Molecular Therapy
Nucleic Acids
Original Article

High expression of uracil DNA glycosylase determines C to T substitution in human pluripotent stem cells

Ju-Chan Park,[1,5] Hyeon-Ki Jang,[2,4,5] Jumee Kim,[1] Jun Hee Han,[3] Youngri Jung,[3] Keuntae Kim,[1] Sangsu Bae,[2,3] and Hyuk-Jin Cha[1]

[1]College of Pharmacy, Seoul National University, 1, Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea; [2]Institute for Convergence of Basic Sciences, Hanyang University, Seoul 04763, Republic of Korea; [3]Department of Chemistry, Hanyang University, 222, Wangsimni-ro, Seongdong-gu, Seoul 04763, Republic of Korea

Precise genome editing of human pluripotent stem cells (hPSCs) is crucial not only for basic science but also for biomedical applications such as *ex vivo* stem cell therapy and genetic disease modeling. However, hPSCs have unique cellular properties compared to somatic cells. For instance, hPSCs are extremely susceptible to DNA damage, and therefore Cas9-mediated DNA double-strand breaks (DSB) induce p53-dependent cell death, resulting in low Cas9 editing efficiency. Unlike Cas9 nucleases, base editors including cytosine base editor (CBE) and adenine base editor (ABE) can efficiently substitute single nucleotides without generating DSBs at target sites. Here, we found that the editing efficiency of CBE was significantly lower than that of ABE in human embryonic stem cells (hESCs), which are associated with high expression of DNA glycosylases, the key component of the base excision repair pathway. Sequential depletion of DNA glycosylases revealed that high expression of uracil DNA glycosylase (UNG) not only resulted in low editing efficiency but also affected CBE product purity (i.e., C to T) in hESCs. Therefore, additional suppression of UNG via transient knockdown would also improve C to T base substitutions in hESCs. These data suggest that the unique cellular characteristics of hPSCs could determine the efficiency of precise genome editing.

INTRODUCTION

With the development of genome editing technologies, human pluripotent stem cells (hPSCs) have become a critical cell resource not only for autologous stem cell therapy but also for disease modeling, thus providing the means to create pathogenic cell models harboring genetic mutations (i.e., an approach referred to as "disease-in-a-dish").[1–3] However, hPSCs are extremely susceptible to double-strand breaks (DSBs) and subsequent p53-dependent cell death through mitochondrial translocation of p53 (referred to as mitochondrial priming to apoptosis),[4,5] which serves as a safeguard to maintain the genome integrity of these cells,[6] all of which affects the editing efficiency of Cas9 nucleases in hPSCs.[7] Several strategies have been tested to enhance Cas9-mediated gene editing activity in hPSCs, including the introduction of survival genes,[8] delivery with Cas9 ribonucleoproteins (RNP),[9] and co-targeting with drug- and/or toxin-resistant genes.[10,11] However, all of these approaches are fundamentally based on the induction of DNA DSBs. Alternatively, DNA base editors (BEs), including cytosine base editor (CBE) and adenine base editor (ABE), can convert single nucleotides without generating DSBs,[12] thus decreasing DNA damage.[13] Therefore, despite the limitations of BEs, such as incompetence of transversion, restricted target sequence accessibility due to the requirement of the protospacer adjacent motif sequence, and undesired mutations of bystander sequences within the target window, BEs have recently garnered much attention and have been applied in hPSCs to produce isogenic disease models with pathogenic mutations[14,15] or gene corrections.[16]

Although no DSB induction occurs by BEs, BEs can induce target DNA damage through base deamination and single-strand breaks, owing to the activities of a Cas9 nickase (nCas9) and conjugated cytidine deaminase or adenosine deaminase.[17,18] Mismatch bases formed by deamination of target bases by BEs are efficiently repaired by the base excision repair (BER) and mismatch repair (MMR) pathways. BER is initiated by the recognition and excision of deaminated bases with base-specific DNA glycosylases (e.g., uracil DNA glycosylase [UNG], thymidine DNA glycosylase [TDG], methyl-CpG binding domain 4 DNA glycosylase [MBD4], and 3-methyladenine-DNA glycosylase [MPG]).[19,20] MMR, an evolutionarily conserved DNA repair process, can be initiated by the recognition of small base mismatches by MutSα, a heterodimer of the MSH2 and MSH6 proteins. Additionally, the MutSα-DNA complex further recruits repair machineries such as endonucleases, DNA polymerase, ligase, and others.[21]

Received 17 June 2021; accepted 28 November 2021;
https://doi.org/10.1016/j.omtn.2021.11.023.

[4]Present address: Stem Cell Convergence Research Center, Korea Research Institute of Bioscience & Biotechnology (KRIBB), Daejeon 34141, Republic of Korea

[5]These authors contributed equally

Correspondence: Sangsu Bae, PhD, Department of Chemistry, Hanyang University, 222, Wangsimni-ro, Seongdong-gu, Seoul 04763, Republic of Korea.
E-mail: sangsubae@hanyang.ac.kr
Correspondence: Hyuk-Jin Cha, PhD, College of Pharmacy, Seoul National University, 1, Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea.
E-mail: hjcha93@snu.ac.kr

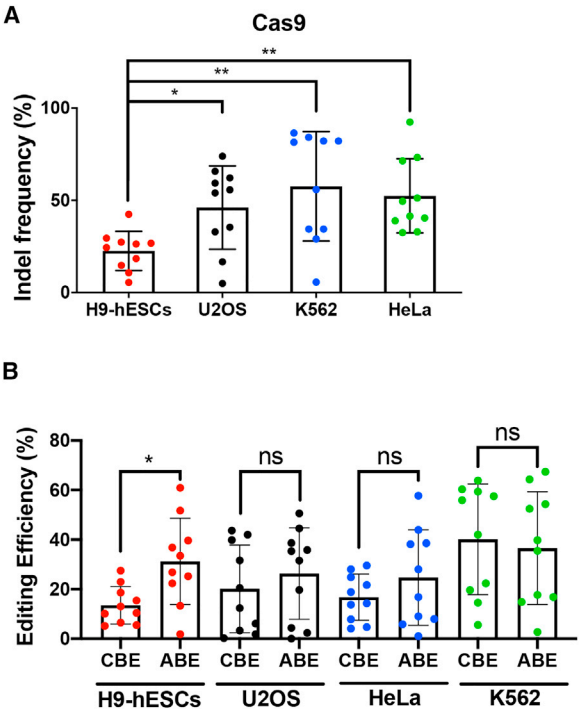


Figure 1. ABE-mediated editing efficiency is distinctively higher than that of CBE in hESCs

(A) Comparison of indel frequencies in hESC, U2OS, K562, and HeLa cells after Cas9-encoding plasmid delivery at 10 genomic sites. Dots represent the mean value of independent biological replicates for each target. Bars represent mean value, and error bars represent the SEM of the dots. *p < 0.05, **p < 0.01, comparing with the H9-hESCs by paired t test. (B) Comparison of base editing efficiencies in hESC, U2OS, K562, and HeLa cells after CBE- and ABE-encoding plasmid delivery at 10 genomic sites. Dots represent the mean value of two independent biological replicates for each target. Bars represent mean value, and error bars represent the SEM of the dots. *p < 0.05, comparing efficiency CBE and ABE by paired t test.

Due to the critical roles of embryonic stem cells (ESCs) in embryo development, a process from which all somatic cells in the body originate, ESCs have evolved molecular mechanisms to safeguard genome integrity and minimize spontaneous mutations.[6,22] Thus, DNA repair systems such as BER, MMR, and others are highly active in hPSCs,[23] which renders unique cellular characteristics to hPSCs but also affects the outcome of genome editing. For example, hPSCs are distinctively susceptible to p53-dependent cell death, which results in low Cas9-mediated gene editing efficacy in these cells.[7] However, to the best of our knowledge, our study is the first to comprehensively analyze gene editing outcomes with CBE and ABE in hPSCs against somatic cell lines.

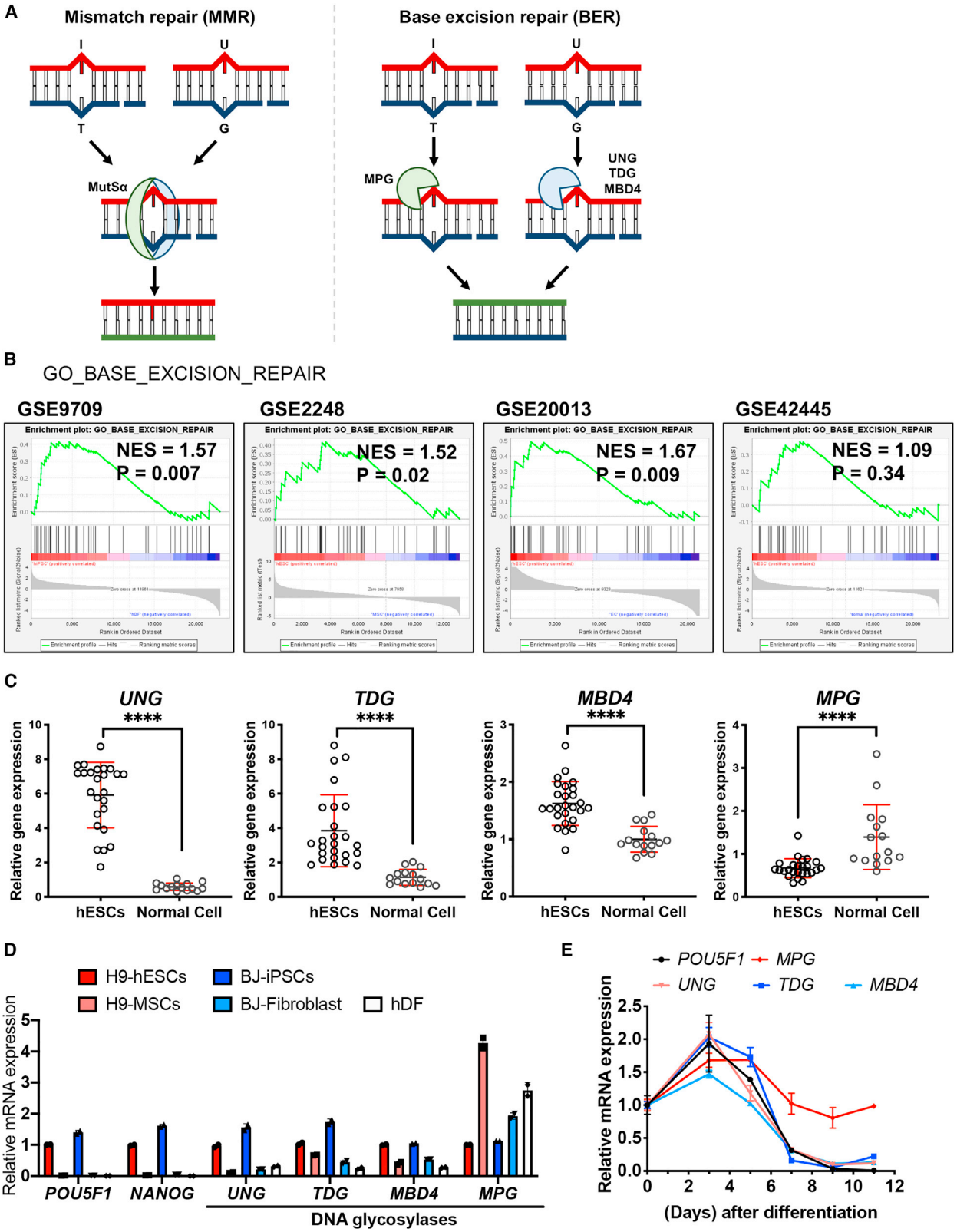
Here, we conducted high-throughput sequencing analysis for tens of endogenous targets in hPSCs as well as in somatic cancer cell lines. Interestingly, the editing efficiency of CBE was lower than that of ABE only in hPSCs, whereas CBEs and ABEs exhibited similar levels of editing efficiencies in several somatic cancer cell lines. Further, high expression levels of DNA glycosylases including UNG, which are responsible for elevated BER activity, prevailed in undifferentiated hPSCs. The genetic perturbation of DNA glycosylases indicated that the relatively low editing activity of CBE compared to ABE in human ESCs (hESCs) was likely attributable to a high UNG expression. Therefore, simple transient depletion of UNG overall increased the editing activity and product purity of CBEs, thus providing a precise means to modulate C to T transition for future disease modeling and *ex vivo* gene correction in hPSCs.

RESULTS

ABE-mediated editing efficiency is distinctively higher than that of CBE in hESCs

We first examined gene editing efficiencies mediated by Cas9 nucleases, CBEs, and ABEs in H9-hESCs, as well as in three different somatic cancer cell lines (HeLa, U2OS, and K562). To this end, we selected ten endogenous target sites for each Cas9 nuclease, CBE, and ABE, where gene editing efficiencies between ABE and CBE were comparable in somatic cell lines[17,18,24,25] and their perturbation gave minimal effect on pluripotency of hESCs.[10] Accordingly, three different plasmids encoding Cas9 from *Streptococcus pyogenes* (SpCas9), as well as AncBE4max (with two UNG inhibitor [UGI] motifs, hereafter BE4) for CBE and ABEmax for ABE were prepared. Three days after the transfection of each genome editing tool, the genomic DNAs were subjected to high-throughput sequencing to evaluate the editing outcomes from bulk populations. For Cas9 nucleases, our results indicated that H9-hESCs exhibited relatively lower insertion and deletion (indel) rates (average 20%) than other cell lines (50% for HeLa, 50% for U2OS, and 50% for K562) (Figure 1A). Consistent with the previous study reporting that Cas9 showed low activity in hPSCs due to p53-dependent cell death during gene editing,[7] Cas9 induced mRNA expression of *PPM1D*, encoding wild-type p53-induced phosphatase1 (WIP1), a well-characterized phosphatase for p53,[26] as well as H2AX,[27] that was markedly induced by Nutlin3 treatment, the MDM2 inhibitor to stabilize p53 (Figure S1A). Intriguingly, both *PPM1D* and *MDM2*, p53 downstream genes, were drastically induced by inhibition of apoptosis with pan-caspase inhibitor (z-VAD) (Figures S1A and S1B), implying that hESCs underwent Cas9-induced cell death. Considering the extremely high susceptibility of hESCs to p53-dependent cell death,[4,5] these results may account for the lower Cas9 efficiency of hESCs compared to the other somatic cancer cell lines, consistent with the previous report[7] (Figure 1A).

Particularly, for CBEs and ABE, we found that the editing activity of ABE was substantially higher than that of CBEs in H9-hESCs (Figures 1B, S1C, and S1D) under a comparable mRNA level of ABE and CBE (Figure S1E), whereas ABEs and CBEs exhibited comparable editing efficiencies in all other cell types. Although the gene editing efficiencies of CBEs and ABEs varied depending on the target sequences, the average activities between CBEs and ABEs were similar in the somatic cancer cell lines. To confirm our findings (i.e., the skewed editing efficiency of CBE in hESCs), we further tested another hESC line,



(legend on next page)

hCHA3, and observed similar distinctive characteristics in these cells (Figure S1F).

Similar to that of Cas9,[7] relevance of the skewed editing efficiencies of CBE compared with those of ABE in hESCs was next examined, in that p53-dependent cell death would be critical for determining editing efficiency in hPSCs. For establishment of TP53 knockout (TP53 KO) hESCs, sgRNA targeting exon 4 of TP53 (gTP53) was designed (Figure S2A). TP53KO hESCs were readily established by simply selecting a surviving colony after introduction of Cas9 and gTP53, followed by Nutlin3 treatment (Figures S2B and S2C), with approximately 100% indel (Figure S2D). As predicted, p53 response, determined by *MDM2* expression, disappeared in TP53KO hESCs by either Nutlin3 or Cas9, unlike that of wild-type hESCs (Figure S2E). While cell death by introduction of Cas9, ABE, and CBE was evidently attenuated (Figure S2F), editing efficiencies of Cas9, ABE, and CBE were significantly improved in TP53 KO hESCs (Figure S2G). These results imply that p53 response would be less relevant to skewed editing efficiency of CBE in hESCs, as shown in Figure 1B.

Distinct expression pattern of DNA glycosylases in hPSCs

We next sought to determine why CBE exhibited lower editing efficiency than ABE in hESCs other than p53. Given the high activity of DNA repair mechanisms in hPSCs due to the high expression of repair genes in BER or MMR compared to differentiated cells, base mismatches formed by BEs would be readily repaired by either BER or MMR. However, MMR is initiated by the recognition of mismatch bases by the MSH protein complex (i.e., MutSα). Of note, it has been reported that base mutations in purine-pyrimidine and purine-purine are equivalently well-repaired by MMR, unlike those in pyrimidine-pyrimidine.[28] As base mutations (e.g., I:T and G:U) that occurred by both ABE and CBE, respectively, belong to purine-pyrimidine, it is likely that MMR similarly affects both ABE and CBE (Figure 2A). Therefore, the MMR did not account for the skewed editing efficiency of CBE compared to ABE in hPSCs and was thus ruled out as the mechanism that mediated the aforementioned discrepancy. Instead, our downstream experiments focused on BER, which is initiated by base-specific DNA glycosylases (Figure 2A).

Consistent with previous reports,[23,29,30] several gene ontology terms associated not only with *DNA repair* but also *base excision repair* were highly enriched in undifferentiated hPSCs (Figures 2B and S3A) based on multiple datasets (Figure S3B). The common differentially expressed genes from these datasets were then visualized in KEGG pathway maps (https://www.genome.jp/). This analysis indicated that a large number of BER-associated genes (Figure S3C) were highly expressed in undifferentiated hPSCs. Interestingly, the KEGG pathway analysis of this dataset indicated that the DNA glycosylases that are responsible for the removal of base lesions by CBE (e.g., *UNG*, *TDG*, and *MBD4*) were upregulated, whereas *MPG*, which is known to remove inosine (I) from DNA, was downregulated in hPSCs compared with their differentiated counterparts (Figure S3C). To further generalize the unique expressions of DNA glycosylases, we took advantage of a transcriptome database of cell lines (http://nextbio.com)[27] and compared the expressions of these DNA glycosylases between 25 hESCs and 15 normal cell lines (Table S1) and identified distinct expression patterns in hESCs (Figure 2C). Similar to the expression profile from the dataset, DNA glycosylases such as *UNG*, *TDG*, and *MBD4* were highly expressed along with *POU5F1* in two independent hPSC models (H9-hESCs and BJ-iPSCs) compared with their differentiated counterparts (mesenchymal stem cells derived from H9-hESCs [hESC-MSCs][31,32] and BJ fibroblast, a parent fibroblast of BJ-iPSCs) (Figure 2D). In contrast, MPG expression was higher in the differentiated cells than in hPSCs (Figure 2D). As predicted, the expression of DNA glycosylases for uracil (e.g., UNG, TDG, and MBD4) was markedly reduced during spontaneous differentiation, whereas MPG exhibited a distinct expression pattern (Figure 2E). Consistent with these observations, the catalytic activity of UNG was significantly higher in hESCs compared to hESC-MSCs (Figure S3D).

Marginal effect of DNA glycosylases on ABE outcomes in hESCs

Due to the high expression of TDG and MBD4 instead of MPG in hPSCs, we first hypothesized that the I:T mismatch caused by ABE would be readily recognized by TDG and MBD4, which preferably recognize G:T mismatches (Figure 3A) due to the structural similarity between inosine and guanine (Figure 3A, inserted panel). However, contrary to our expectation that the prompt excision of T from I:T mismatches by MBD4 and/or TDG may cause higher A to G mutation rates in hESCs, perturbation of MBD4 and/or TDG expression failed to significantly affect ABE outcomes in multiple targets (Figures 3B and 3C). Next, we hypothesized that the relatively low expression of MPG (Figure 2C) might affect ABE efficiency. Unexpectedly, ectopic expression of MPG (Figure S4) had only a marginal effect on A to G editing efficiency in multiple targets (Figure 3D), and no significant differences were observed between conditions (Figure 3E). Therefore, ABE outcome was less affected by the expression level of DNA glycosylases in hESCs.

Figure 2. Distinct expression pattern of DNA glycosylases in hPSCs

(A) Brief scheme of mismatch repair and base excision repair pathway. Deaminated bases are colored in red, and newly synthesized DNA strand through DNA repair pathway is colored in green. (B) Gene set enrichment analysis for gene ontology of base excision repair pathway from indicated gene sets. (C) Expression of UNG, TDG, MBD4, and MPG in several hESCs cell lines from NextBio portal. Normal cells were analyzed as controls. Bars represent mean values, and error bars represent the SD. ****$p < 0.0001$, comparing with the hESCs by unpaired t test. (D) mRNA expression of POU5F1, NANOG, UNG, TDG, MBD4, and MPG of hESC-derived MSCs (H9-MSCs), hESCs (H9-ESCs), human fibroblast (BJ fibroblast), BJ-fibroblast-induced iPSCs (BJ-iPSCs), and human dermal fibroblasts (hDF). Bars represent mean values, and error bars represent the SD (n = 2). (E) mRNA expression of POU5F1, UNG, TDG, MBD4, and MPG from H9-hESCs at indicative days after differentiation. Dots represent mean values, and error bars represent the SD (n = 2).

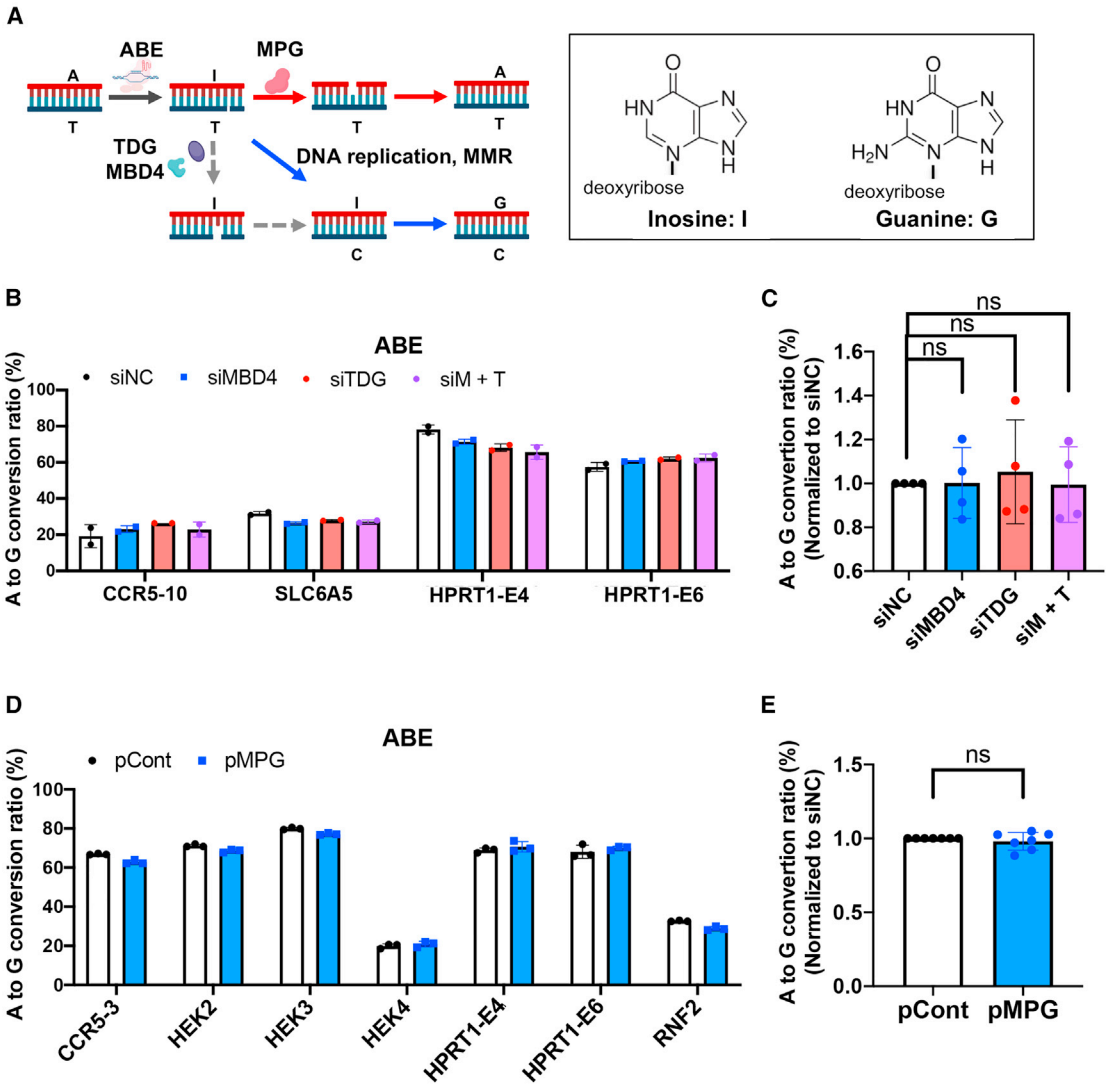
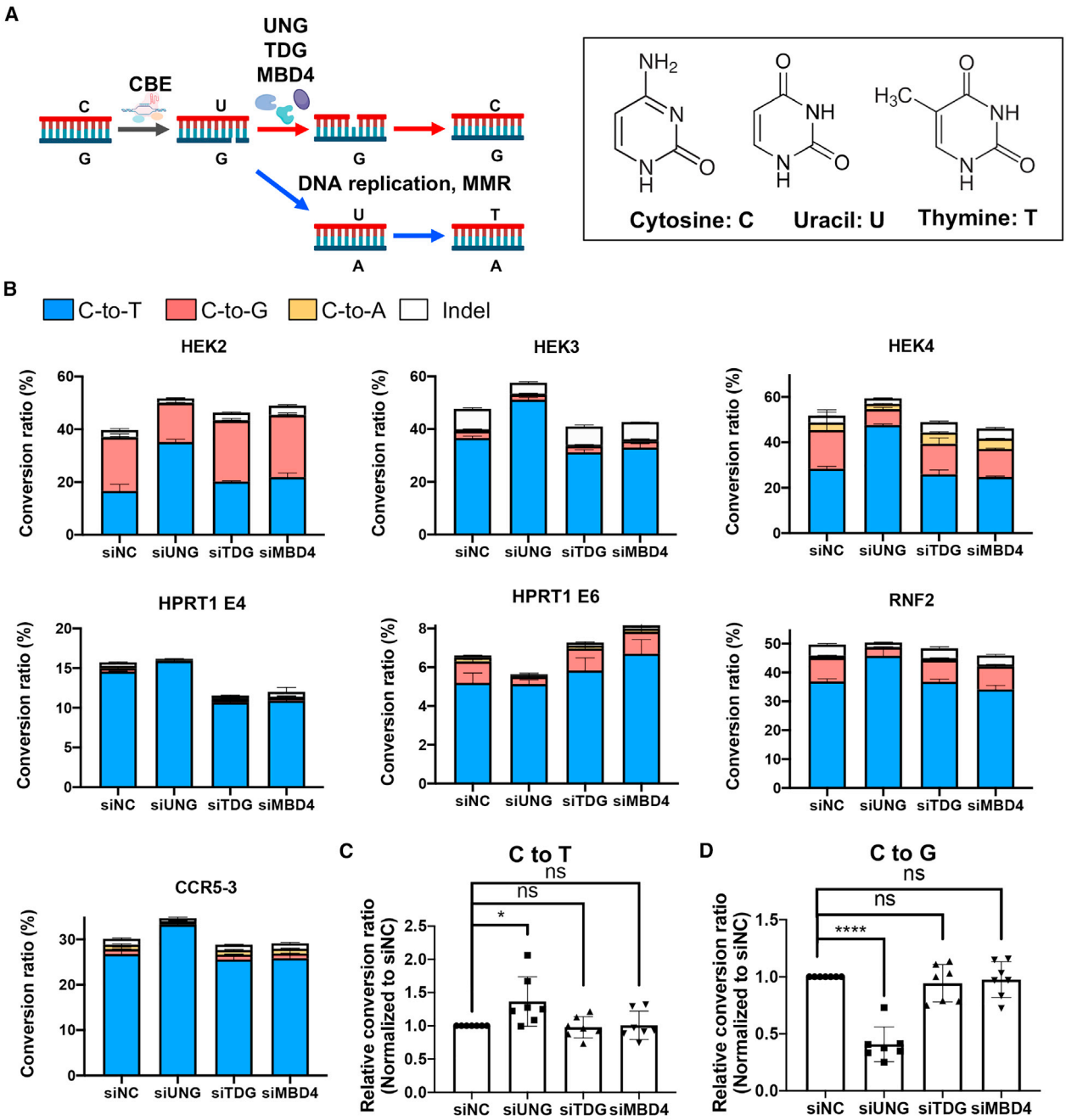


Figure 3. Marginal effect of DNA glycosylases on ABE outcomes in hESCs

(A) Scheme of ABE-mediated base editing process. (B) Base editing in H9-hESCs after ABE-encoding plasmid delivery with siRNA targeting MBD4 (siMBD4), TDG (siTDG) and both siRNA at the same time (siM + T) at four genomic sites. Bars represent mean values, and error bars represent the SD of two independent biological replicates. (C) Comparison of ABE efficiencies in H9-hESCs after ABE-encoding plasmid delivery with siRNA targeting DNA glycosylases at four genomic sites. The editing efficiency of ABE is normalized to non-targeting siRNA (siNC) delivered efficiency. Error bars represent the SEM of two independent biological replicates. (D) Base editing in H9-hESCs after ABE-encoding plasmid delivery with MPG expression (pMPG) or control plasmid (pCont) plasmid at seven genomic sites. Bars represent mean values, and error bars represent the SD of two independent biological replicates. (E) Comparison of ABE efficiencies in H9-hESCs after ABE-encoding plasmid delivery with MPG overexpression (pMPG) or control plasmid (pCont) at seven genomic sites. The editing efficiency was normalized to pCont. Bars represent mean values, and error bars represent the SD of three independent biological replicates.

UNG expression for efficiency and product purity of CBE in hESCs

Given that the expression level of DNA glycosylases was less associated with ABE (Figure 3), we next sought to determine whether this factor accounted for the skewed efficiency of CBE in hESCs. Given that uracil results from the deamination of cytosine by CBE, after which it is recognized and excised by UNG, TDG, or MBD4 (Figure 4A), depletion of one of these DNA glycosylases would alter the editing outcome of CBE. Unlike in ABE, knock down of UNG but not the other DNA glycosylases (Figure S5A) was likely to improve C to T editing in most of the tested targets in hESCs (Figure 4B). The overall improvement of C to T editing efficiency by depletion of UNG was statistically significant (Figure 4C). These results suggest that high UNG expression in hESCs likely explained the low CBE editing efficiency, resulting in skewed BE editing. Next, we examined the effect of transient depletion of DNA glycosylases in off-target editing



E

DNA glycosylase	Targetability				Main preference
	U in ssDNA	U in dsDNA	G:T mismatch	5-hmU in dsDNA	
UNG	Yes	Yes	No	No	G:U and A:U mismatch, U in ssDNA
TDG	No	Yes	Yes	Yes	G:T mismatch, 5hmU:G mismatch
MBD4	No	Yes	Yes	Yes	G:T and G:U mismatch in CpG context

(legend on next page)

of CBE by analyzing 12 off-target sites from CCR5-3, HEK2, and HEK3 (Table S2). Of note, no significant alterations of off-target editing by transient knockdown of DNA glycosylase were observed (Figures S5B–S5D). We also observed that depletion of UNG not only improved C to T efficiency but also product purity. The undesirable C to G and C to A conversion was reduced in multiple targets with different degrees by transient depletion of UNG in hESCs (Figures 4B and 4C). Therefore, the normalized product impurity (e.g., C to G [Figure 4D], C to A [Figure S5E], and indel [Figure S5F]) was significantly diminished by the simple knockdown of UNG but not any other DNA glycosylase. Furthermore, different expression level of UGI (by BE4 without UGI, BE4, and BE4 with additional ectopic expression of UGI) in CBE editing determined both editing efficiency and product purity (Figures S5G and S5H). We thus concluded that the distinct base editing pattern (e.g., lower CBE efficiency) in hESCs resulted from the high expression of UNG. Therefore, transient siRNA-mediated depletion of UNG would be beneficial for efficient C to T substitution of hESCs with CBE to further establish disease models or correct genetic mutations in hPSCs.

DISCUSSION

BER is an evolutionally conserved DNA damage repair mechanism that safeguards genome integrity even immediately after fertilization. BER is the primary repair system through which DNA damage is repaired in the developing zygote[33] and is also highly activated in ESCs compared to differentiated cells, serving as an important genome safeguard.[23,34]

Unlike in somatic cancer cell lines, we demonstrated that the editing efficiency of CBE was significantly lower than that of ABE in hESCs (Figure 1). Base-specific DNA glycosylases were distinctively expressed in hPSCs (Figure 2), which can account for the skewed efficiency of CBE compared to ABE in hPSCs. The temporary genetic perturbation of DNA glycosylases (UNG, TDG, MPG, and MBD4) indicated that the efficiency of CBE but not ABE (Figure 3) was determined by the high UNG expression in hPSCs (Figure 4). The fact that UNG depletion only affected CBE efficiency (Figure 4) suggests that the U:G mismatch produced by CBE was mostly repaired by UNG but not TDG and MBD4 in hESCs. UNG mostly recognizes uracil produced from the deamination of cytosine, whereas TDG preferentially recognizes 5-hydromethyl uracil (5hmU) or thymidine produced by 5hmC and 5-methyl cytosine (5mC) deamination (Figure 4E).[35] Likewise, MBD4 preferentially recognizes T:G and U:G mismatches in CpG regions (Figure 4E).[36] Given the preference of the rat APOBEC1 deaminase of CBE (BE4max) toward unmodified cytosine,[37,38] it is highly plausible that high UNG activity rather than TDG or MBD4 in hESCs competes with the activity of APOBEC1 toward unmodified cytosine despite the presence of UGI in BE4max.

In contrast, the genetic perturbations in DNA glycosylase expression observed herein had a marginal effect on ABE editing efficiency (Figure 3). Despite the structural similarity of inosine with guanine, temporary ectopic expression of MPG, which is responsible for removing G from G:T mismatch, failed to enhance ABE editing efficiency in hPSCs (Figure 3D). These findings were consistent with previous results demonstrating that MPG knockout had a marginal effect on ABE editing efficiency and product purity.[18] Although several cytosine deaminases act on DNA in mammalian cells, adenosine deaminases in mammalian cells are only involved in nucleotide metabolism and RNA.[39] This is why adenosine deaminases in ABE are genetically engineered from the pre-existing adenosine deaminases acting from RNA,[18] which might explain why editing efficiency is less affected by natural MPG than UNG. We also examined the depletion of TDG and/or MBD4 during ABE treatment under the assumption that TDG and MBD4 would similarly recognize I:T and G:T mismatches. However, transition depletion of TDG and/or MBD4 showed marginal effects on ABE efficiency (Figure 3B), suggesting that TDG and MBD4 can discriminate between I:T and G:T mismatches.

The distinct DNA damage responses of hESCs may not only affect the editing outcome of CBE but also that of other types of DNA editing tools. For example, the recently developed glycosylase base editor (GBE) enables C to G transversion in eukaryotes by inducing apurinic (AP) sites via cytidine deaminase and UNG conjugated onto nCas9.[40] We speculate that hPSCs would likely exhibit distinct GBE editing outcomes compared with other somatic cell lines given that the AP sites produced by GBE can be recognized by MBD4,[41] which is highly expressed in hPSCs. Thus, for efficient genome editing in hPSCs either for disease modeling or *ex vivo* stem cell therapy, the unique characteristics of specific cell lines must be considered.

MATERIALS AND METHODS

Statistical analysis

The quantitative data are expressed as the mean values ± standard deviation (SD). Paired t tests and Student's unpaired t test were performed to analyze the statistical significance of gene editing efficiency and gene expression comparison, respectively, using the PRISM. For comparison of gene editing efficiency between cell lines or editing conditions (e.g., ABE versus CBE), paired t test was performed. For comparison of mRNA expression and live cell ratio, Student's unpaired t test was performed. Values less than 0.05 were considered statistically significant (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$ and ****$p < 0.0001$).

Figure 4. UNG expression for efficiency and product purity of CBE in hESCs

(A) Scheme of CBE-mediated base editing process. (B) Percentage of sequencing reads with specific edited nucleotides that have been converted from target Cs in hESCs after CBE-encoding plasmid delivery with siRNA targeting UNG (siUNG), TDG (siTDG), and MBD4 (siMBD4). Non-targeting siRNA (siNC) was used as control. Blue, C to T; red, C to G; yellow, C to A. Error bars represent the SEM of three independent biological replicates. (C and D) Comparison of C to T (C) and C to G (D) conversion efficiencies in hESC after CBE-encoding plasmid delivery with siRNA targeting DNA glycosylase at seven genomic sites. The conversion efficiency after CBE-encoding plasmid delivery with each DNA glycosylase-targeting siRNA delivery is normalized to the conversion efficiency after CBE-encoding plasmid delivery with siNC. (E) Table of targetability and main preference of UNG, TDG, and MBD4.

Plasmid construction

The plasmids in this study were provided from Addgene, including p3s-Cas9-HN (addgene no.104171), pCMV_ABEmax (addgene no.112095), pCMV-AncBE4max (addgene no. 112094), pUGI-NLS (addgene no.101091), and pRG2 (Addgene no. 104174). Sequences corresponding to sgRNAs were cloned into BsaI-digested pRG2 vector (Addgene no. 104174). For this step, oligos containing the spacer sequence were annealed to form double-stranded DNA fragments with compatible overhangs and ligated using T4 ligase (Enzynomics). To construct pCMV-AncBE4max without UGI, the c-terminal part of BE4max contacting 2× UGI was digested by Cas9 and AgeI endonuclease (NEB) and Gibson cloned using NEBuilder HiFi DNA Assembly master mix (NEB). All plasmids used for transfection experiments were prepared using a NucleoBond Xtra Midi Plus EF kit (MN).

Cell culture and transfection

HeLa (ATCC CCL-2) and U-2OS (ATCC HTB-96) cells were cultured in DMEM supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin (WELGENE). K562 cells (ATCC CRL-3343) were cultured in RPMI-1640 supplemented with 10% FBS and 1% penicillin-streptomycin (WELGENE). For Cas9-, ABE-, or CBE-mediated genome editing, Cas9-, ABE-, or CBE-encoding plasmids (0.5 μg) and sgRNA-encoding plasmids (0.17 μg) were mixed with cells (1.5×10^5) and electroporated via Neon Transfection System. H9 (WA09, WiCell Research Institute) and CHA3 hESCs were cultured on Matrigel (BD Biosciences) coated dishes fed with StemMACS media (Miltenyi-Biotec) added with 50 μg/mL Gentamicin (Gibco). For Matrigel coating, 200 μL of Matrigel was diluted in 16 mL of cold DMEM/F-12 media (Gibco). Diluted Matrigel was distributed to culture plate and incubated in a cell culture incubator for 1 h. For transfer, hESCs were rinsed with Dulbecco's phosphate buffered saline (DPBS) and detached with Accutase solution (561527, BD Biosciences). Detached cells were washed with DMEM/F-12 media for three times. Washed cells were resuspended with 1 mL of StemMACS media and plated on Matrigel-coated plate with StemMACS media added with 10 μM of Y27632 (Gibco). For transfection, hESCs were rinsed with DPBS and detached with Accutase solution (561527, BD Biosciences). Cells were washed with Opti-MEM (31985070, Gibco) three times and diluted to a concentration of 1×10^6 cells in 100 μL of Opti-MEM (31985070, Gibco). 2 μg of Cas9 or BE vectors (Cas9, BE4max, and ABEmax cloned in pCMV vector) and 2 μg of sgRNA vector were added to the cell mixture. For siRNA or overexpression vector, an additional 2 μg of siRNA or overexpression vector was additionally added to the cell mixture. Electroporation was performed by NEPA-21. Poring pulse was 175 V and transfer pulse was 2.5 mV.

Targeted deep sequencing

For analysis of editing efficiency, genomic DNA were extracted from Cas9-, ABE-, or CBE-transfected cells using a NucleoSpin Tissue kit (MN) at 3 days after transfection. Target sites were amplified using a KOD Multi & Epi PCR kit (TOYOBO) for sequencing library generation. These libraries were sequenced using MiniSeq with a TruSeq HT Dual Index system (Illumina) as previously described. Briefly, equal amounts of the PCR amplicons were subjected to paired-end read sequencing using an Illumina MiniSeq platform. After MiniSeq, paired-end reads were analyzed by comparing wild-type and mutant sequences using BE-analyzer. High-throughput sequencing data have been deposited in the NCBI Sequence Read Archive database (https://www.ncbi.nlm.nih.gov/sra) under accession number PRJNA787724.

Editing efficiency averaging and normalization

Average value of editing efficiency in each target was calculated via arithmetic mean of individual experiments. Average value of editing efficiency in each cell line was calculated via arithmetic mean of individual targets. Normalization of editing efficiency was conducted by dividing the average value of editing efficiency of each target from perturbation (e.g., siUNG or pMPG) by that of the control (e.g., siNC or pCont).

RT-qPCR analysis

For total RNA extraction, Easy-BLUETM RNA isolation kit (iNtRON Biotechnology) was used. RNA was extracted form cell pellets via Easy-BLUETM RNA isolation kit, following the supplier's instructions. cDNA was synthesized by PrimeScriptTM RT reagent kit (TaKaRa). 2 μL of PrimeScriptTM RT reagent kit was added to 500 μg of RNA samples in 8 μL of distilled water (DW) and reacted for 15 min at 37°C. Light Cycler-480®II (Loche) and SYBR® Green PCR reagents (Life Technologies) were used for quantitative real-time PCR analysis, following the supplier's instructions.

Generation of *TP53* knockout hPSCs

Cells were transfected with Cas9 and gTP53 vector with the same conditions as in the "cell culture and transfection" section. To enrich *TP53* KO cells, 10 μM of nutlin3 was treated for 48 h after 5 days from transfection.

Data availability

Source data are available from the corresponding authors upon request.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omtn.2021.11.023.

ACKNOWLEDGMENTS

This research was supported by the National Research Foundation of Korea (NRF) no. NRF-2021R1A2C3012908 (S.B.) and NRF-2020R1A2C2005914 (H.J.C.) and a grant from the Research of Korea Centers for Disease Control and Prevention no. 2020-ER6902-00 (H.J.C.).

AUTHOR CONTRIBUTIONS

H.J.C. and S.B conceived the overall study design and led the experiments. J.C.P. and H.K.J. mainly conducted the experiments, data analysis, and critical discussion of the results. J.H.H. and Y.J. contributed to NGS analysis and provided gene editing techniques. J.K. and K.T.K. contributed to hPSCs characterization.

DECLARATION OF INTERESTS

The authors declare no conflict of interest.

REFERENCES

1. Tiscornia, G., Vivas, E.L., and Izpisua Belmonte, J.C. (2011). Diseases in a dish: modeling human genetic disorders using induced pluripotent cells. Nat Med *17*, 1570–1576.
2. Hotta, A., and Yamanaka, S. (2015). From Genomics to Gene Therapy: Induced Pluripotent Stem Cells Meet Genome Editing. Annu Rev Genet *49*, 47–70.
3. Hockemeyer, D., and Jaenisch, R. (2016). Induced Pluripotent Stem Cells Meet Genome Editing. Cell Stem Cell *18*, 573–586.
4. Liu, J.C., Guan, X., Ryan, J.A., Rivera, A.G., Mock, C., Agrawal, V., Letai, A., Lerou, P.H., and Lahav, G. (2013). High mitochondrial priming sensitizes hESCs to DNA-damage-induced apoptosis. Cell Stem Cell *13*, 483–491.
5. Lee, M.O., Moon, S.H., Jeong, H.C., Yi, J.Y., Lee, T.H., Shim, S.H., Rhee, Y.H., Lee, S.H., Oh, S.J., Lee, M.Y., et al. (2013). Inhibition of pluripotent stem cell-derived teratoma formation by small molecules. Proc Natl Acad Sci U S A *110*, E3281–E3290.
6. Weissbein, U., Benvenisty, N., and Ben-David, U. (2014). Quality control: Genome maintenance in pluripotent stem cells. The Journal of cell biology *204*, 153–163.
7. 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.
8. Li, X.L., Li, G.H., Fu, J., Fu, Y.W., Zhang, L., Chen, W., Arakaki, C., Zhang, J.P., Wen, W., Zhao, M., et al. (2018). Highly efficient genome editing via CRISPR-Cas9 in human pluripotent stem cells is achieved by transient BCL-XL overexpression. Nucleic Acids Res *46*, 10195–10215.
9. Martin, R.M., Ikeda, K., Cromer, M.K., Uchida, N., Nishimura, T., Romano, R., Tong, A.J., Lemgart, V.T., Camarena, J., Pavel-Dinu, M., et al. (2019). Highly Efficient and Marker-free Genome Editing of Human Pluripotent Stem Cells by CRISPR-Cas9 RNP and AAV6 Donor-Mediated Homologous Recombination. Cell Stem Cell *24*, 821–828 e825.
10. 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.
11. Li, S., Akrap, N., Cerboni, S., Porritt, M.J., Wimberger, S., Lundin, A., Moller, C., Firth, M., Gordon, E., Lazovic, B., et al. (2021). Universal toxin-based selection for precise genome engineering in human cells. Nat Commun *12*, 497.
12. 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.
13. Wang, X., Ding, C., Yu, W., Wang, Y., He, S., Yang, B., Xiong, Y.C., Wei, J., Li, J., Liang, J., et al. (2020). Cas12a Base Editors Induce Efficient and Specific Editing with Low DNA Damage Response. Cell Rep *31*, 107723.
14. 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.
15. Park, J.-C., Kim, J., Jang, H.-K., Lee, S.-Y., Kim, K.-T., Park, S., Lee, H.S., Choi, H.-J., Park, S.-J., and Moon, S.-H. (2020). Efficient GNE myopathy disease modeling with mutation specific phenotypes in human pluripotent stem cells by base (bioRxiv).
16. 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.
17. 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.
18. Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I., and Liu, D.R. (2017). Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. Nature *551*, 464–471.
19. Krokan, H.E., and Bjoras, M. (2013). Base excision repair. Cold Spring Harb Perspect Biol *5*, a012583.
20. Krokan, H.E., Drablos, F., and Slupphaug, G. (2002). Uracil in DNA–occurrence, consequences and repair. Oncogene *21*, 8935–8948.
21. Li, G.M. (2008). Mechanisms and functions of DNA mismatch repair. Cell Res *18*, 85–98.
22. Stambrook, P.J. (2007). An ageing question: do embryonic stem cells protect their genomes? Mech Ageing Dev *128*, 31–35.
23. Fu, X., Cui, K., Yi, Q., Yu, L., and Xu, Y. (2017). DNA repair mechanisms in embryonic stem cells. Cell Mol Life Sci *74*, 487–493.
24. 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.
25. Cho, S.W., Kim, S., Kim, Y., Kweon, J., Kim, H.S., Bae, S., and Kim, J.S. (2014). Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome Res *24*, 132–141.
26. Lowe, J., Cha, H., Lee, M.O., Mazur, S.J., Appella, E., and Fornace, A.J., Jr. (2012). Regulation of the Wip1 phosphatase and its effects on the stress response. Front Biosci (Landmark Ed) *17*, 1480–1498.
27. Cha, H., Lowe, J.M., Li, H., Lee, J.S., Belova, G.I., Bulavin, D.V., and Fornace, A.J., Jr. (2010). Wip1 directly dephosphorylates gamma-H2AX and attenuates the DNA damage response. Cancer Res *70*, 4112–4122.
28. Modrich, P. (1997). Strand-specific mismatch repair in mammalian cells. J Biol Chem *272*, 24727–24730.
29. Tichy, E.D., Pillai, R., Deng, L., Liang, L., Tischfield, J., Schwemberger, S.J., Babcock, G.F., and Stambrook, P.J. (2010). Mouse embryonic stem cells, but not somatic cells, predominantly use homologous recombination to repair double-strand DNA breaks. Stem Cells Dev *19*, 1699–1711.
30. Tichy, E.D., Liang, L., Deng, L., Tischfield, J., Schwemberger, S., Babcock, G., and Stambrook, P.J. (2011). Mismatch and base excision repair proficiency in murine embryonic stem cells. DNA Repair (Amst) *10*, 445–451.
31. Kim, K.T., Jeong, H.C., Kim, C.Y., Kim, E.Y., Heo, S.H., Cho, S.J., Hong, K.S., and Cha, H.J. (2017). Intact wound repair activity of human mesenchymal stem cells after YM155 mediated selective ablation of undifferentiated human embryonic stem cells. J Dermatol Sci *86*, 123–131.
32. Hong, K.S., Bae, D., Choi, Y., Kang, S.W., Moon, S.H., Lee, H.T., and Chung, H.M. (2015). A porous membrane-mediated isolation of mesenchymal stem cells from human embryonic stem cells. Tissue Eng Part C Methods *21*, 322–329.
33. Khokhlova, E.V., Fesenko, Z.S., Sopova, J.V., and Leonova, E.I. (2020). Features of DNA Repair in the Early Stages of Mammalian Embryonic Development. Genes (Basel) *11*.
34. Mani, C., Reddy, P.H., and Palle, K. (2020). DNA repair fidelity in stem cell maintenance, health, and disease. Biochim Biophys Acta Mol Basis Dis *1866*, 165444.
35. Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A., Le Coz, M., Devarajan, K., Wessels, A., Soprano, D., et al. (2011). Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. Cell *146*, 67–79.
36. Petronzelli, F., Riccio, A., Markham, G.D., Seeholzer, S.H., Genuardi, M., Karbowski, M., Yeung, A.T., Matsumoto, Y., and Bellacosa, A. (2000). Investigation of the substrate spectrum of the human mismatch-specific DNA N-glycosylase MED1 (MBD4): fundamental role of the catalytic domain. J Cell Physiol *185*, 473–480.
37. Nabel, C.S., Jia, H., Ye, Y., Shen, L., Goldschmidt, H.L., Stivers, J.T., Zhang, Y., and Kohli, R.M. (2012). AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. Nat Chem Biol *8*, 751–758.
38. Wang, X., Li, J., Wang, Y., Yang, B., Wei, J., Wu, J., Wang, R., Huang, X., Chen, J., and Yang, L. (2018). Efficient base editing in methylated regions with a human APOBEC3A-Cas9 fusion. Nat Biotechnol *36*, 946–949.
39. Samuel, C.E. (2020). Adenosine Deaminases. Reference Module in Life Sciences (Elsevier).
40. Zhao, D., Li, J., Li, S., Xin, X., Hu, M., Price, M.A., Rosser, S.J., Bi, C., and Zhang, X. (2021). Glycosylase base editors enable C-to-A and C-to-G base changes. Nat Biotechnol *39*, 35–40.
41. Morera, S., Grin, I., Vigouroux, A., Couve, S., Henriot, V., Saparbaev, M., and Ishchenko, A.A. (2012). Biochemical and structural characterization of the glycosylase domain of MBD4 bound to thymine and 5-hydroxymethyluracil-containing DNA. Nucleic Acids Res *40*, 9917–9926.