Molecular Therapy
Nucleic Acids
Original Article

MutSα and MutSβ as size-dependent cellular determinants for prime editing in human embryonic stem cells

Ju-Chan Park,[1] Yun-Jeong Kim,[1] Jun Hee Han,[2] Dayeon Kim,[1] Mihn Jeong Park,[1] Jumee Kim,[1] Hyeon-Ki Jang,[3] Sangsu Bae,[4] and Hyuk-Jin Cha[1,5]

[1]College of Pharmacy, Seoul National University, Seoul, Republic of Korea; [2]Department of Chemistry, Hanyang University, Seoul, Republic of Korea; [3]Division of Chemical Engineering and Bioengineering, College of Art Culture and Engineering, Kangwon National University, Chuncheon, South Korea; [4]College of Medicine, Seoul National University, Seoul, Republic of Korea; [5]Research Institute of Pharmaceutical Sciences, Seoul National University, Seoul, Republic of Korea

Precise genome editing in human pluripotent stem cells (hPSCs) has potential applications in isogenic disease modeling and *ex vivo* stem cell therapy, necessitating diverse genome editing tools. However, unlike differentiated somatic cells, hPSCs have unique cellular properties that maintain genome integrity, which largely determine the overall efficiency of an editing tool. Considering the high demand for prime editors (PEs), it is imperative to characterize the key molecular determinants of PE outcomes in hPSCs. Through homozygous knockout (KO) of MMR pathway key proteins MSH2, MSH3, and MSH6, we reveal that MutSα and MutSβ determine PE efficiency in an editing size-dependent manner. Notably, *MSH2* perturbation disrupted both MutSα and MutSβ complexes, dramatically escalating PE efficiency from base mispair to 10 bases, up to 50 folds. Similarly, impaired MutSα by *MSH6* KO improved editing efficiency from single to three base pairs, while defective MutSβ by *MSH3* KO heightened efficiency from three to 10 base pairs. Thus, the size-dependent effect of MutSα and MutSβ on prime editing implies that MMR is a vital PE efficiency determinant in hPSCs and highlights the distinct roles of MutSα and MutSβ in its outcome.

INTRODUCTION

With continuous advancements in the clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 technology, patient-derived induced pluripotent stem cells (iPSCs) have been recognized for their potential in isogenic disease modeling[1,2] and *ex vivo* autologous cell therapy following pathogenic mutation correction.[3] As most pathogenic genetic variants found in humans are point mutations (58%) and deletions (25%),[4] precise genome editing in hPSCs to correct these genetic variations becomes crucial for clinical applications. Thus, various Cas9-based genome editing technologies, including homology-directed repair (HDR)-based knockin,[5] base editors (BEs),[6] and prime editors (PEs),[7] have been applied in human pluripotent stem cells (hPSCs) and optimized to improve efficiency.[8–10]

Unlike knockout (KO) by insertion/deletion (indel) introduction, the Cas9 enzyme with supplement of donor DNA conducts the desirable precise genome editing via HDR.[11] However, Cas9's endonuclease activity, inducing double-strand breaks (DSBs), causes extensive deletion and complex chromosomal rearrangements, resulting in unexpected DNA changes and pathogenic consequences.[12–14] Alternatively, BEs achieve base transitions without DSB through deaminase conjugation to nickase Cas9 (nCas9) or catalytically dead Cas9 (dCas9).[15,16] Thus, capability of single base substitution, with no large deletion,[17] renders BEs beneficial and secure tools for correcting hPSCs' pathogenic mutations. However, the requirement of a protospacer adjacent motif (PAM) sequence and the inability of transversion substitution restricts the editing coverage of pathogenic variants with BEs.[6] On the other hand, a PE enables highly flexible genome editing, such as indels and base substitutions, through synthesizing a DNA strand into a target site via reverse transcriptase (RT) and an engineered PE guide RNA (pegRNA).[18] Due to a PE's broader coverage (i.e., transition/transversion and short insertion/deletion correcting approximately 89% of pathogenic mutations[18]) than that of BEs, extensive studies have focused on improving PE editing efficiency.[10,19,20]

Embryonic stem cells, which give rise to somatic cells during embryogenesis, have evolved molecular mechanisms in maintaining genome integrity, relying on a high susceptibility to DNA damage and active DNA repair pathways.[21–23] These unique characteristics are major determinants for genome editing tool's outcomes in hESCs and iPSCs. For instance, the relatively low efficiency of Cas9-mediated genome editing in hPSCs[24] was recently explained by a high p53-dependent cell death upon Cas9-induced DSB.[25] Additionally, the overall lower efficiency of cytosine base editors (CBEs) compared with an adenine base editor in hPSCs results from active base excision repair (BER) through elevated uracil DNA glycosylase expression.[26] Despite recent examinations of PEs in hPSCs,[7,27,28] little is known about hPSCs' unique cellular properties in PE outcomes.

Received 4 October 2022; accepted 10 May 2023;
https://doi.org/10.1016/j.omtn.2023.05.015.

Correspondence: Hyuk-Jin Cha, College of Pharmacy, Seoul National University, Seoul, Republic of Korea.
E-mail: hjcha93@snu.ac.kr

DNA mismatch repair (MMR) is a highly conserved biological pathway that plays a crucial role in maintaining replication fidelity through multiple processes, including mismatch recognition, mismatch removal, and repair DNA synthesis.[29] DNA mismatch is recognized by mismatch repair complexes MutSα (MSH2-MSH6 subunits) and MutSβ (MSH2-MSH3 subunits).[29] MutSα initiates base mispairs (e.g., G-U mispair from cytosine deamination) and small insertion/deletion loops (IDLs) of one or two extrahelical nucleotides to cover most replication error repairs.[30,31] In contrast, MutSβ governs most large IDLs repair[29] up to 13 nucleotides long.[32]

Herein, we demonstrated using homozygous *MSH2* (MutS homology 2), *MSH3* (MutS homology 3), or *MSH6* (MutS homology 6) KO hESCs that MutSα and MutSβ are cellular determinants for PE efficiency according to editing sizes. Considering the critical role of MMR in regulating the genomic integrity of hPSCs, transient modulation of MutSα or MutSβ components is a potential strategy to improve prime editing of hPSCs, depending on the desired editing size.

RESULTS

MMR pathway's high enrichment in hPSCs

As briefly summarized in Figure 1A, nCas9 in a PE produces a single-strand break (SSB) near the PAM sequence (a). Next, RT synthesizes edited DNA (3′ flap containing "edit") based on the RT template sequence of pegRNA (b). The 3′ flap is annealed to the non-edited strand (c), producing a 5′ flap (lacking "edit") through an equilibrium process. Following the DNA damage repair machinery's removal of the liberated 5′ flap and ligation (d), the intended edit is incorporated into the target DNA sequence (e). The 5′ flap excised intermediate product is susceptible to MMR, resulting in newly synthesized sequence removal (f) (Figure 1A). This implies that MMR activity that varies in a cell type-specific manner[33] determines PE efficiency. Recent pooled CRISPRi screening consistently reveals that repression of major components of MMR, including *MSH2, MSH6, PMS2* or *MLH1*, enhances PE efficiency.[10] Accordingly, the improved PE version (i.e., PE4 and PE5) was made through the transient modulation of MMR activity.[10] As CBE efficiency was determined by active BER,[26] we hypothesized that hPSCs' unique cellular characteristics with a highly activated MMR may affect PE efficiency. As expected, MMR-associated gene sets were exceptionally enriched in undifferentiated hPSCs compared with the differentiated counterparts consistent in all datasets (i.e., GEO: GSE9709, GSE2248, GSE16963, and GSE42445) (Figures 1B and 1C).

Distinctive MSH2 and MSH6 gene expression levels in hPSCs

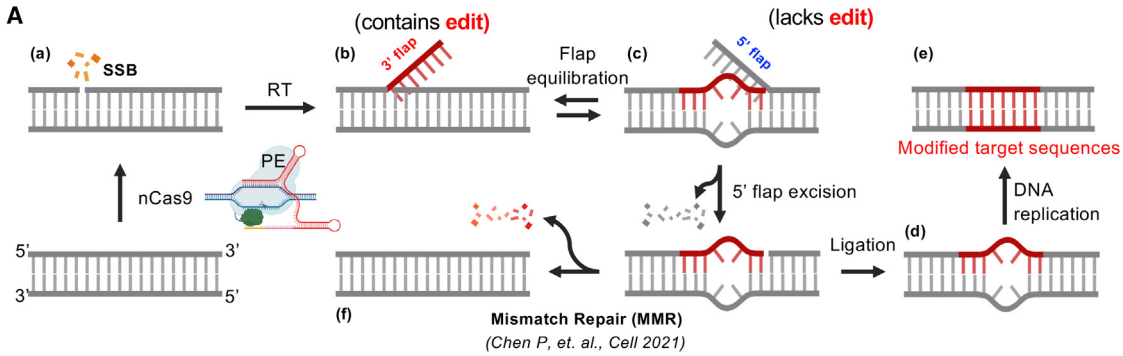
To illustrate significant MMR activity components in hPSCs, undifferentiated hPSC upregulated genes were depicted in the KEGG pathway map. *MSH2* and *MSH6* were highly expressed in undifferentiated hPSCs (Figures 2A–2D) among known cellular PE efficiency determinants (e.g., *MLH1*, *PMS2*, *MSH2*, and *MSH6*).[10] *MSH2's* and *MSH6's* high expression levels of 25 different hESCs lines compared with 14 normal cells were again validated by the NextBio dataset (http://nextbio.com) (see Table S1 for a complete cell line list). As predicted by hPSC transcriptome datasets (Figures 2A–2E), *MSH2* and *MSH6* gene expressions were highly expressed in hESCs (of which *SOX2* and *LIN28* expression levels determined pluripotency) compared with human mesenchymal stem cells derived from hESCs (hESC-MSCs)[34] and human dermal fibroblasts (hDFs) (Figure 2F). Furthermore, MSH2 and MSH6 protein levels were also heightened in undifferentiated hESCs (Undiff.) compared with the spontaneously differentiated cells from hESCs (Diff.) (Figure 2G). In addition, *MSH2* and *MSH6* mRNA levels were markedly repressed during spontaneous hESC differentiation (Figure 2H). *MSH2* and *MSH6* unique expression patterns were also substantiated in iPSCs (e.g., BJ-iPSCs) compared with BJ-fibroblast, an iPSC parent somatic cell[35] (Figure 2I). Considering the correlation between MMR component protein expression levels and overall MMR activities in cell lines,[33] high MSH2 and MSH6 expression in hPSCs would indicate elevated MMR activity.

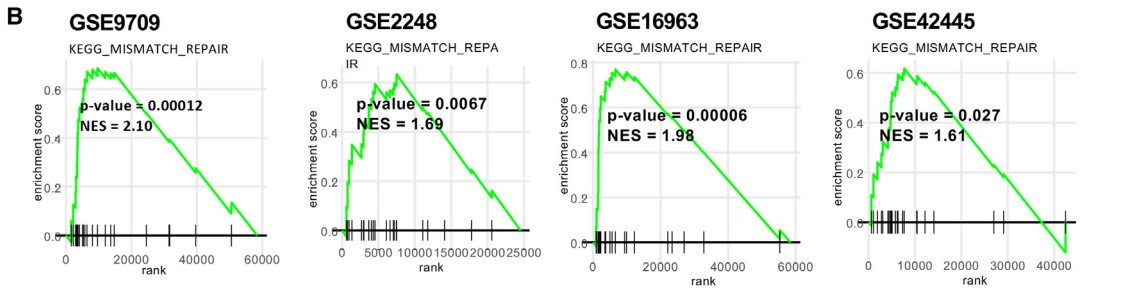
MSH2, MSH3, and MSH6 genetic perturbation in hESCs with CRISPR-Cas9

MSH2, MSH3, and *MSH6* were each genetically perturbed through CRISPR-Cas9 (hereafter Cas9) to examine their effects of MMR components on PE efficiency in hPSCs. To improve Cas9-mediated editing outcomes, un-edited hPSCs were selectively eliminated by YM155 treatment, a small molecule used to selectively kill undifferentiated hPSCs[36] through SLC35F2-dependent cellular import,[37] to account for undifferentiated hPSCs selective cell death.[9] The YM155 enriched selection (YES) procedure was performed as previously demonstrated[9] for efficient gene of interest knockout (Figure S1A). This approach successfully enriches edited hESCs and establishes homozygous *MSH2*, *MSH3*, and *MSH6* knockout (M2KO, M3KO, and M6KO respectively), along with three corresponding SLC35F2-KO control hESCs (Cont-2, Cont-3, and Cont-6 respectively) (Figures 3A, S1B, and S1C). Complete loss of MSH2 (Figure 3B), MSH3 (Figure 3C), or MSH6 (Figure 3D) in each KO hESC was corroborated through immunoblotting analysis. Notably, both MSH3 and MSH6 proteins were drastically reduced in M2KO hESCs (Figure 3B). This finding was consistent with data from previous studies demonstrating that MSH3 and MSH6 are rapidly destabilized by MSH2 absence.[38,39] Typical pluripotency gene (i.e., *POU5F1* and *NANOG*) (Figure 3E) and protein (i.e., OCT4) (Figure 3F) expression levels revealed that *MSH2, MSH3,* or *MSH6* depletion only marginally affected hESC pluripotency. Typical colonial morphology and alkaline phosphatase activity demonstrated that hESCs lacking MSH2, MSH6, or MSH6 (along with SLC35F2 KO) maintained pluripotency similar to wild type (WT) (Figure 3G).

Editing size discrimination by MSH2, MSH3, and MSH6 in prime editing

As previously mentioned, MSH2/MSH6- and MSH2/MSH3-composed MutSα and MutSβ by forming a complex with MLH1/PMS2 govern size-dependent DNA errors (from a base mispair to large IDLs, respectively). MutSα recognizes a base mismatch and small insertion-deletion loops of 1–2 base pairs (bps). Conversely, MutSβ preferentially recognizes most IDLs up to 15 bps (Figure 4A).[29,40] Considering MutSα′s and MutSβ′s distinct roles in base mispair and IDL (with different sizes) recognition,[29] different PE outcomes after *MSH2, MSH3*, and *MSH6 KO* were predicted (Figure 4B).





C

GEO study	Cell lines	Sample size	PMID
GSE9709	human iPSCs from neonetal skin derived cells(hiPSCs)	3	19383391
	human neonatal dermal fibroblast(hDF)	2	
GSE2248	hESCs	3	15971941
	Mesenchymal stem cells(MSCs)	1	
GSE16963	iPSC of Human Third Molars	4	20595386
	human Third Molas cell line	2	
GSE42245	hESCs	10	24259714
	Somatic cells	6	

Figure 1. MMR pathway's high enrichment in hPSCs

(A) Graphical illustration of prime editing process. (B) Gene set enrichment analysis (GSEA) for gene ontology of mismatch repair pathway from indicated gene sets. (C) Detailed information of gene sets used for GSEA and KEGG analysis.

We examined the PE efficiency of multiple differently sized editing (i.e., base mispair, small IDLs: 1–3 bps; large IDLs: 5–10 bps; larger IDLs: 15–34 bps) in each established KO hESC (Figure S2A). As predicted, MutSα abrogation by either M2KO or M6KO markedly improved base substitution (through base mispair) (Figures 4C) and insertion or deletion of 1 bp (e.g., 1-bp IDL) (Figures 4D and S3B). On the other hand, a PE's insertion of 3 bps (e.g., 3-bp IDL) increased in M2KO, M3KO, and M6KO (Figures 4E and S3C), while large IDLs, such as deletions of 5 and 10 bps, were only affected in M2KO and M3KO (Figures 4F and S3D). It is noteworthy that M2KO led to MutSα- and MutSβ-dependent MMR improved PE efficiency from base mispair to 10-bp IDLs (Figures 4C–4F). However, IDLs over 15 bps (Figure S2A) remained unaffected by MutSα or MutSβ loss (Figures 4G and S3E), consistent with larger IDLs over 15 nucleotides failing to be recognized by either MutSα or MutSβ.[41,42]

Notably, average Cas9 indel frequencies on HEK2 and HEK4 sites remained comparable in M2KO, M3KO, and M6KO hESCs compared with control hESCs (Figure S2F). Additionally, MSH2 reintroduction in M2KO markedly lowered G to T substitution (Figure S2G). These

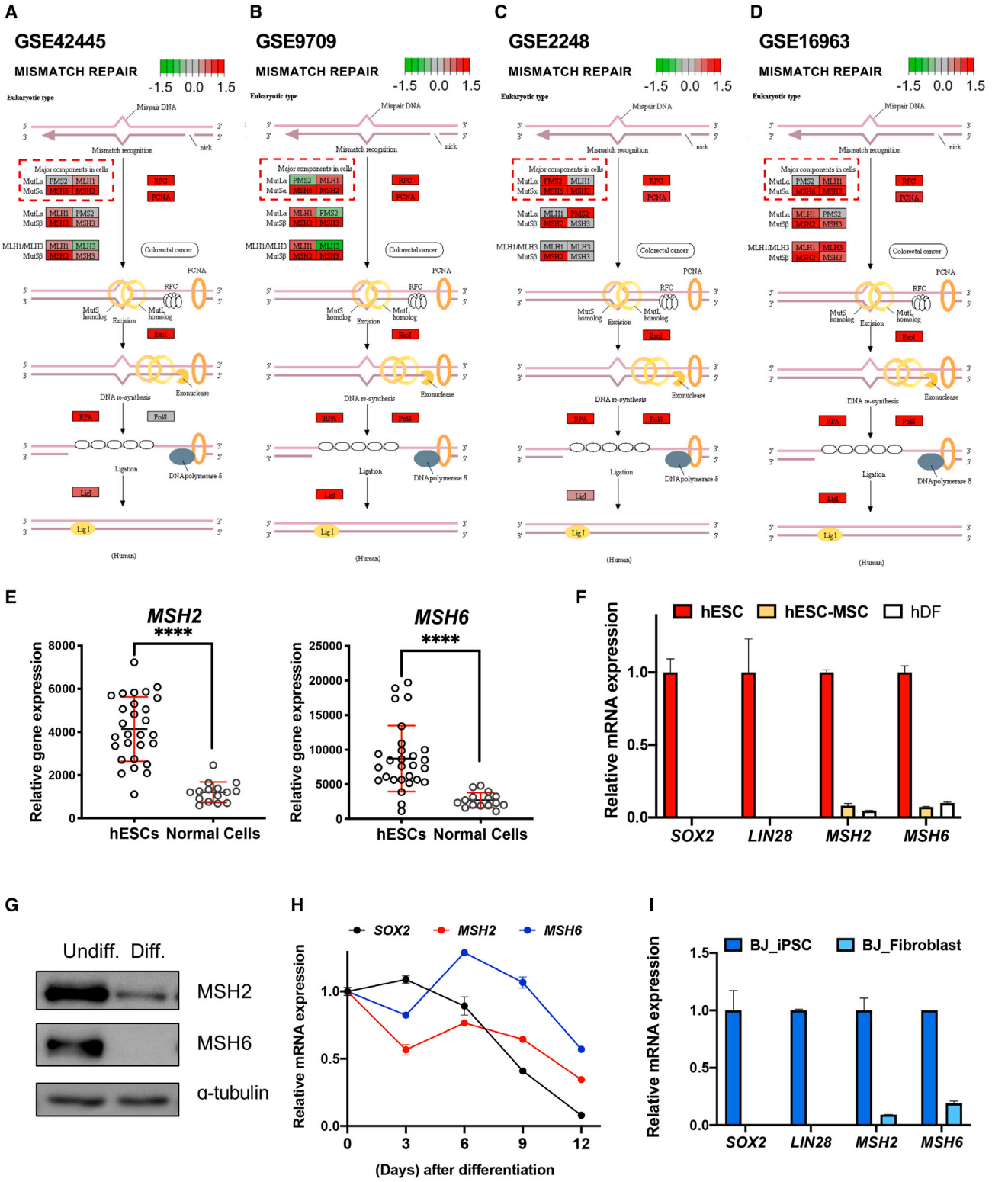


Figure 2. Distinctive MSH2 and MSH6 genes expression levels in hPSCs

(A–D) KEGG pathway analysis of eukaryotic mismatch repair from indicated gene sets; highly expressed genes are colored in red, while lowly expressed genes are colored in green. (E) Expression of MSH2 and MSH6 in multiple hESC cells and normal cells from NextBio portal. Bars represent mean values, and error bars represent the SD,

(legend continued on next page)

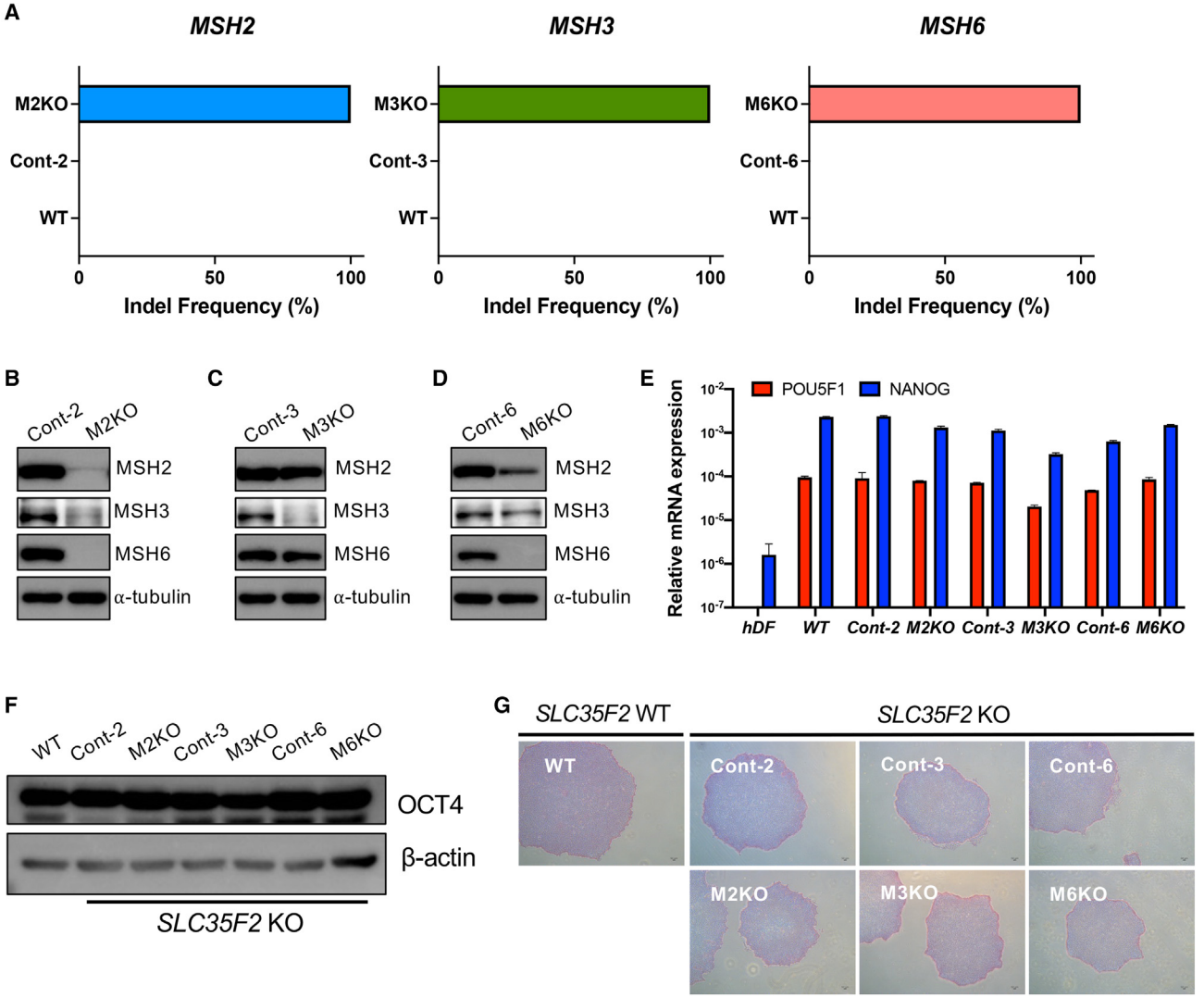


Figure 3. MSH2, MSH3, and MSH6 genetic perturbation in hESCs with CRISPR-Cas9

(A) Deep sequencing result of *MSH2* in MSH2 KO (M2KO), control for M2KO (Cont-2), and wild-type (WT) hESCs, *MSH3* in MSH3 KO (M3KO), control for M3KO (Cont-3), and WT hESCs, *MSH6* in MSH6 KO (M6KO), control for M6KO (Cont-6), and WT hESCs. (B–D) Immunoblotting for MSH2, MSH3, and MSH6 in MSH2 KO (M2KO) and control for M2KO (Cont-2) (B), MSH3 KO (M3KO) and control for M3KO (Cont-3) (C), and MSH6 KO (M6KO) and control for M6KO (Cont-6) hESCs (E) mRNA expression of *POU5F1* and *NANOG* in human dermal fibroblast (hDF), wild-type (WT), MSH2 KO (M2KO), control for M2KO (Cont-2), MSH3 KO (M3KO), control for M3KO (Cont-3), MSH6 KO (M6KO), and control for M6KO (cont-6) hESCs, Bars represent mean values, and error bars represent the SD (n = 2). (F) Immunoblotting for OCT4 in wild-type (WT), MSH2 KO (M2KO), control for M2KO (Cont-2), MSH3 KO (M3KO), control for M3KO (Cont-3), MSH6 KO (M6KO), and control for M6KO (Cont-6) hESCs. (G) Alkaline phosphatase activity assay of wild-type (WT), MSH2 KO (M2KO), control for M2KO (Cont-2), MSH3 KO (M3KO), control for M3KO (Cont-3), MSH6 KO (M6KO), and control for M6KO (Cont-6) hESCs.

results suggest that the editing size-dependent effect on PE efficiency occurred from MSH6 (for defective MutSα) or MSH3 (for defective MutSβ) loss.

DISCUSSION

Prime editing has had a substantial impact on hPSC-based disease modeling due to its high capability to correct approximately 89% of

****p < 0.0001, comparing with the hESCs by unpaired t test. (F) mRNA expression of *SOX2, LIN28, MSH2*, and *MSH6* in hESCs and human mesenchymal stem cells derived from hESCs (hESC-MSCs) and human dermal fibroblasts (hDFs). Bars represent mean values, and error bars represent the SD (n = 2). (G) Immunoblotting for MSH2 and MSH6 in undifferentiated hESCs (Undiff.) and differentiated cells from hESCs (Diff.). α-tubulin for equal loading control (H) mRNA expression of *SOX2, MSH2*, and *MSH6* at indicative days after spontaneous differentiation of hESCs. Dots represent mean values, and error bars represent the SD (n = 2). (I) mRNA expression of *SOX2, LIN28, MSH2*, and *MSH6* of human fibroblast (BJ-Fibroblast) and BJ-fibroblast-induced iPSCs (BJ-iPSCs). Bars represent mean values, and error bars represent the SD (n = 2).

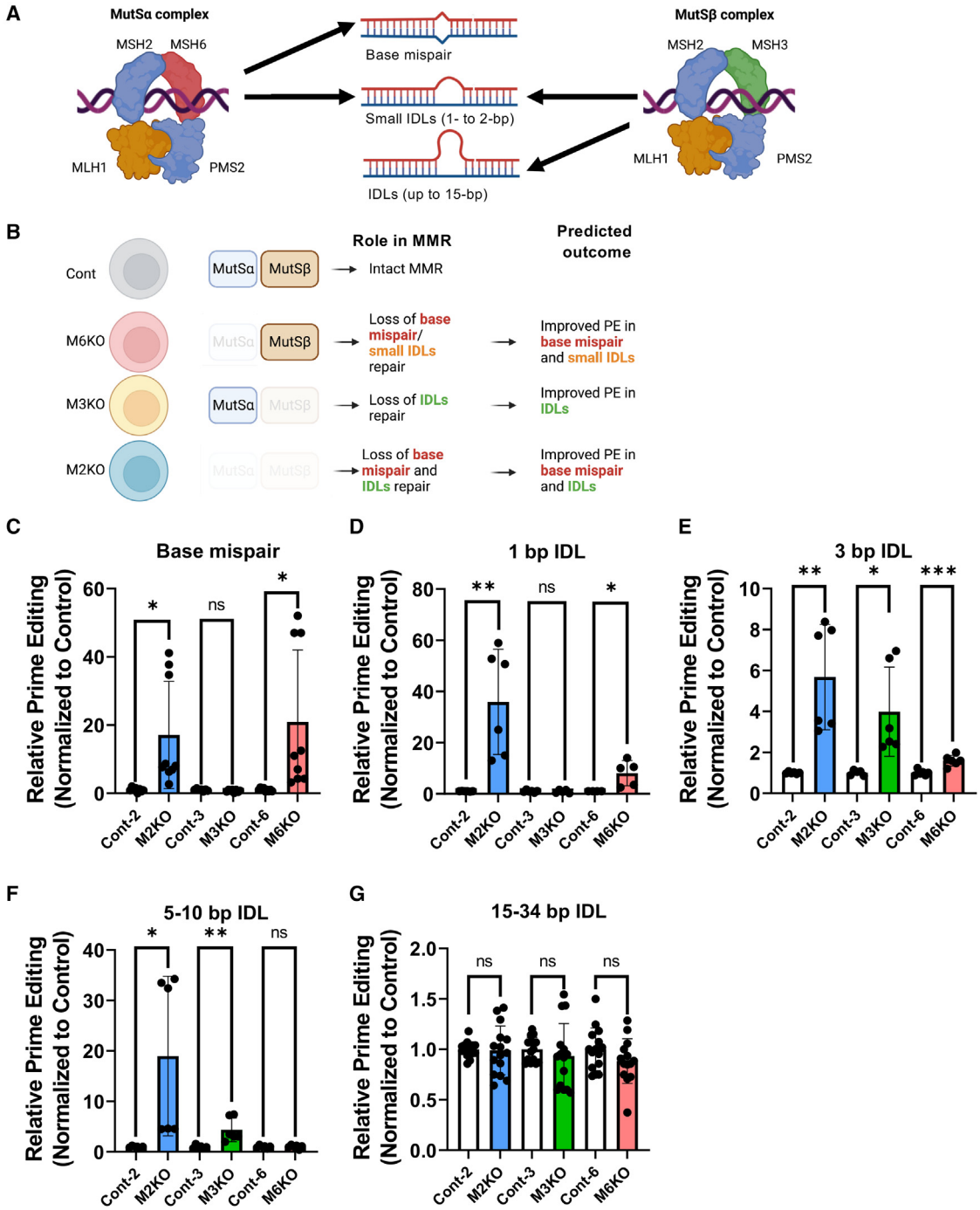


Figure 4. Editing size discrimination by MSH2, MSH3, and MSH6 in prime editing

(A and B) Graphical illustration of mismatch recognition by MutSα and MutSβ complex (A) and predicted outcome of each indicated cell model (B). (C–G) Prime editing efficiency of base mispair (C), 1-bp indel loop (IDL) (D), 3-bp IDL (E), 5–10 bp IDL (F), and 15–34 bp IDL (G) in MSH2 KO (M2KO), control for M2KO (Cont-2), MSH3 KO (M3KO), control for M3KO (Cont-3), MSH6 KO (M6KO), and control for M6KO (Cont-6) hESCs, (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$, and ns, not significant).

pathogenic mutations.[18] However, its low editing efficiency warrants developing techniques to improve this and avoid laborious clonal selection.[43] The identification of MSH2 and MSH6 (components of MutSα complex), as well as PMS2 and MLH1 (forming MutL complex), as critical factors to determine PE efficiency through pooled CRISPRi screens based on 1-bp mismatch,[10] clearly supports the

editing size-dependent effect of MutSα or MutSβ complex on PE efficiency shown in our studies. Considering the effect of MMR activity for PE efficiency, hPSCs where overall MMR activity is higher than the differentiated cells[44] would be highly affected by loss of MutSα or MutSβ complex.

To improve PE efficiency in hPSCs through MMR transient modulation, chemical perturbation with specific inhibitors of MutSα or MutSβ is desirable. However, aside from cadmium[45] inducing severe DNA damage,[46] no commercially available chemical compounds can specifically inhibit the MMR pathway on demand. Alternatively, siRNA transient knockdown of *MSH6* (for MutSα) or *MSH3* (for MutSβ) in hPSCs was attempted. However, siRNA failed to improve PE editing efficiency even after successful RNA repression (data not shown) possibly due to high protein stability of MSH2, MSH3, and MSH6 in hPSCs tightly regulated by ubiquitination-proteosome system.[47] Thus, we suggest that the most feasible approach is MutSα or MutSβ transient interruption with the dominant-negative form of MMR proteins (e.g., dnMLH1, used in the previous study[10]).

As previously reported,[10] PE3, by introducing a nick on the non-edited strand, may sequester the MMR machinery to the non-edited strand, which could be beneficial in hPSCs with high MMR activity. Therefore, it would be intriguing to investigate whether the higher editing outcomes of PE3 compared with PE2, as observed in hPSCs,[43] are a result of MMR strand discrimination. However, it should be noted that PE3 also introduces additional SSBs that have been shown to correlate with increased indel formation.[10] Considering the unique cellular characteristics of hPSCs, which undergo massive p53-dependent cell death in response to DNA damage, PE3 may potentially induce severe p53-dependent cell death, similar to Cas9.[25] This is why we hypothesize that simultaneous expression of p53 dominant-negative mutants could markedly increase the editing outcome of PE3 in hPSCs.[43]

Unlike cancer cell lines where the editing outcomes of newly developed editing tools are frequently monitored, hPSCs where precise genome editing is performed for mostly translational purpose (i.e., disease modeling and *ex vivo* cell therapy) should avoid inadvertent mutations. Thus, it would be necessary to consider the editing tools to minimize the possible off-target effects as well as editing efficiency. It is noteworthy that *MLH1* (or *MSH2*) loss causes a higher mutation frequency than other MMR component absences, such as *MSH3* or *MSH6*.[48] PE4 or PE5 applications are newly developed PE tools based on transient inhibition of MLH1, an essential MMR protein for impeding both MutSα- and MutSβ-dependent MMR. Accordingly, these applications necessitate close examination of whether any unintended mutations occur from simultaneous MutSα and MutSβ inhibition in rapidly proliferating hPSCs with a lack of cell-cycle checkpoints.[23] It has been well documented that p53-dominant mutations occur in hPSCs during *in vitro* culture, and the p53 mutant clones soon become dominant due to strong selective advantage.[49] In parallel, p53 mutant hPSCs, which survive Cas9-mediated DNA damage, are readily enriched after Cas9 genome editing.[50] The recent study to demonstrate that loss of MSH2 (leading to defective MMR) is sufficient to promote a selective advantage of hESCs[51] would imply that complete inhibition of MMR activity (i.e., with dnMLH1) during prime editing increases the possibility of enrichment of aberrant clones in hPSCs. This study demonstrates that MutSα and MutSβ are size-dependent cellular PE outcome determinants. Thus, rather than disrupt MLH1 or MSH2 that cause dual MutSα and MutSβ inhibition, MutSα or MutSβ transient inhibition according to desired editing size would achieve a safer PE hPSC editing. Additionally, we noted a variation in human cell lines' basal MMR activity with SW480, CaCo2, HEL, HeLa, MRC, and WI38 being MMR proficient and HCT116, LoVo, DLD1, HCT15, and SW48 being MMR deficient.[38] Accordingly, that PE4 or PE5 editing designed for MMR modulation works in a cell-type-specific manner underlies the gravity in considering cellular context when selecting appropriate editing tools.

Taken together, we demonstrated that elevated MMR activity in hPSCs served as a major determinant for overall PE editing outcomes. The distinct role of MutSα and MutSβ in MMR determines the size-dependent editing outcome of a PE in hPSCs.

MATERIALS AND METHODS

Cell line, culture and transfection

hESCs (WA09, WiCell Research Institute) were cultured on a Matrigel-coated (BD Biosciences) cell culture dish in StemMACS media (Miltenyi-Biotec) added with 50 μg/mL Gentamicin (Gibco). At 70%–80% of confluency, cells were rinsed with DPBS (Gibco) and exposed to Accutase (BD Biosciences) for detachment. Detached cells were washed with DMEM/F-12 (Gibco) media and plated to a Matrigel-coated plate fed with StemMACS media added with 50 μg/mL Gentamicin (Gibco) and 10 μM of Y27632 (Gibco). Cells were detached with Accutase solution (561527, BD Biosciences) and washed with Opti-MEM (31985070, Gibco) three times and diluted to a concentration of 1 x 10^6 cells in 100 μL of Opti-MEM for transfection. 2 μg of PE2 of Cas9 vector and 3 μg of sgRNA or pegRNA vector were added to the cell mixture. Electroporation was performed by NEPA-21 with 175 V of poring pulse and 2.5 ms of pulse length. All procedures were approved by the institutional Review Board at Seoul National University (SNU IRB protocol #2305/003-014).

Targeted deep sequencing

For analysis, genomic DNA (gDNA) of each sample was extracted by NucleoSpin Tissue kit (MN) following the supplier's instructions. To generate sequencing libraries, target sites were amplified using a TOYOBO KOD Multi & Epi PCR kit. As previously disclosed, these libraries were sequenced utilizing MiniSeq and an Illumina TruSeq HT Dual Index system. Using an Illumina MiniSeq technology, equal amounts of PCR amplicons were submitted to paired-end read sequencing. Paired-end reads were analyzed by comparing mutant and WT sequences using BE-analyzer after MiniSeq.

qRT-PCR

Easy-BLUE RNA isolation kit (iNtRON Biotechnology) was used for total RNA extraction following the supplier's instructions. cDNA was

synthesized by adding 2 μL of PrimeScript RT reagent kit (TaKaRa) to 500 μg of RNA samples in 8 μL of distilled water (DW) and reacted for 15 min in 37°C. Synthesized cDNA was diluted in 200 μL of DW, and 9 μL of diluted cDNA was added to a 96-well qPCR plate. 1 μL of qPCR primer and 10 μL of SYBR Green PCR reagents (Life Technologies) were added to each well. qPCR was performed by Light Cycler-480II (Loche).

Gene set enrichment analysis (GSEA) and KEGG pathway analysis

Four selected gene expression profiles by RNA microarray (GEO: GSE9709, GSE2248, GSE16963, GSE42445) were downloaded from the Gene Expression Omnibus database. Ranks for differently expressed RNA/genes between hESCs and differentiated cells were yielded by GEO2R (https://www.ncbi.nlm.nih.gov/geo/geo2r/). Using the gene list as input, GSEA analysis was performed using the KEGG_MISMATCH_REPAIR gene sets from MSigDB via the R package "fgsea," and KEGG pathway analysis was performed by the R package "pathview" (https://doi.org/10.1093/bioinformatics/btt285). Significantly upregulated or downregulated GO terms were selected with nominal p value < 0.05 and a |normalized enrichment score| > 1.5.

Statistical analysis

The quantitative data are expressed as the mean values ± standard deviation (SD). Student's unpaired t tests were performed to analyze the statistical significance of each response variable using the PRISM. p values less than 0.05 were considered statistically significant (*p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001).

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.2023.05.015.

ACKNOWLEDGMENTS

This research was supported by the National Research Foundation of Korea (NRF) no. NRF-2020R1A2C2005914 and a grant from the Korean Fund for Regenerative Medicine funded by Ministry of Science and ICT, and Ministry of Health and Welfare no. RS-2022-00070316.

AUTHOR CONTRIBUTIONS

H.J.C. conceived the overall study design and led the experiments. J.C.P. conducted the experiments and critical discussion of the results. Y.J.K., D.K., M.J.P., and J.K. contributed to establishing KO hESC lines. J.H.H. performed next generation sequencing (NGS). H.K.J. and S.B. provided key materials and analyzed NGS data. All authors contributed to manuscript writing and revising and endorsed the final manuscript.

DECLARATION OF INTERESTS

The authors declare no conflict of interest.

REFERENCES

1. Merkle, F.T., and Eggan, K. (2013). Modeling human disease with pluripotent stem cells: from genome association to function. Cell Stem Cell *12*, 656–668.
2. Musunuru, K. (2013). Genome editing of human pluripotent stem cells to generate human cellular disease models. Dis. Model. Mech. *6*, 896–904.
3. Rowe, R.G., and Daley, G.Q. (2019). Induced pluripotent stem cells in disease modelling and drug discovery. Nat. Rev. Genet. *20*, 377–388.
4. 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.
5. 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.
6. Park, J.C., Kim, J., Jang, H.K., Lee, S.Y., Kim, K.T., Kwon, E.J., Park, S., Lee, H.S., Choi, H., Park, S.Y., et al. (2022). Multiple isogenic GNE-myopathy modeling with mutation specific phenotypes from human pluripotent stem cells by base editors. Biomaterials *282*, 121419.
7. Habib, O., Habib, G., Hwang, G.H., and Bae, S. (2022). Comprehensive analysis of prime editing outcomes in human embryonic stem cells. Nucleic Acids Res. *50*, 1187–1197.
8. Zhang, J.P., Li, X.L., Li, G.H., Chen, W., Arakaki, C., Botimer, G.D., Baylink, D., Zhang, L., Wen, W., Fu, Y.W., et al. (2017). Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. Genome Biol. *18*, 35.
9. 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.
10. Chen, P.J., Hussmann, J.A., Yan, J., Knipping, F., Ravisankar, P., Chen, P.F., Chen, C., Nelson, J.W., Newby, G.A., Sahin, M., et al. (2021). Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. Cell *184*, 5635–5652.e29.
11. Sander, J.D., and Joung, J.K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. Nat. Biotechnol. *32*, 347–355.
12. 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.
13. Cullot, G., Boutin, J., Toutain, J., Prat, F., Pennamen, P., Rooryck, C., Teichmann, M., Rousseau, E., Lamrissi-Garcia, I., Guyonnet-Duperat, V., et al. (2019). CRISPR-Cas9 genome editing induces megabase-scale chromosomal truncations. Nat. Commun. *10*, 1136.
14. Simkin, D., Papakis, V., Bustos, B.I., Ambrosi, C.M., Ryan, S.J., Baru, V., Williams, L.A., Dempsey, G.T., McManus, O.B., Landers, J.E., et al. (2022). Homozygous might be hemizygous: CRISPR/Cas9 editing in iPSCs results in detrimental on-target defects that escape standard quality controls. Stem Cell Rep. *17*, 993–1008.
15. 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.
16. 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.
17. Song, Y., Liu, Z., Zhang, Y., Chen, M., Sui, T., Lai, L., and Li, Z. (2020). Large-fragment deletions induced by Cas9 cleavage while not in the BEs system. Mol. Ther. Nucleic Acids *21*, 523–526.
18. 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., and Liu, D.R. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. Nature *576*, 149–157.

19. Ferreira da Silva, J., Oliveira, G.P., Arasa-Verge, E.A., Kagiou, C., Moretton, A., Timelthaler, G., Jiricny, J., and Loizou, J.I. (2022). Prime editing efficiency and fidelity are enhanced in the absence of mismatch repair. Nat. Commun. *13*, 760.

20. Nelson, J.W., Randolph, P.B., Shen, S.P., Everette, K.A., Chen, P.J., Anzalone, A.V., An, M., Newby, G.A., Chen, J.C., Hsu, A., and Liu, D.R. (2022). Engineered pegRNAs improve prime editing efficiency. Nat. Biotechnol. *40*, 402–410.

21. 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.

22. 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.

23. Weissbein, U., Benvenisty, N., and Ben-David, U. (2014). Quality control: genome maintenance in pluripotent stem cells. J. Cell Biol. *204*, 153–163.

24. 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.

25. 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.

26. 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.

27. 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 *11*.

28. Li, H., Busquets, O., Verma, Y., Syed, K.M., Kutnowski, N., Pangilinan, G.R., Gilbert, L.A., Bateup, H.S., Rio, D.C., Hockemeyer, D., and Soldner, F. (2022). Highly efficient generation of isogenic pluripotent stem cell models using prime editing. Elife *11*, e79208.

29. Jiricny, J. (2006). The multifaceted mismatch-repair system. Nat. Rev. Mol. Cell Biol. *7*, 335–346.

30. Palombo, F., Gallinari, P., Iaccarino, I., Lettieri, T., Hughes, M., D'Arrigo, A., Truong, O., Hsuan, J.J., and Jiricny, J. (1995). GTBP, a 160-kilodalton protein essential for mismatch-binding activity in human cells. Science *268*, 1912–1914.

31. Drummond, J.T., Li, G.M., Longley, M.J., and Modrich, P. (1995). Isolation of an hMSH2-p160 heterodimer that restores DNA mismatch repair to tumor cells. Science *268*, 1909–1912.

32. Verma, R., Agarwal, A.K., Sakhuja, P., and Sharma, P.C. (2019). Microsatellite instability in mismatch repair and tumor suppressor genes and their expression profiling provide important targets for the development of biomarkers in gastric cancer. Gene *710*, 48–58.

33. Yeh, C.C., Lee, C., and Dahiya, R. (2001). DNA mismatch repair enzyme activity and gene expression in prostate cancer. Biochem. Biophys. Res. Commun. *285*, 409–413.

34. 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. C Methods *21*, 322–329.

35. Bang, J.S., Choi, N.Y., Lee, M., Ko, K., Lee, H.J., Park, Y.S., Jeong, D., Chung, H.M., and Ko, K. (2018). Optimization of episomal reprogramming for generation of human induced pluripotent stem cells from fibroblasts. Anim. Cell Syst. *22*, 132–139.

36. 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. USA *110*, E3281–E3290.

37. Go, Y.H., Lim, C., Jeong, H.C., Kwon, O.S., Chung, S., Lee, H., Kim, W., Suh, Y.G., Son, W.S., Lee, M.O., et al. (2019). Structure-activity relationship analysis of YM155 for inducing selective cell death of human pluripotent stem cells. Front. Chem. *7*, 298.

38. Chang, D.K., Ricciardiello, L., Goel, A., Chang, C.L., and Boland, C.R. (2000). Steady-state regulation of the human DNA mismatch repair system. J. Biol. Chem. *275*, 18424–18431.

39. de Wind, N., Dekker, M., Claij, N., Jansen, L., van Klink, Y., Radman, M., Riggins, G., van der Valk, M., van't Wout, K., and te Riele, H. (1999). HNPCC-like cancer predisposition in mice through simultaneous loss of Msh3 and Msh6 mismatch-repair protein functions. Nat. Genet. *23*, 359–362.

40. Sharma, M., Predeus, A.V., Kovacs, N., and Feig, M. (2014). Differential mismatch recognition specificities of eukaryotic MutS homologs, MutSalpha and MutSbeta. Biophys. J. *106*, 2483–2492.

41. Warren, J.J., Pohlhaus, T.J., Changela, A., Iyer, R.R., Modrich, P.L., and Beese, L.S. (2007). Structure of the human MutSalpha DNA lesion recognition complex. Mol. Cell *26*, 579–592.

42. Gupta, S., Gellert, M., and Yang, W. (2011). Mechanism of mismatch recognition revealed by human MutSbeta bound to unpaired DNA loops. Nat. Struct. Mol. Biol. *19*, 72–78.

43. Li, M., Zhong, A., Wu, Y., Sidharta, M., Beaury, M., Zhao, X., Studer, L., and Zhou, T. (2022). Transient inhibition of p53 enhances prime editing and cytosine base-editing efficiencies in human pluripotent stem cells. Nat. Commun. *13*, 6354.

44. Lin, B., Gupta, D., and Heinen, C.D. (2014). Human pluripotent stem cells have a novel mismatch repair-dependent damage response. J. Biol. Chem. *289*, 24314–24324.

45. Banerjee, S., and Flores-Rozas, H. (2005). Cadmium inhibits mismatch repair by blocking the ATPase activity of the MSH2-MSH6 complex. Nucleic Acids Res. *33*, 1410–1419.

46. Bertin, G., and Averbeck, D. (2006). Cadmium: cellular effects, modifications of biomolecules, modulation of DNA repair and genotoxic consequences (a review). Biochimie *88*, 1549–1559.

47. Zhang, M., Xiang, S., Joo, H.Y., Wang, L., Williams, K.A., Liu, W., Hu, C., Tong, D., Haakenson, J., Wang, C., et al. (2014). HDAC6 deacetylates and ubiquitinates MSH2 to maintain proper levels of MutSalpha. Mol. Cell *55*, 31–46.

48. Hegan, D.C., Narayanan, L., Jirik, F.R., Edelmann, W., Liskay, R.M., and Glazer, P.M. (2006). Differing patterns of genetic instability in mice deficient in the mismatch repair genes Pms2, Mlh1, Msh2, Msh3 and Msh6. Carcinogenesis *27*, 2402–2408.

49. Merkle, F.T., Ghosh, S., Kamitaki, N., Mitchell, J., Avior, Y., Mello, C., Kashin, S., Mekhoubad, S., Ilic, D., Charlton, M., et al. (2017). Human pluripotent stem cells recurrently acquire and expand dominant negative P53 mutations. Nature *545*, 229–233.

50. Enache, O.M., Rendo, V., Abdusamad, M., Lam, D., Davison, D., Pal, S., Currimjee, N., Hess, J., Pantel, S., Nag, A., et al. (2020). Cas9 activates the p53 pathway and selects for p53-inactivating mutations. Nat. Genet. *52*, 662–668.

51. Madden-Hennessey, K., Gupta, D., Radecki, A.A., Guild, C., Rath, A., and Heinen, C.D. (2022). Loss of mismatch repair promotes a direct selective advantage in human stem cells. Stem Cell Rep. *17*, 2661–2673.