YMTHE, Volume *30*

Supplemental Information

Broadening prime editing toolkits

using RNA-Pol-II-driven engineered pegRNA

Shisheng Huang, Zhenwu Zhang, Wanyu Tao, Yao Liu, Xiangyang Li, Xiaolong Wang, Javad Harati, Peng-Yuan Wang, Xingxu Huang, and Chao-Po Lin

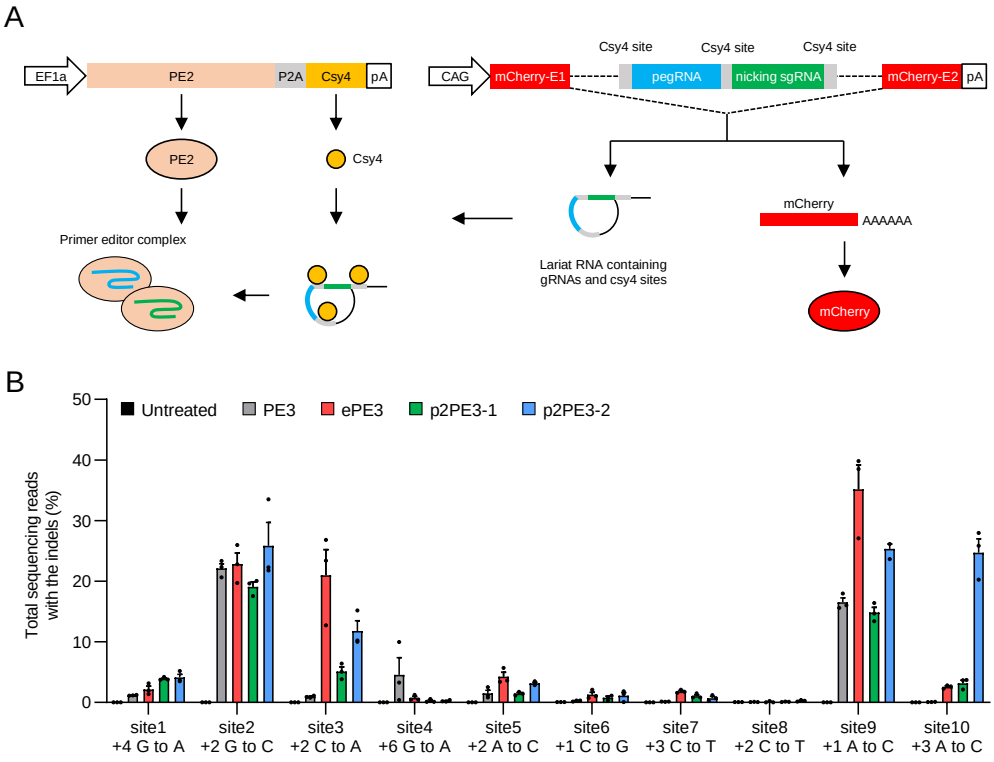

Figure S1. Analysis of pol II-mediated transcription of pegRNA editing system. (**A**) A diagram shown how functional pegRNAs were generated from introns in pol II-mediated transcripts. mCherry was used as the harboring gene. A P2A peptide was used to link Csy4 with PE2. E: Exon. (**B**) Comparison of indels with PE3, ePE3, p2PE3-1 and p2PE3-2 at ten endogenous target sites in HEK293T cells. Data were represented as the mean ± SEM (n = 3 from independent experiments).

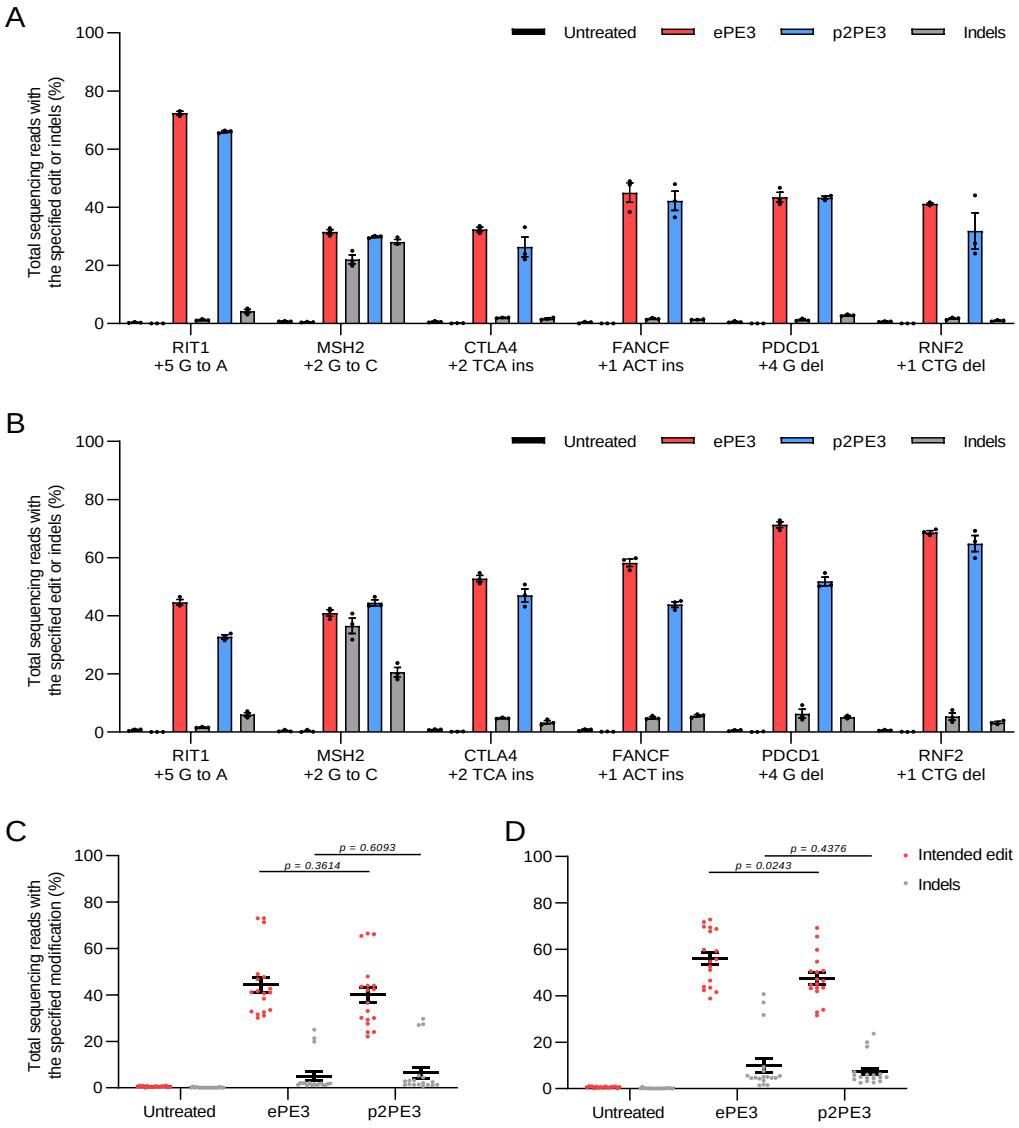

Figure S2. The performance of ePE3 and p2PE3 in U2OS and Hela. (**A, B**) Comparison of editing and indel frequencies with ePE3 and p2PE3 targeting 6 individual sites (2 substitutions, 2 insertions and 2 deletions) in U2OS (A) and Hela (B) cells. (**C, D**) Statistical analysis of intended editing and indel frequencies in U2OS (C) and Hela (D) cells. Data were represented as the mean ± SEM (n = 3 from independent experiments). Two-tailed Student's t-tests were performed.

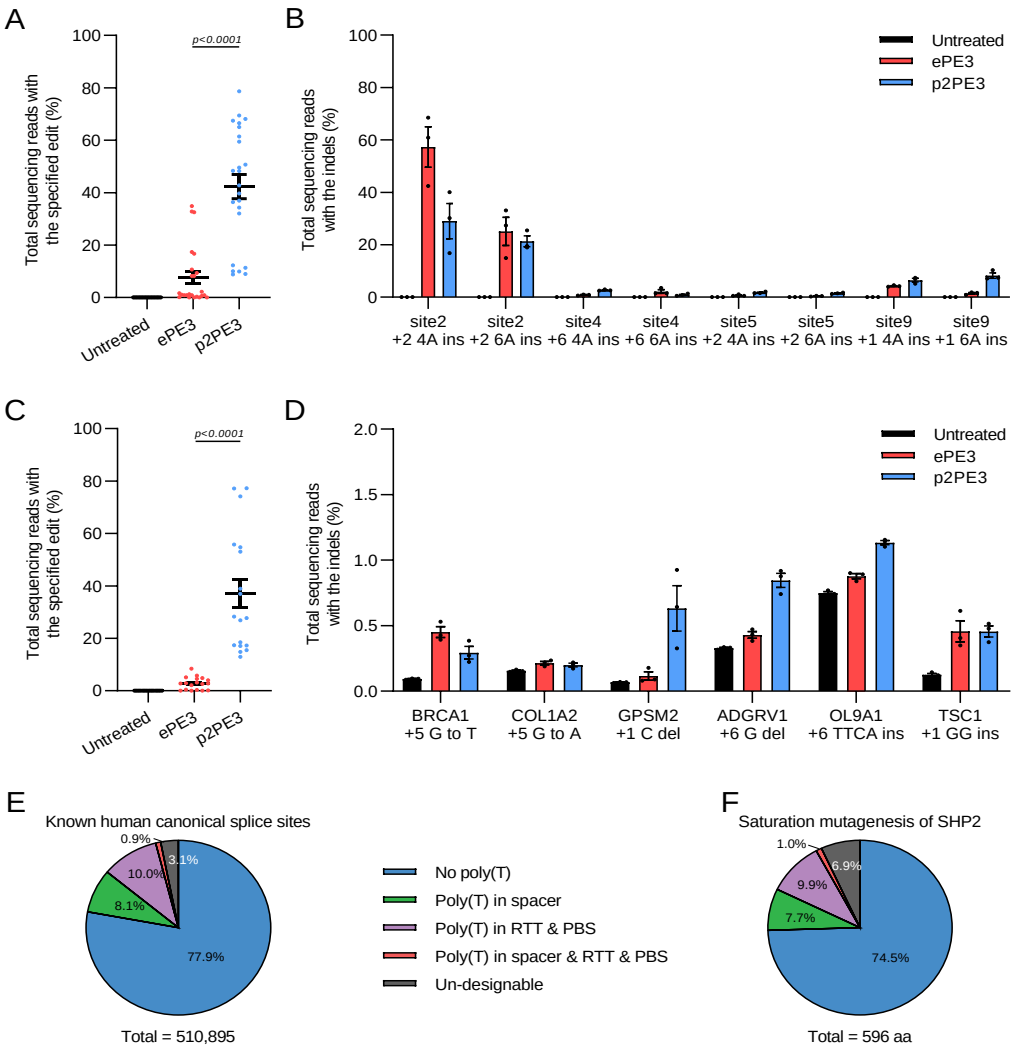

Figure S3. Analysis of ePE3 and p2PE3-mediated editing with pegRNA containing poly(T). (**A, B**) Summary of editing efficiency (**A**) and indels (**B**) of ePE3 and p2PE3-mediated insertions of poly(A) at four endogenous target sites in HEK293T cells. (**C, D**) Summary of editing efficiency (**C**) and indels (**D**) of ePE3 and p2PE3-mediated editing with pegRNA containing poly(T) at six human pathogenic sites in HEK293T cells. Data and error bars indicated the mean ± SEM of three independent experiments. Two-tailed Student's t-tests were performed. (**E**) The proportion of pegRNA containing poly(T) to target known human canonical splice sites. (**F**) the proportion of poly(T) in the pegRNAs targeting SHP2 gene to make saturation mutagenesis library.

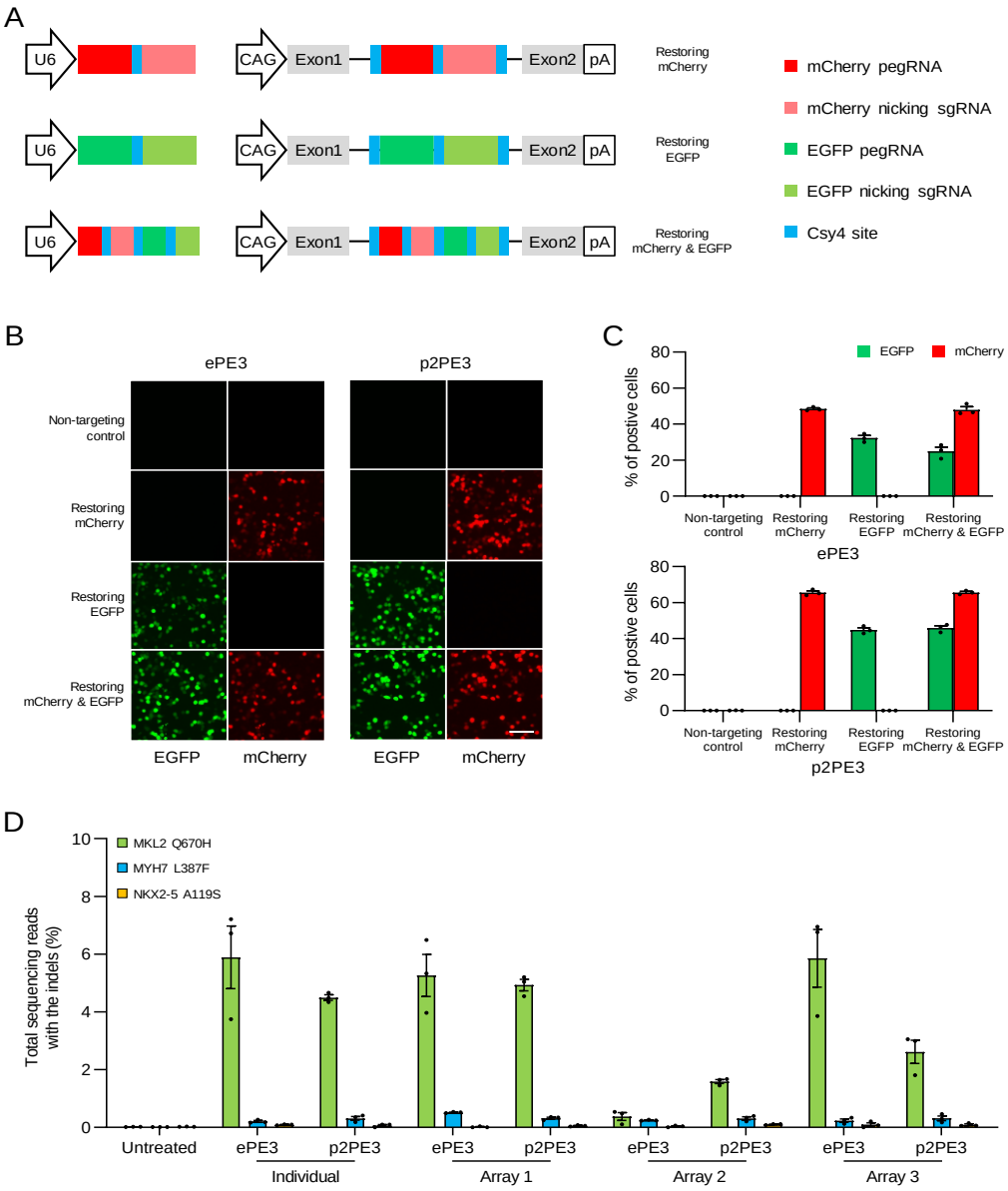

Figure S4. The performance of ePE3 and p2PE3 in combinatorial genetic editing. (**A**) Diagram of the pegRNA-nicking sgRNA cassettes restoring the fluorescence of mCherry, EGFP or the dual-color fluorescence. Harboring gene of p2PE3 was replaced with irrelevant transcript. (**B**) The fluorescence of restored mCherry and EGFP determined by microscope images. *Scale bars*, 100 μm. (**C**) Analysis of fluorescence by flow cytometry. Additional blue fluorescent protein (BFP) fluorescence was used to indicate cell transfection. The percentage was calculated as the number of positive cells/total transfected cells. (**D**) Indels of ePE3 and p2PE3 to model childhood-onset cardiomyopathy mutations. Data were represented as the mean ± SEM (n = 3 from three independent experiments).

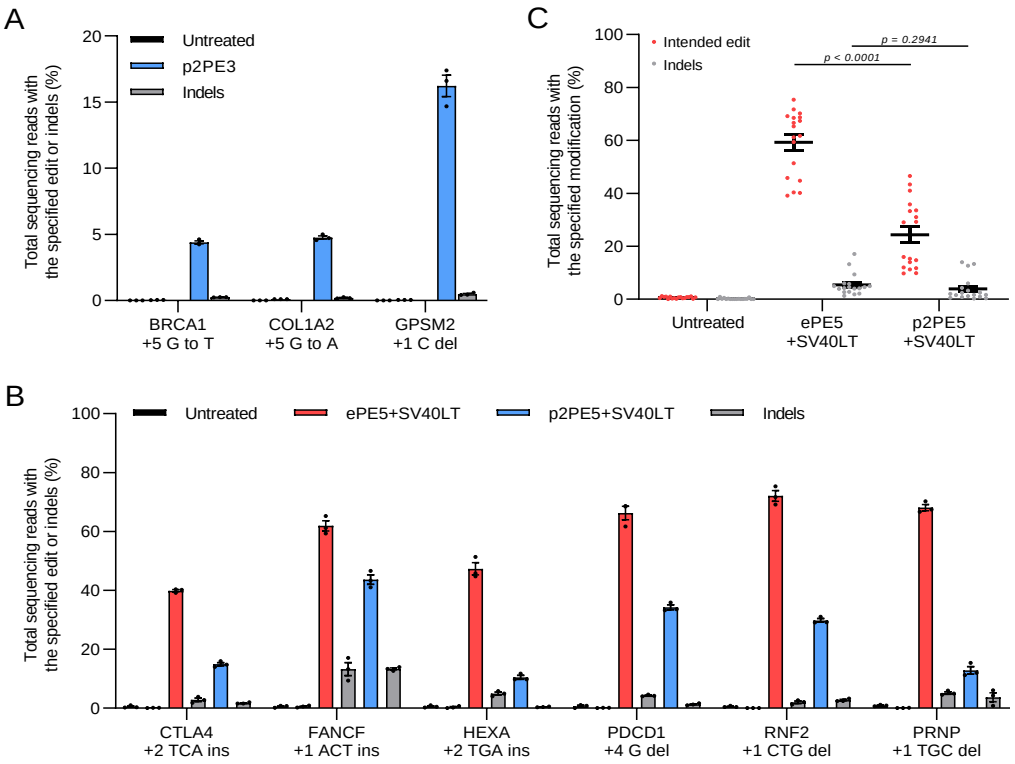

Figure S5. The comparison of editing efficiency of ePE3 and p2PE3 in hESCs. (**A**) Efficiency of p2PE3-mediated base substitution and small targeted deletion in hESCs. (**B**) Comparison of editing and indel frequencies with ePE3 and p2PE3 targeting 6 individual sites (2 substitutions, 2 insertions and 2 deletions) in the presence of hMLH1^{NTD}-NLS and SV40LT (ePE5+SV40LT and p2PE5+SV40LT) in hESCs. (**C**) Statistical analysis of intended editing and indel frequencies in (B). Data were represented as the mean ± SEM (n = 3 from independent experiments). Two-tailed Student's t-tests were performed.

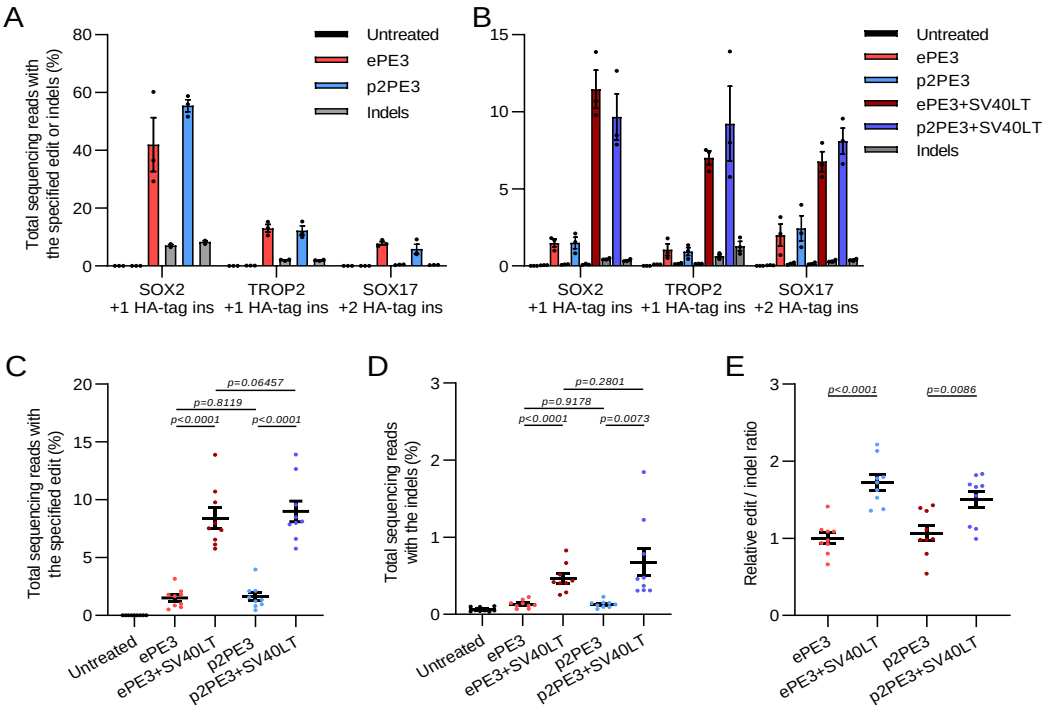

Figure S6. Editing efficiency of long targeted insertions. (**A**) Efficiency of ePE3 and p2PE3-mediated long targeted insertion (HA-tag insertion, 27 bp) in HEK293T. (**B**) Transient p53 inhibition increased efficiency of ePE3 and p2PE3-mediated long targeted insertion in hESCs. (**C, D**) Statistical analysis of the editing frequency (C) and indels (D) in (B). (**E**) Relative edit:indel ratios associated with ePE3 and p2PE3-mediated editing enhanced by SV40LT. The levels of the ePE3 group were set as 1. The results correspond to those shown in (B). Data were represented as the mean ± SEM (n = 3 from independent experiments). Two-tailed Student's t-tests were performed.

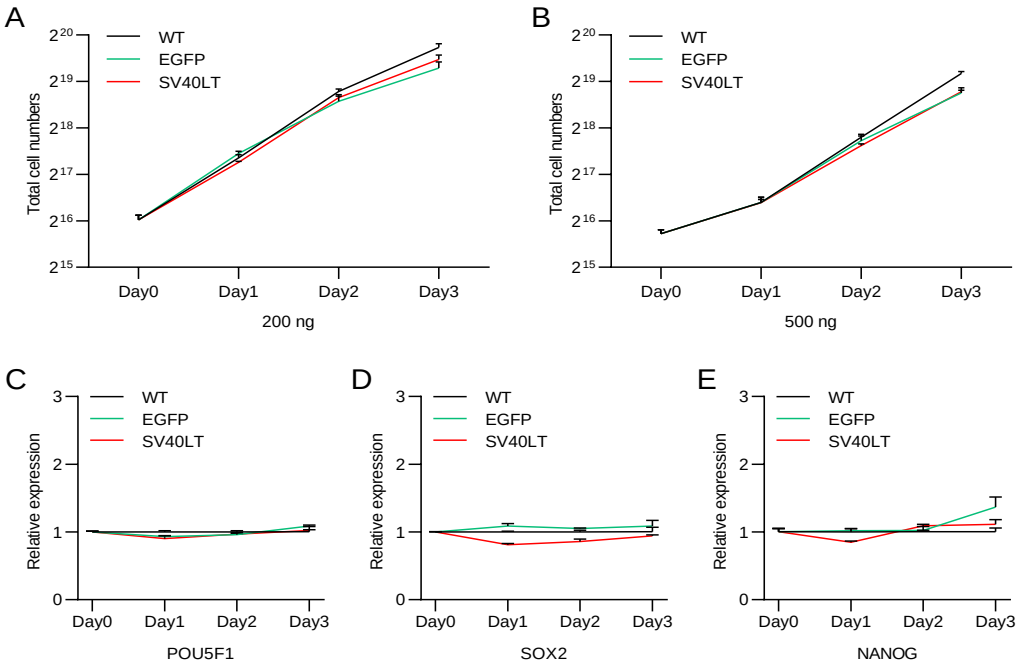

Figure S7. Evaluating the effect of p53 inhibition on hESCs. (**A, B**) Cells were assayed for growth 0, 24, 48 and 72 h after transfection with 200 ng (A) or 500 ng (B) of EGFP or SV40LT expression plasmid. (**C, D, E**) The expression of stemness gene at different time points in (B). Data were represented as the mean ± SEM (n = 3 from independent experiments).

A

Sample	Sequenced reads	Mapping rate	Duplication	Mean depth
Untreated	671,094,320	99.91%	18.75%	26.10X
Control	657,824,626	99.45%	15.49%	26.51X
Edited	628,782,768	99.48%	15.18%	25.42X

B

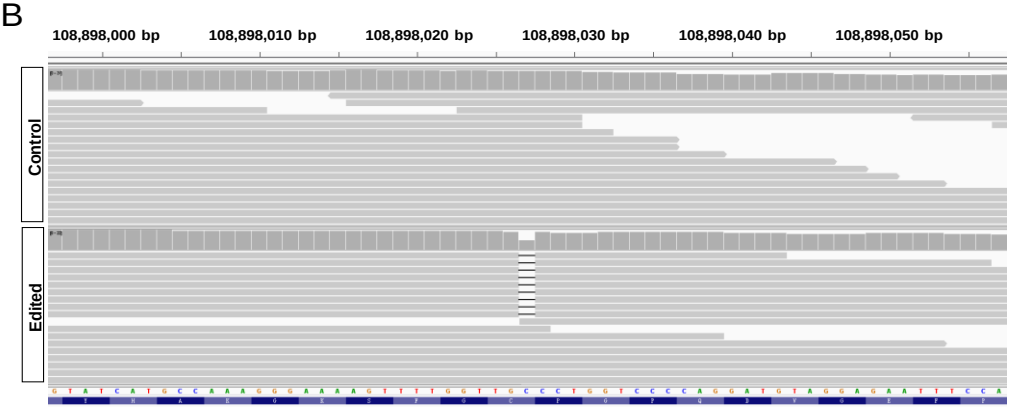

Figure S8. The information of whole genome sequencing. (**A**) Sequencing statistics of whole genome sequencing samples. Control: H1 hESCs were transfect with EGFP plasmid; Edited: H1 hESCs were transfect with pEF1a-PE2-P2A-Csy4, p2PE3 gRNA cassette targeting GPSM2 site and pEF1a-SV40LT-P2A-hMLH1^{NTD}-NLS plasmids. (**B**) Confirmation of the on-target editing (GPSM2 +1 C del) by analyzing the whole-genome sequencing results.

A

Sample	Sequenced reads	Mapped reads	Mapping rate
Untreated	23,183,089	21,755,534	93.84%
Control-1	22,189,192	20,699,561	93.29%
Control-2	22,386,445	20,949,765	93.58%
Control-3	23,145,508	21,692,466	93.72%
Edited-1	22,065,185	20,543,345	93.10%
Edited-2	23,947,068	22,237,410	92.86%
Edited-3	23,700,356	22,109,047	93.29%

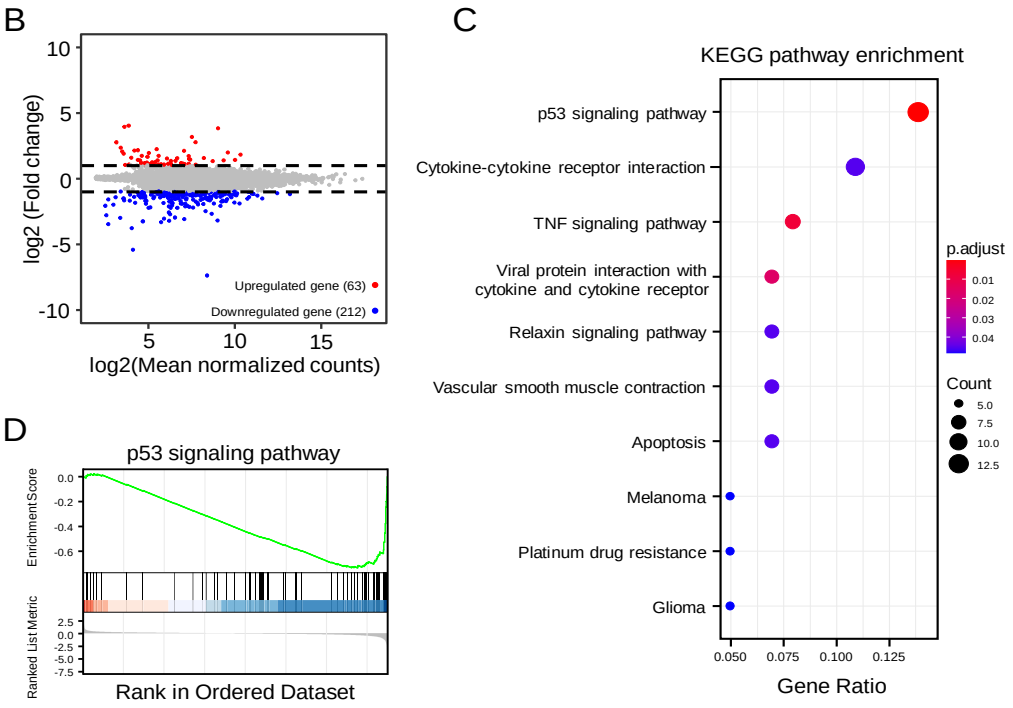

Figure S9. The differential expression analysis of RNA sequencing. (**A**) Sequencing statistics of RNA sequencing samples. Control: H1 hESCs were transfect with EGFP plasmid; Edited: H1 hESCs were transfect with pEF1a-PE2-P2A-Csy4, p2PE3 gRNA cassette targeting GPSM2 site and pEF1a-SV40LT-P2A-hMLH1^{NTD}-NLS plasmids. (**B**) The differentially expressed genes of edited samples comparing with control. (**C**) A dot plot to visualize the rich KEGG pathway from the differentially expressed genes. (**D**) The GSEA enrichment plot showed down-regulation of p53 signaling pathway in edited samples.

Table S2. Primers used in this study

Name	Forward primer	Reverse primer
site1	GTATGGAAAGGTAAGGCACTG	TCCATGCTTCCTGTCAAATGG
site2	CTCAGCATTCAGTGCTCTCC	TCGCATTTGCACTAGTCCTC
site3	GGCTGAGAGGACTGATCTTTCT	TCGACCTCGAGGAGACAATG
site4	ACGTAGGAATTTTGGTGGGAC	GTTTACACGTCTCATATGCC
site5	GAGCCCTCACTTTGGGTGTT	CCTCATTGCCAATGGATCAG
site6	CTCTAGGTGATGCTCAAGATG	ATTTTGGGCAAGGTCTGCGT
site7	AGAAAACAGGATGACCCCGATG	GACATTGTGGCCATCATTCC
site8	ACGTTGAGCTGTGCAGAGAA	TTGAAGCCAACCCACACAGT
site9	GGCAAACAAGGGAGTAATTC	AGAGAGACGGGAAGCCATTG
site10	CTGCTGGGAGATGTAGTCCAT	TTTGTGCGGTCGGGTAAACA
BRCA1	TAGCTTCTTAGGACAGCACTTC	GGTAACTCAGACTCAGCATCAG
COL1A2	ACAGAAACCACAGACTAGGGA	GTGTGGTTCTTAGATGAATGCT
GPSM2	CAGATTAGGTAGCATGTCTCTC	AGTCAGCTGTTGGGACAATC
ADGRV1	GCAGCTGTCCTCTGAAAGATA	GAAAGCCGCCTATCGGAAAG
OL9A1	GATGTCTTCATTTAGGTGGGAGA	CTTAGATGGGCCTACATGACTG
TSC1	GGCACATTGGTCTTTGAACC	TGGTATGGAGCACTCTGTTG
MKL2	CTTGGCTCCTCCATCAAAGAT	CTGCGTGGTCAGTAAAGCCT
MYH7	TCTCATCCCACCATGCCAGT	ACCAACTTTGCTACTTGCCT
NKX2-5	CCTCCACGAGGATCCCTTAC	GGTACCGCTGCTGCTTGAA
SOX2	TTCACATGTCCCAGCACTAC	TCATTTGCTGTGGGTGATGG
TROP2	GCTGCACACGGTCATCTTG	CCTGCAGACCATCCCAGA
SOX17	TGGGTACGCTGTAGACCAGA	TCTGGTCGTCACTGGCGTAT
RIT1	GTATGGAAAGGTAAGGCACTG	TCCATGCTTCCTGTCAAATGG
MSH2	CTCAGCATTCAGTGCTCTCC	TCGCATTTGCACTAGTCCTC
CTLA4	ATGCATCTCCAGGCAAAGCC	CTTGCAGATGTAGAGTCCCG
FANCF	CCTGCGCCACATCCATCGGC	TGCACCAGGTGGTAACGAGC
HEXA	GAGAGCTCGCCCAACATCGC	CCTGTTCTTGCCAGCAGGGC
PDCD1	GCACTGCCTCTGTCACTCT	CCGACCCCACCTACCTAAGA
PRNP	TGAGCAGCTGATACCATTGC	GCGGTTGCCTCCAGGGCTGC
RNF2	CCTCGCTCGCTCGCTCCTTC	CAGCCCAGGGCTCCGCTGGC
GAPDH-qPCR	GGCCCCCTCAAGGGCATCCT	GGGCCATGAGGTCCACCACC
NANOG-qPCR	TGTTTGGGATTGGGAGGCTT	GCACAACCAACAAATTAGGGGA
OCT4-qPCR	GTGGAGGAAGCTGACAACAATG	TCTCACTCGGTTCTCGATACTGG
SOX2-qPCR	GACCAGCTCGCAGACCTACATG	ACTTGACCACCGAACCCATG