Stem Cell Reports, Volume *15*

Supplemental Information

Inhibition of Apoptosis Reduces Diploidization of Haploid Mouse Embryonic Stem Cells during Differentiation

Wenhao Zhang, Yaru Tian, Qian Gao, Xu Li, Yanni Li, Jinxin Zhang, Chunmeng Yao, Yuna Wang, Haoyu Wang, Yiding Zhao, Qian Zhang, Luyuan Li, Yang Yu, Yong Fan, and Ling Shuai

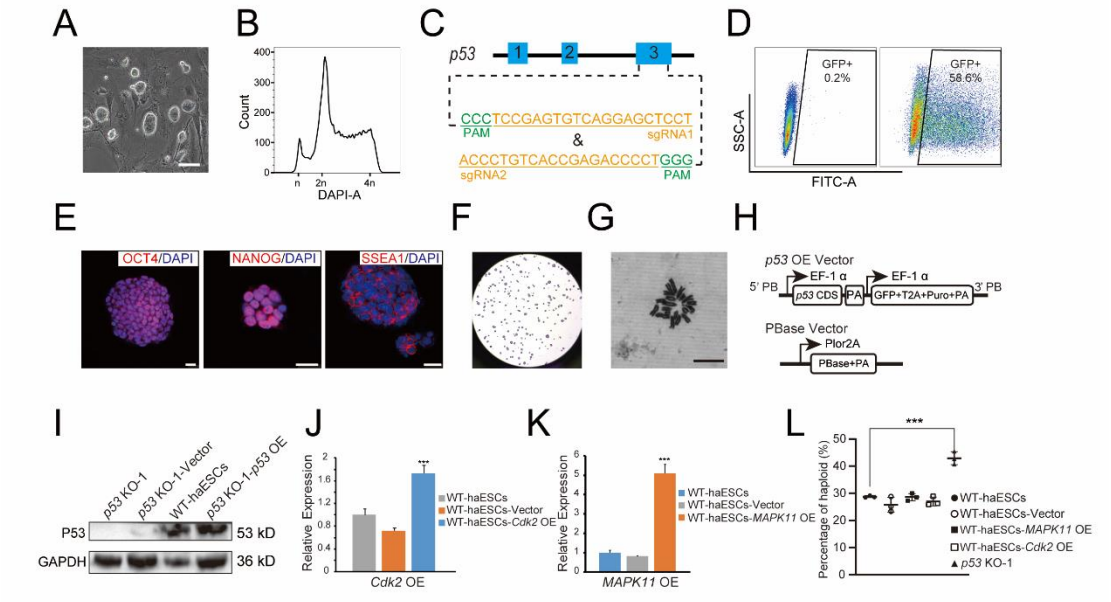


Figure S1. Gene editing in haESCs and characterization. Related to Figure 1

(A) Colonies of WT-haESCs in bright field. Scale bar, 100 μm.

(B) DNA content analysis of WT-haESCs for *p53*-deletion. The percentage of the 1n peak (G0/G1 phase) in WT-haESCs was 7.0%.

(C) DNA information for *p53* knocking out using the CRISPR/Cas9 system.

(D) Percentage of Cas9-GFP positive cells in transfected WT-haESCs 2 days after transfection by FACS indicating. WT-haESCs without transfection were used as control.

(E) Immunofluorescence staining of pluripotent markers, including OCT4, NANOG and SSEA1 in *p53*-KO haESCs. DNA was stained with DAPI. Scale bar, 50 μm.

(F) Alkaline phosphatase (AP) staining of *p53*-KO haESCs.

(G) Chromosome spread analysis of *p53*-KO haESCs. Scale bar, 10 μm.

(H) Schematic overview of the vectors used for overexpression of *p53*.

(I) Western blotting analysis of P53 in *p53* KO-1, *p53* KO-1-Vector, WT-haESCs and *p53* KO-1-*p53* OE cells. GAPDH was used as a loading control.

(J) Expression levels of *Cdk2* in WT-haESCs, WT-haESCs-Vector and WT-haESCs-*Cdk2* OE cells determined by qPCR. (n=3 independent experiments). t test, ***p < 0.001. Data was represented as the mean ± SEM.

(K) Expression levels of *MAPK11* in WT-haESCs, WT-haESCs-Vector and WT-haESCs-*MAPK11* OE cells determined by qPCR. (n=3 independent experiments). t test, ***p < 0.001. Data was represented as the mean ± SEM.

(L) Percentage of haploid cells in WT-haESCs, WT-haESCs-Vector, WT-haESCs-*MAPK11*-OE, WT-haESCs-*Cdk2*-OE and *p53*-KO haESCs at day 10 after sorting. Nearly 29% of sorted cells remained in haploidy in WT-haESCs-*Cdk2*-OE, WT-haESCs-*MAPK11*-OE, WT-haESCs and WT-haESCs-Vector, whereas *p53* KO-1 still had 44% of cells stayed in haploidy. (n=3 independent experiments). t test, ***p < 0.001. Data was represented as the mean ± SEM.

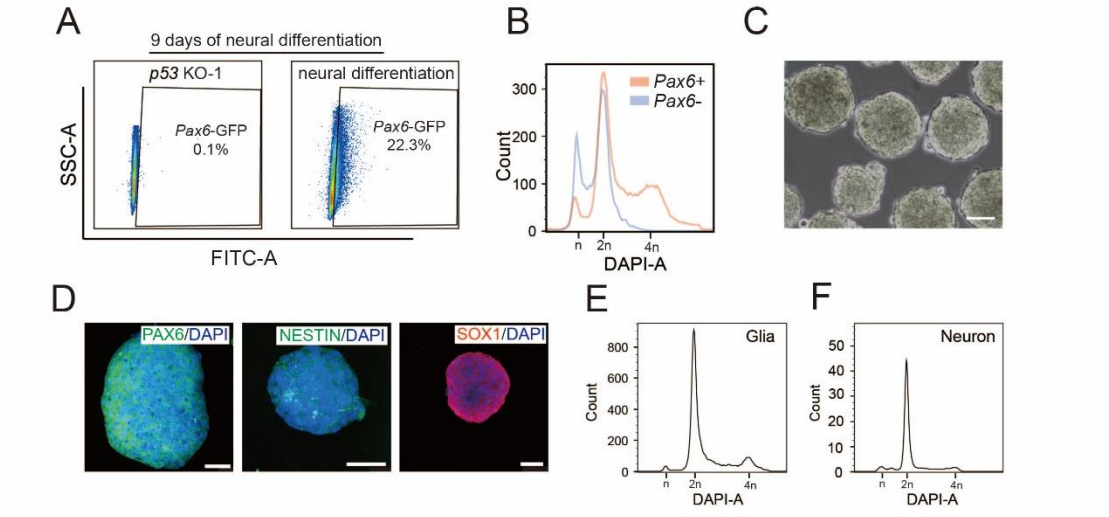


Figure S2. *In vitro* differentiation potentials of *p53*-KO haESCs. Related to Figure 2

(A) FACS analysis of *Pax6*-GFP-positive cells on 9-day of neural differentiation. *p53*-KO haESCs was used as the *Pax6*-GFP-negative control.

(B) DNA content analysis of *Pax6*-GFP-positive and -negative cells derived from *p53*-KO haESCs. The percentage of the 1n (G0/G1) peak in *Pax6*-GFP-positive cells was 6.2%, and that in *Pax6*-GFP-negative cells was 27.8%.

(C) The haNSCLCs spontaneously aggregate to form compact neural sphere structures in suspension culture. Scale bar, 100 μm.

(D) Immunostaining of PAX6, NESTIN and SOX1 in the neural spheres derived from *p53*-KO haESCs. Scale bar, 50 μm.

(E) DNA content analysis of glial mixture differentiated from haNSCLCs. 2.0% of cells remained in the 1n peak.

(F) DNA content analysis of neuronal mixture differentiated from haNSCLCs. 4.5% of cells remained in the 1n peak.

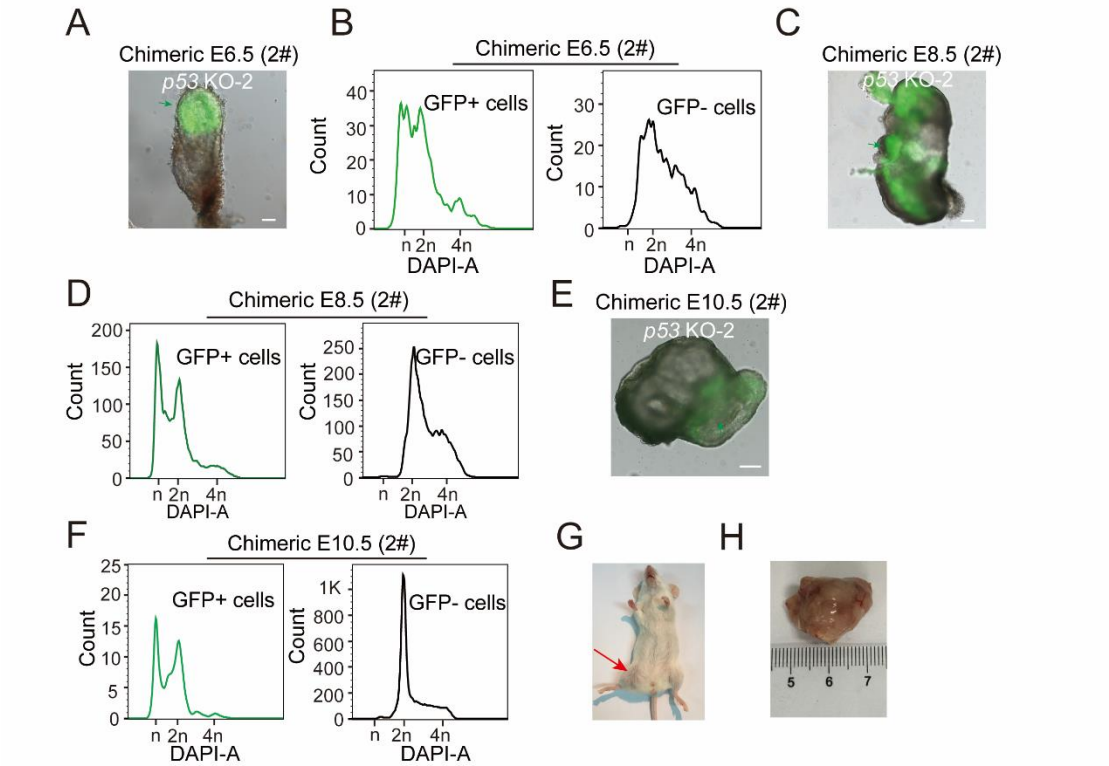


Figure S3. *In vivo* differentiation potentials of *p53*-KO haESCs. Related to Figure 3

(A) Morphology of chimeric embryo at E6.5 derived from GFP labelling *p53* KO-2 cells. Scale bar, 100 μm.

(B) DNA content analysis of GFP+ cells and GFP- cells in chimeric E6.5 embryo derived from *p53* KO-2 cells. The percentage of the 1n (G0/G1) peak in GFP+ cells was 24.5%.

(C) Morphology of chimeric embryo at E8.5 derived from GFP labelling *p53* KO-2 cells. Scale bar, 100 μm.

(D) DNA content analysis of GFP+ cells and GFP- cells in chimeric E8.5 embryo derived from *p53* KO-2 cells. The percentage of the 1n (G0/G1) peak in GFP+ cells was 24.3%.

(E) Morphology of chimeric embryo at E10.5 derived from GFP labelling *p53* KO-2 cells. Scale bar, 100 μm.

(F) DNA content analysis of GFP+ cells and GFP- cells in chimeric E10.5 with *p53* KO-2 cells. The percentage of the 1n (G0/G1) peak in chimeric GFP+ cells was 22.3%.

(G-H) Teratoma derived from *p53*-KO haESCs.

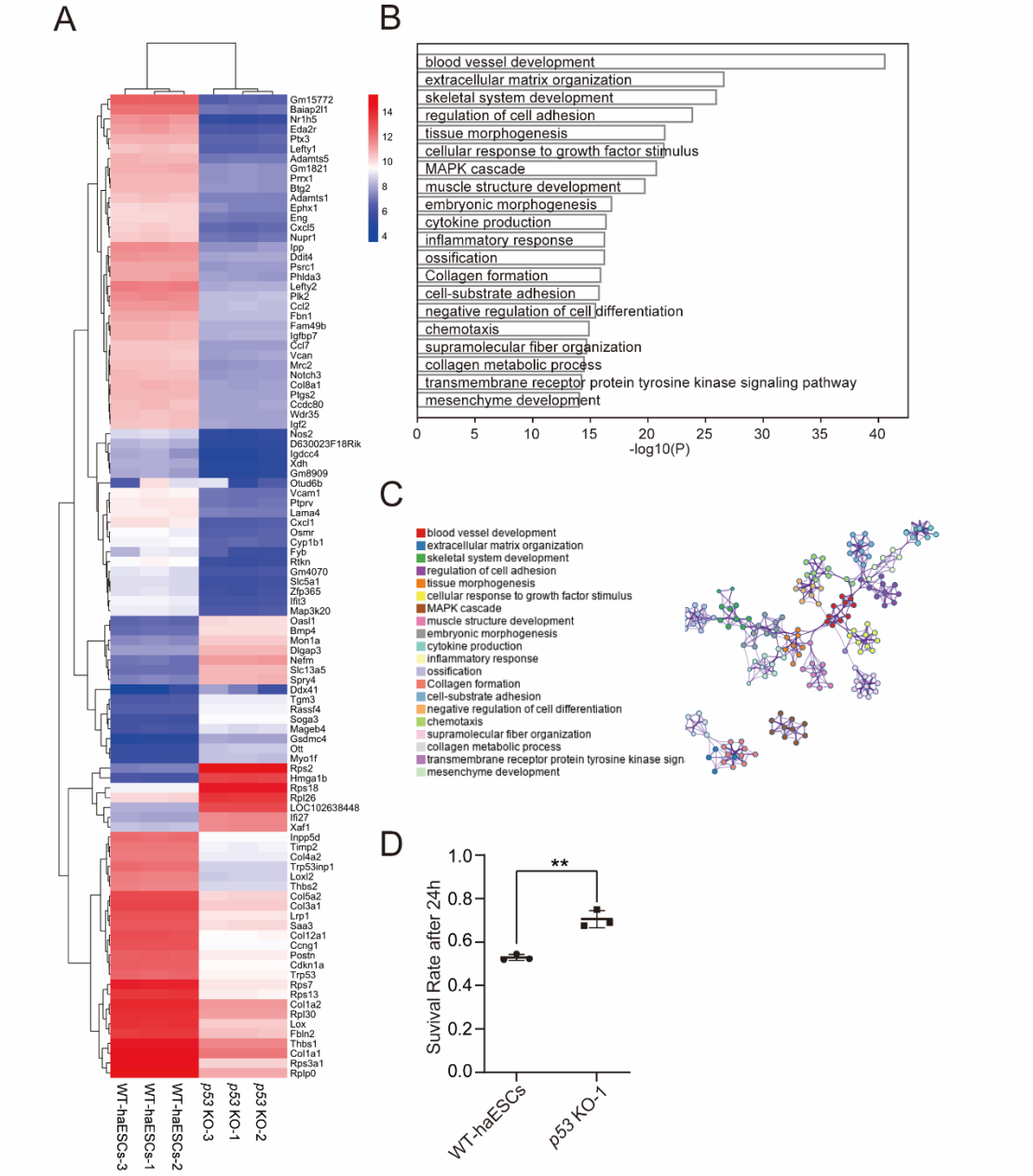


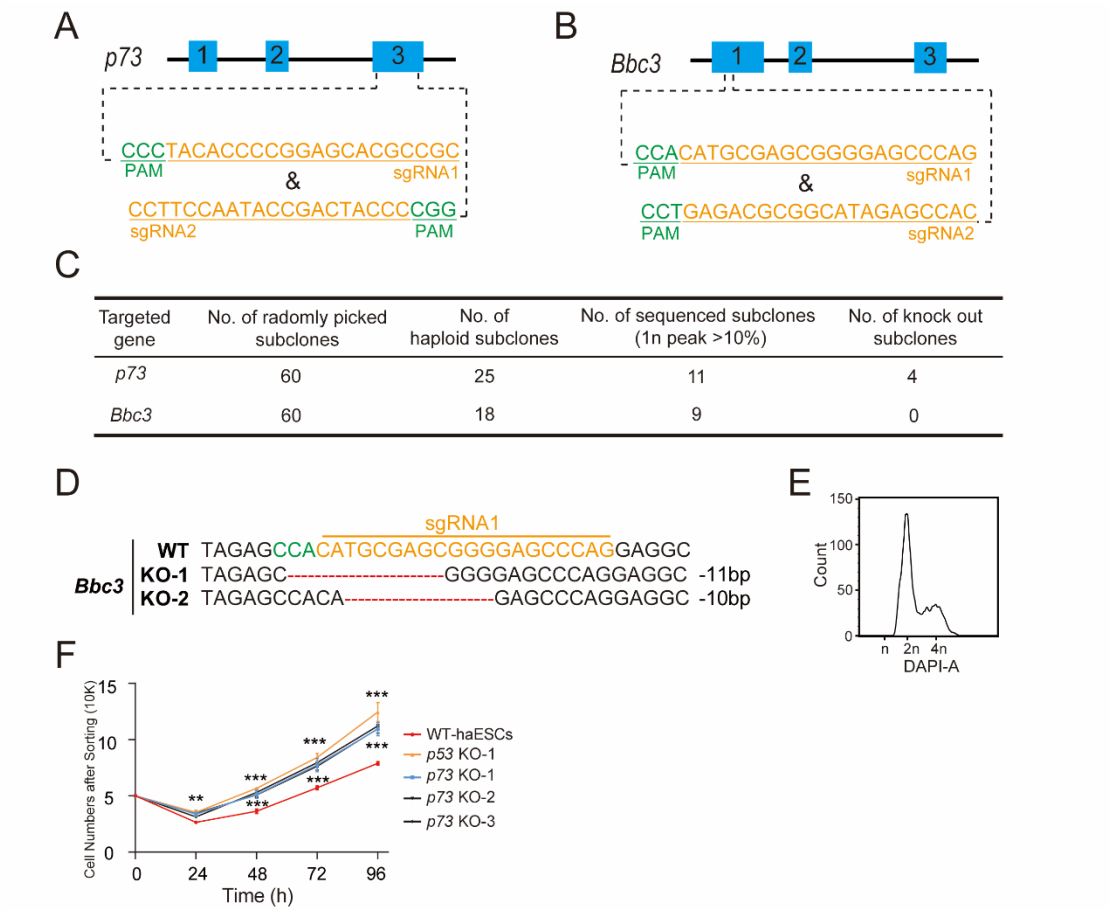
Figure S4. RNA-seq and survival analysis of *p53*-KO haESCs. Related to Figure 4

(A) Hierarchical clustering (heat map) showed the top 100 differentially expressed genes among WT-ESCs and *p53*-KO haESCs, considering a fold change of 2.0 and $p < 0.05$. To observe the differences between samples, we selected the top 50 variant genes.

(B) Enriched terms of differentially expressed genes in *p53*-KO haESCs (more than 2 folds) compared to WT-haESCs, statistical significance was represented by a discrete color scale.

(C) Enrichment network visualization for differential expressed genes between WT-haESCs and *p53*-KO haESCs. Each node represents an enriched term and cluster annotations are shown in color.

(D) Cell survival rates of WT-haESCs and *p53*-KO haESCs for 24 h after sorting. Nearly 70.5% of sorted cells survived in *p53*-KO haESCs, whereas that in WT-haESCs exhibited only 52.9%. (n=3 independent experiments). t test, $**p < 0.01$. Data was represented as the mean ± SEM.



Targeted gene	No. of radomly picked subclones	No. of haploid subclones	No. of sequenced subclones (1n peak >10%)	No. of knock out subclones
p73	60	25	11	4
Bbc3	60	18	9	0

Figure S5. Gene function validation of *p73* and *Bbc3* in haESCS. Related to Figure 5

(A-B) DNA information for knocking out of *p73* and *Bbc3* in WT-haESCs using the CRISPR/Cas9 system.

(C) Summary of haploidy and gene mutations in *p73*-KO and *Bbc3*-KO ESCs. 25/60 subclones of *p73*-KO and 18/60 subclones of *Bbc3*-KO were haploid. The subclones whose 1n peak >10% were sequenced, and there were 4 *p73*-KO subclones but no *Bbc3*-KO subclone.

(D) *Bbc3*-KO genotypes in the mixture *Bbc3*-KO ESCs.

(E) DNA content analysis of *Bbc3*-KO ESCs. The percentage of the 1n peak (G0/G1 phase) was 0%.

(F) Cell counts of *p53* KO-1, *p73* KO-1, *p73* KO-2 and *p73* KO-3 for 24 h, 48 h, 72 h and 96 h after sorting, with WT-haESCs as a control. (n=3 independent experiments). t test, **p < 0.01, ***p < 0.001. Data are represented as the mean ± SEM.

Table S1. Contribution of *p53*-KO ESCs in chimeric embryos. **Related to Figure 3.**

	Percentage of GFP+ cells (%)[a]		Percentage of cells in the 1n peak (%)[b]
	p53-KO haploid	*p53*-KO diploid	*p53*-KO haploid
E6.5	17.2	ND[c]	28.5
	55.5	ND	24.5
E8.5	12.4	24.9	25.3
	37.6	40.3	24.3
E10.5	1.6	ND	21.4
	3.3	ND	22.3

[a]The percentage of GFP+ cells were calculated based on all cells dissected from respective chimeric embryo by FACS analysis.

[b]The percentage of cells in the 1n peak were calculated based on GFP+ cells from respective chimeric embryo by FACS analysis.

[c]ND, not determined.

Table S4. Summary of used primers. Related to Experimental Procedures and Figures 1-5, S1 and S5

Type	Target Name	Sequence 5'-3'
***p53* knockout sgRNA**	sgRNA1-1	CACCGAGGAGCTCCTGACACTCGGA
	sgRNA1-2	AAACTCCGAGTGTCAGGAGCTCCTC
	sgRNA2-1	CACCGACCCTGTCACCGAGACCCCT
	sgRNA2-2	AAACAGGGGTCTCGGTGACAGGGTC
***Bbc3* knockout sgRNA**	sgRNA1-1	CACCGGTGGCTCTATGCCGCGTCTC
	sgRNA1-2	AAACGAGACGCGGCATAGAGCCACC
	sgRNA2-1	CACCGCTGGGCTCCCCGCTCGCATG
	sgRNA2-2	AAACCATGCGAGCGGGGAGCCCAGC
***p73* knockout sgRNA**	sgRNA1-1	CACCGCCTTCCAATACCGACTACCC
	sgRNA1-2	AAACGGGTAGTCGGTATTGGAAGGC
	sgRNA2-1	CACCGGCGGCGTGCTCCGGGGTGTA
	sgRNA2-2	AAACTACACCCCGGAGCACGCCGCC
***p53* overexpression**	*p53*	F: GCTCTAGAATGACTGCCATGGAGGAGTC
	p53	R: CGGAATTCTCAGTCTGAGTCAGGCCCCA
***Cdk2* overexpression**	*Cdk2*	F: GCTCTAGAATGTCGGGTCCGCGCGCG
	Cdk2	R: CGGAATTCTCACTGCTCAATTTCATGGGTGCCAG
***MAPK11* overexpression**	*MAPK11*	F: CGTACCGGTATGGAGAACTTCCAAAAGGTGGA
	MAPK11	R: CGGAATTCTCAGAGCCGAAGGTGGGGC
qPCR	*Gapdh*	F: AGGTCGGTGTGAACGGATTTG
	Gapdh	R: TGTAGACCATGTAGTTGAGGTCA
	Oct4	F: GGATGGCATACTGTGGACCTC
	Oct4	R: TTTCATGTCCTGGGACTCCTCG
	Nanog	F: CCAGGGCTATCTGGTGAACG
	Nanog	R: CCCGAAGTTATGGAGCGGAG
	p53	F: CATGAACCGCCGACCTATCC
	p53	R: GCAGTTCAGGGCAAAGGACT
	p21	F: CCCGAGAACGGTGGAACTTT
	p21	R: AGAGTGCAAGACAGCGACAA
	Cyclin B	F: GACGCAGATAGTGGGGCTG
	Cyclin B	R: GCAATAAACATGGCCGTTACACC
	Cdk2	F: CCTGCTTATCAATGCAGAGGG
	Cdk2	R: TGCGGGTCACCATTTCAGC
	Rex1	F: CCCTCGACAGACTGACCCTAA
	Rex1	R: TCGGGGCTAATCTCACTTTCAT
	Hand1	F: GGCAGCTACGCACATCATCA
	Hand1	R: GCATCGGGACCATAGGCAG
	Cdx2	F: GTCCCTAGGAAGCCAAGTGAA
	Cdx2	R: TTGGCTCTGCGGTTCTGAAA
	FGF5	F: GAAGCGTCTCACTCCCGAAG
	FGF5	R: GAAGAAAACGTCGCGCTACT

	Nestin	F: CCCTGAAGTCGAGGAGCTG
	Nestin	R: CTGCTGCACCTCTAAGCGA
	T	F: GCTTCAAGGAGCTAACTAACGAG
	T	R: CGTCACGAAGTCCAGCAAGA
	Klf4	F: TCGCAGTCATGTGTGTGCTC
	Klf4	R: GCCTCCGGTCCTTCACTAGG
	Otx2	F: ACACACGAGGTAGTGACGC
	Otx2	R: TCTGAGTCGCCGTGGAAAAG
	p73	F: ATGCTTTACGTCGGTGACCC
	p73	R: GCACTGCTGAGCAAATTGAAC
	MAPK11	F: GCGGGATTCTACCGGCAAG
	MAPK11	R: GAGCAGACTGAGCCGTAGG
	Phlda3	F: CCGTGGAGTGCGTAGAGAG
	Phlda3	R: TCTGGATGGCCTGTTGATTCT
	Bbc3	F: AGCAGCACTTAGAGTCGCC
	Bbc3	R: CCTGGGTAAGGGGAGGAGT
	Bax	F: TGAAGACAGGGGCCTTTTTG
	Bax	R: AATTCGCCGGAGACACTCG

Supplemental Experimental Procedures

Mice

All the specific pathogen-free (SPF)-grade mice were purchased from Beijing Vital River Laboratory Animal Technology Co., Ltd. and housed at the Nankai University Animal Center. All experiments were guided by the Institutional Animal Care and Use Committee of Nankai University, and all experiments conformed to the relevant regulatory standards.

Cell Culture

HaESCs were derived from chemically activated oocytes with a 129Sv/Jae genetic background as described previously (Gao et al., 2018). The culture medium of the haESCs referred to a previous protocol (Shuai et al., 2014). For the purification of haploids, haESCs were incubated with 5 μg/ml Hoechst 33342 (Invitrogen, H3570) for 20 min at 37℃. Subsequently, the haploid (1n) peak was gated to sort cells with a diploid control on a Moflo XDP sorter (Beckman). For DNA content analysis, cell samples were fixed with 75% ethanol overnight and then incubated with 50 μg/ml propidium iodide (Sigma, P4170) for 30 min at 37℃. Subsequently, the haploid (1n) peak was analyzed with a diploid control on a BD LSR Fortessa flow cytometer.

Vector Construction

For the *p53*-, *Bbc3*- or *p73*- knockout vector, sgRNAs were designed with reference to the CRISPR design tool (www.crispr.mit.edu), then annealed and inserted into SpCas9n (BB)-2A-GFP (PX461, Addgene, #48140) vectors. For the *p53-, Cdk2*- and *MAPK11*- overexpression vector, the cDNA of *p53, Cdk2* and *MAPK11* were amplified from the total cDNA of mouse ESCs by using Phanta Max Super-Fidelity DNA Polymerase (Vazyme, P505-d2). Then, *p53, Cdk2* and *MAPK11* cDNA was linked in the reconstructive *PiggyBac* Dual promoter (SBI, PB513B-1) separately by using the BamHI and EcoRI sites (*p53*), AgeI and EcoRI sites (*Cdk2*) and XbaI and EcoRI sites (*MAPK11*). All the primers used are listed in Table S2.

Quantitative Real-time PCR

Following the manufacturer's protocol to purify total RNAs from cells using TRIzol Reagent (Invitrogen, 15596018), cDNA was obtained by reverse transcription using a Prime Script™ RT reagent Kit with a gDNA Eraser (Takara, RR047A). Quantitative real-time PCR was performed on an ABI QuantStudio™ 6 Flex machine with FS Universal SYBR Green Master (Roche, 4913914001). Relative expression levels were normalized to *Gapdh*. Averages and SD values were calculated based on three independent experiments. All primers used are listed in Table S2.

Western Blotting

For western blotting experiments, protein samples were extracted from cells by using RIPA lysis solution (Solarbio, R0020) following the manufacturer's protocol. Equal amounts of the samples were separated by SDS-PAGE and transferred to a polyvinylidene fluoride (PVDF) membrane (GE, RPN303F). The membranes were blocked in 5% nonfat dry milk for 2 hours, washed three times with PBST, and incubated with primary antibodies overnight at 4°C. Then, the membranes were washed and incubated with appropriate secondary antibodies for 1 hour at room temperature. The signal was detected with Enlight™ (Engreen, 29050) and imaged. The primary antibodies used in western blotting included anti-P53 (Santa, sc-126), anti-P73 (Abcam, ab40658) and anti-GAPDH (Santa, sc-365062).

T7ENI Assay

Following the manufacturer's recommended protocol for T7 Endonuclease I (NEB, E3321), the DNA fragments containing sgRNA target sites were amplified from transfected cells and wild-type cell genomic DNA by using Phanta Max Super-Fidelity DNA Polymerase (Vazyme). These two kinds of

DNA fragments were hybridized in Buffer 2.0 (NEB). After the hybridization reaction, T7 Endonuclease I was added to the mixture, and digestion was performed for 2 hours at 37 °C. The products were analyzed by agarose gel electrophoresis.

Immunostaining, AP staining, and Karyotype Analysis

For immunofluorescence analysis, cells were fixed with 4% paraformaldehyde at 4°C overnight. The samples were washed three times with PBS. Then, the cells were permeabilized with 0.3% Triton X-100 (Sigma, T9284) for 40 minutes and incubated with 1% BSA (Sigma, B2064) for 40 minutes at room temperature. Next, the samples were incubated with primary antibodies against OCT4 (Abcam, ab18976), SSEA1 (Santa, sc-21702), NANOG (Abcam, ab80892), TFE3 (Sigma, HPA023881), PAX6 (Abcam, ab5790), SOX1 (Abcam, ab87775), NESTIN (Abcam, ab6320), O4 (R&D Systems, MAB1326), NEUN (Abcam, ab104225), GFAP (Abcam, ab7260), MAP2 (Abcam, ab32454) and TUJ1 (Abcam, ab18207) at 4°C overnight. The primary antibodies were diluted in blocking buffer (1% BSA). After three wash steps, the cells were stained with secondary antibodies at room temperature for 1 hour. After another three wash steps, the nuclei were stained with DAPI (YEASEN, 40727ES10) for 10 minutes at room temperature. Immunofluorescence images were taken with a TCS SP8 confocal laser scanning microscope (Leica). AP staining was performed according to the manufacturer's instructions for the alkaline phosphatase kit (Beyotime, P0321). For karyotype analysis, cells were incubated with 0.2 μg/mL nocodazole (MCE, HY-13520) overnight. The samples treated with hypotonic solution were fixed in methanol: acetic acid (3:1) for 20 min twice and dropped onto precooled slides. Then, the cells were stained with Giemsa for 15 min before observation.

Analysis of RNA-seq Data

All the RNA-seq data were sequenced by a local company (Novogene). For data analysis, quality control was performed for the raw sequencing reads using the FastQC application. Raw data were trimmed using Trimmomatic to remove low-quality bases (quality score < 5) or reads (low-quality bases > 50% of the read size)(Bolger et al., 2014). Then, the abundance of each transcript was quantified using Salmon (Patro et al., 2017). Differential gene expression analysis was performed using the DEseq2 package, and gene expression was normalized using the relative-log-expression (RLE) in DESeq2 (Love et al., 2014). Genes with summed reads greater than 10 across all samples were kept. The R function of heatmap was used to generate the hierarchical clustering of the top 50 variant genes. Scatterplots were visualized with ggplot2 and ggrepel.

For Gene Ontology analysis, metascape (http://metascape.org) was employed to perform gene enrichment and functional annotation analyses in both WT ESCs and *p53* KO-1 cells with the following ontology sources: KEGG Pathway, GO Biological Processes, and Reactome Gene Sets(Zhou et al., 2019). Genes with p-values < 0.001, a minimum count of 3, and an enrichment factor > 1.5 were identified and grouped into clusters based on their membership similarities.

Cell Survival Test, Cell Growth Curve and Caspase-3 Activity

For the cell survival test, 50,000 haploid cells were sorted and plated in a 24-well plate under feeder-free conditions. Each group included three repetitions. After 24 hours, the cells were digested into single cells and counted with a cell counter (Logos Biosystems, LUNAII). To describe the cell growth curve, the number of living cells was recorded at 24, 48, 72 and 96 hours. Averages and SD values were calculated based on three independent experiments.

To analyze the rate of diploidization, purified *p53*-KO, *p73*-KO and WT-haESCs were passaged every 2-3 days. Then, the haploid proportions were analyzed at 10 days, 14 days, 28 days and 35 days after sorting and calculated according to a previous report (Xu et al., 2017). Averages and SD values were

calculated based on three independent experiments. For the Caspase-3 activity, 5×10^5 haploid cells were sorted and plated in a 35 mm petri-dish under feeder-free condition. After 24 hours, the Caspase-3 activity was detected by the Caspase 3 Activity Assay Kit (Beyotime, C1116) follow the instructions. Averages and SD values were calculated based on three independent experiments.

Supplemental References

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114-2120.

Gao, Q., Zhang, W., Ma, L., Li, X., Wang, H., Li, Y., Freimann, R., Yu, Y., Shuai, L., and Wutz, A. (2018). Derivation of Haploid Neural Stem Cell Lines by Selection for a Pax6-GFP Reporter. Stem cells and development *27*, 479-487.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology *15*, 550.

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nature methods *14*, 417-419.

Shuai, L., Li, W., Wan, H., Zhao, X.Y., Wang, L., and Zhou, Q. (2014). Generation of Mammalian offspring by haploid embryonic stem cells microinjection. Current protocols in stem cell biology *31*, 1A 6 1-15.

Xu, H., Yue, C., Zhang, T., Li, Y., Guo, A., Liao, J., Pei, G., Li, J., and Jing, N. (2017). Derivation of Haploid Neurons from Mouse Androgenetic Haploid Embryonic Stem Cells. Neuroscience bulletin *33*, 361-364.

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nature communications *10*, 1523.