding,N., Sun,Y., Yuan,T., Li,J., Yuan,Q., Liu,L., Yang,J., Wang,Q., Kolomeisky,A.B. *et al.* (2020) Single C-to-T substitution using engineered APOBEC3G-nCas9 base editors with minimum genome- and transcriptome-wide off-target effects. *Sci. Adv.*, **6**, eaba1773.
24. 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.
25. Anzalone,A.V., Randolph,P.B., Davis,J.R., Sousa,A.A., Koblan,L.W., Levy,J.M., Chen,P.J., Wilson,C., Newby,G.A., Raguram,A. *et al.* (2019) Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*, **576**, 149–157.
26. Surun,D., Schneider,A., Mircetic,J., Neumann,K., Lansing,F., Paszkowski-Rogacz,M., Hanchen,V., Lee-Kirsch,M.A. and Buchholz,F. (2020) Efficient generation and correction of mutations in human iPS cells utilizing mRNAs of CRISPR base editors and prime editors. *Genes (Basel)*, **11**, 511.
27. Hwang,G.H., Jeong,Y.K., Habib,O., Hong,S.A., Lim,K., Kim,J.S. and Bae,S. (2021) PE-Designer and PE-Analyzer: web-based design and analysis tools for CRISPR prime editing. *Nucleic Acids Res.*, **49**, W499–W504.
28. Koblan,L.W., Doman,J.L., Wilson,C., Levy,J.M., Tay,T., Newby,G.A., Maianti,J.P., Raguram,A. and Liu,D.R. (2018) Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat. Biotechnol.*, **36**, 843–846.
29. 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.

30. Perlmutter,D.H. (2009) Autophagic disposal of the aggregation-prone protein that causes liver inflammation and carcinogenesis in alpha-1-antitrypsin deficiency. *Cell Death Differ.*, **16**, 39–45.
31. Gooptu,B. and Lomas,D.A. (2009) Conformational pathology of the serpins: themes, variations, and therapeutic strategies. *Annu. Rev. Biochem.*, **78**, 147–176.
32. Schene,I.F., Joore,I.P., Oka,R., Mokry,M., van Vugt,A.H.M., van Boxtel,R., van der Doef,H.P.J., van der Laan,L.J.W., Verstegen,M.M.A., van Hasselt,P.M. *et al.* (2020) Prime editing for functional repair in patient-derived disease models. *Nat. Commun.*, **11**, 5352.
33. Lin,Q., Zong,Y., Xue,C., Wang,S., Jin,S., Zhu,Z., Wang,Y., Anzalone,A.V., Raguram,A., Doman,J.L. *et al.* (2020) Prime genome editing in rice and wheat. *Nat. Biotechnol.*, **38**, 582–585.
34. Xu,W., Zhang,C., Yang,Y., Zhao,S., Kang,G., He,X., Song,J. and Yang,J. (2020) Versatile nucleotides substitution in plant using an improved prime editing system. *Mol Plant*, **13**, 675–678.
35. Hua,K., Jiang,Y., Tao,X. and Zhu,J.K. (2020) Precision genome engineering in rice using prime editing system. *Plant Biotechnol. J.*, **18**, 2167–2169.
36. Lin,Q., Jin,S., Zong,Y., Yu,H., Zhu,Z., Liu,G., Kou,L., Wang,Y., Qiu,J.L., Li,J. *et al.* (2021) High-efficiency prime editing with optimized, paired pegRNAs in plants. *Nat. Biotechnol.*, **39**, 923–927.
37. Geurts,M.H., de Poel,E., Pleguezuelos-Manzano,C., Carrillo,L., Andersson-Rolf,A., Boretto,M., Beekman,J.M. and Clevers,H. (2021) Evaluating CRISPR-based prime editing for cancer modeling and CFTR repair in intestinal organoids. *Life Sci. Alliance.*, **4**, e202000940.
38. Bosch,J.A., Birchak,G. and Perrimon,N. (2021) Precise genome engineering in Drosophila using prime editing. *Proc. Natl. Acad. Sci. U.S.A.*, **118**, e2021996118.
39. Liu,Y., Li,X., He,S., Huang,S., Li,C., Chen,Y., Liu,Z., Huang,X. and Wang,X. (2020) Efficient generation of mouse models with the prime editing system. *Cell Discov.*, **6**, 27.
40. Aida,T., Wilde,J.J., Yang,L., Hou,Y., Li,M., Xu,D., Lin,J., Qi,P., Lu,Z. and Feng,G. (2020) Prime editing primarily induces undesired outcomes in mice. bioRxiv doi: https://doi.org/10.1101/2020.08.06.239723, 06 August 2020, preprint: not peer reviewed.
41. Jang,H., Jo,D.H., Cho,C.S., Shin,J.H., Seo,J.H., Yu,G., Gopalappa,R., Kim,D., Cho,S.R., Kim,J.H. *et al.* (2021) Application of prime editing to the correction of mutations and phenotypes in adult mice with liver and eye diseases. *Nat. Biomed. Eng.*, https://doi.org/10.1038/s41551-021-00788-9.
42. Kim,D.Y., Moon,S.B., Ko,J.H., Kim,Y.S. and Kim,D. (2020) Unbiased investigation of specificities of prime editing systems in human cells. *Nucleic Acids Res.*, **48**, 10576–10589.
43. Zhi,S., Chen,Y., Wu,G., Wen,J., Wu,J., Liu,Q., Li,Y., Kang,R., Hu,S., Wang,J. *et al.* (2021) Dual-AAV delivering split prime editor system for in vivo genome editing. *Mol. Ther.*, https://doi.org/10.1016/j.ymthe.2021.07.011.
44. Kim,Y., Hong,S.A., Yu,J., Eom,J., Jang,K., Yoon,S., Hong,D.H., Seo,D., Lee,S.N., Woo,J.S. *et al.* (2021) Adenine base editing and prime editing of chemically derived hepatic progenitors rescue genetic liver disease. *Cell Stem Cell*, **28**, 1614–1624.
45. Kim,H.K., Yu,G., Park,J., Min,S., Lee,S., Yoon,S. and Kim,H.H. (2021) Predicting the efficiency of prime editing guide RNAs in human cells. *Nat. Biotechnol.*, **39**, 198–206.