Derivation of Haploid Trophoblast Stem Cells via Conversion In Vitro

HIGHLIGHTS
A haploid cell line of extraembryonic lineages with self-renewal ability
haiTSCs have multipotency to functional trophoblast lineages both in vitro and in vivo
High-throughput screening of spongiotrophoblast specification-related genes in haiTSCs
Htra1 is a blocker for spongiotrophoblast-specific differentiation

Peng et al., iScience 11, 508–518
January 25, 2018 © 2018 The Author(s).
https://doi.org/10.1016/j.isci.2018.12.014
Derivation of Haploid Trophoblast Stem Cells via Conversion In Vitro

Keli Peng,1,6 Xu Li,1,6 Congyu Wu,2,6 Yuna Wang,1 Jian Yu,3 Jinxin Zhang,1 Qian Gao,1,5 Wenhao Zhang,1 Qian Zhang,1 Yong Fan,4 Yang Yu,5 and Ling Shuai1,7,*

SUMMARY
Owing to their single genome, haploid cells are powerful to uncover unknown genes by performing genetic screening in mammals. However, no haploid cell line from an extraembryonic lineage has been achieved yet, which limits the application of haploid cells in placental genetic screening. Here, we show that overexpression of Cdx2 can convert haploid embryonic stem cells to trophoblast stem cells (TSCs). p53 deletion reduces diploidization during the conversion and guarantees the generation of haploid-induced TSCs (haiTSCs). haiTSCs not only share the same molecular characterization with trophoderm-derived TSCs but also possess multipotency to placental lineages in various procedures. In addition, haiTSCs can maintain haploidy in the long term, assisted by periodic sorting and with reliance on FGF4 and heparin. Finally, we perform piggyBac-mediated high-throughput mutation in haiTSCs and use them in trophoblast lineage genetic screening. Deep sequencing analysis and validation experiments prove that Htra1 is a blocker for spongiotrophoblast specification.

INTRODUCTION
Haploid cells serve as a powerful tool in forward and reverse genetic screening owing to their single-set chromosome feature (Shuai and Zhou, 2014). To date, haploid embryonic stem cells (haESCs) have been achieved in many species assisted by Hoechst 33342 staining and fluorescence-activated cell sorting (FACS) (Leeb and Wutz, 2011; Sagi et al., 2016), which are important for recent gene discovery (Elling et al., 2011; Leeb et al., 2014). Recent derivation of haploid somatic cell lines has facilitated lineage-specific genetic screening (He et al., 2017; Gao et al., 2018). Nevertheless, all haploid cell cultures prefer to double back to diploids, the mechanism of which is still unclear. According to previous reports, the addition of cell cycle inhibitors (Takahashi et al., 2014; He et al., 2017) or editing of specific genes (Olbrich et al., 2017; He et al., 2018) can stabilize the haploid genome and reduce self-diploidization to some degree. However, no haploid cell line has been reported for an extraembryonic lineage. Trophoblast stem cells (TSCs) are one type of placental progenitor cell and are derived from blastocysts or the extraembryonic ectoderm of implantation embryos (Tanaka et al., 1998). They can self-renew by relying on FGF4 and heparin (F4H in vitro and retain the potential to contribute exclusively to the placenta (Oda et al., 2006). Therefore the derivation of TSCs provides considerable insight into the mechanisms that regulate extraembryonic lineage specification and placental development.

Previously, haploid cells were shown to be detectable in preimplantation blastocysts (Liu et al., 2002) and implantation epiblast-stage embryos (Shuai et al., 2015), which indicated that haploid cells were reasonable in trophoderm (TE) lineages. By regulating the expression of Oct4 and Cdx2, embryonic stem cells (ESCs) and TSCs could switch from one to the other easily (Niwa et al., 2005; Wu et al., 2011). In addition, transcriptional induction reprograms somatic cells to functional TSCs (Kubaczka et al., 2015; Benchetrit et al., 2015), suggesting that overexpression of TSC-specific transcription factors can commit cell fate to TE lineages. Hence, haESCs may have the potential to be converted to haploid-induced TSCs (haiTSCs) via overexpression of Cdx2.

Here, we overexpressed Cdx2 in haESCs by using a Tet-On inducible system to alternate cell fate. We demonstrated that haiTSCs were generated from p53-deleted haESCs in vitro under defined conditions. haiTSCs maintained haploidy and contributed to the placenta in a chimeric experiment, proving that they potentially differentiated into functional trophoblast terminal cells. Then we performed a proof-of-concept screening in haiTSCs to identify key genes regulating spongiotrophoblast specification.
RESULTS

Overexpression of Cdx2 Converts haESCs to TSCs

To generate haiTSCs from haESCs by conversion in vitro, we adopted an inducible overexpression strategy. Parthenogenetic haESC lines were established from 129Sv/Jae background chemical-activated oocytes, and one line with a high percentage of haploid cells was chosen to perform subsequent experiments (Figure 1A). We then designed two piggyBac (PB) vectors to introduce Tet-On inducible Cdx2 overexpression into haESCs: vector 1 had the rtTA (Tet-On Advanced transactivator) and neomycin selection genes, driven separately by an SV40 promoter, and vector 2 had the Cdx2 and puromycin resistance genes, driven by a tetracycline response element with a minimal cytomegalovirus promoter (Figure S1A).

We transfected these two PB vectors and a PBase vector into haESCs by electroporation. Transfected cells were selected in 2i/L (inhibitor PD0325901, inhibitor CHIR99021, and mLif) medium (Ying et al., 2008), supplemented with G418 (250 μg/mL) for 6 days. To evaluate the pluripotency of the transfected haESCs (which we termed OE-Cdx2 haESCs), we performed immunofluorescent staining of pluripotent markers and alkaline phosphatase (AP) staining. The results showed that OE-Cdx2 haESCs were positive not only for Oct4, Nanog, and SSEA1 (Figure 1B) but also for AP (Figure 1C), which demonstrated that vector insertion did not jeopardize haESC pluripotency. Thereafter we cultured newly sorted haploid OE-Cdx2 haESCs for 23 days in 70% CM + DOX (70CM+DOX) medium (Figure 1D).

We performed qPCR analysis of pluripotent marker genes (Oct4 and Rex1) in iTSCs, WT-TSCs, and WT-ESCs (Figure 1F). The expression levels of Oct4 and Rex1 were significantly higher in WT-TSCs than in WT-ESCs (t test, ***p < 0.001, mean ± SEM). We also performed qPCR analysis of TSC marker genes (Eomes and Elf5) (Figure 1G). The expression levels of Eomes and Elf5 were significantly higher in iTSCs than in WT-TSCs (t test, ***p < 0.001, mean ± SEM).

DNA content analysis of iTSCs derived from haESCs indicated that there were no haploid cells in the iTSCs (Figure 1H). We also performed chromosome spreads of iTSCs, each single cell spread had 40 chromosomes (Figure 1I).

See also Figures S1 and S2.
haESCs in standard TSC culture medium supplemented with doxycycline (Dox) and puromycin to induce Cdx2 overexpression (Figure 1D). Obvious morphological change was observed 5 days after Dox induction, and many cells died during puromycin selection. Approximately 11 days after induction, typical TSC-like colonies were formed (Figure 1E) and expanded with trypsin, which meant that an inducible TSC (iTSC) line was established. Immunofluorescence results revealed that iTSCs expressed the TSC-specific markers Cdx2 and Eomes, rather than the ESC marker Oct4 (Figure 1F). Quantitative PCR (qPCR) results confirmed that iTSCs did not express pluripotent genes (Oct4 and Rex1) relative to wild-type (WT) ESCs (Figure 1I) and showed similar expression levels of TSC markers (Eomes and Elf5) relative to WT-TSCs (Figure 1G). Therefore iTSCs were derived by conversion of haESCs in vitro through Cdx2 overexpression.

We utilized Hoechst 33342 staining and FACS to isolate haploid cells during conversion. After optimization, we first sorted the haploid cells on day 11 post conversion; the 1n peak (haploid cells) was 4.38% and was further expanded with good viability (Figure S1B). There was no haploid cell left among the iTSCs according to the second round of sorting by FACS (Figures 1H and S2C) and chromosome spread analysis (Figure 1I). We reasoned that different PB insertions might affect haploid maintenance ability; thus we randomly picked six subclones from among OE-Cdx2 haESCs. We genotyped the subclones, and the results suggested that all subclones carried Cdx2 and rT3A (Figure S1C). Among the six subclones, #1 and #2 were stable in terms of haploid maintenance and were further assessed (Figure S2A). Although #1 and #2 carried a few insertions (Figure S1D) and could be converted to typical iTSCs easily (Figure S2B), neither of them could generate haiTSC lines (Figure S2C). We reasoned that the nature of diploidization in haESCs during conversion hindered the derivation of haiTSCs.

**Deletion of p53 Facilitates Derivation of HaiTSCs**

A previous study showed that p53 gene deficiency could stabilize haploidy in mouse haESCs by promoting the viability and proliferation of haploid cells in daily culture (Olbrich et al., 2017). To achieve haiTSCs in our OE-Cdx2 system, we knocked out p53 in OE-Cdx2 haESCs through CRISPR/Cas9-mediated non-homologous end joining. We transfected plasmids carrying Cas9-GFP and two single guide RNAs (sgRNAs, targeting the third exon of p53) into OE-Cdx2 haESCsubclone #1, which had a high percentage of haploid cells (Figure 2A). Approximately 36 hr after transfection, haploid cells expressing Cas9-GFP were enriched by FACS (Figure S3A). To address whether p53 deletion occurred, we randomly picked four subclones (we termed them PO1, PO2, PO3, and PO4) and performed T7 endonuclease I (T7EN1) cleavage analysis. The results showed that all subclones underwent gene editing (Figure S3C). Further sequencing results confirmed that all four subclones carried mutations with small deletions at the target sites (Figure S3D). We also detected p53 at the protein level in PO1 and PO4 by western blot and found that the p53 protein was absent in these subclones (Figure 2B). Furthermore, we analyzed the expression levels of p53-related genes in PO1 and PO4 by qPCR relative to a WT-ESC line. p53 and P21 in PO1 and PO4 were downregulated, whereas other cell cycle genes, including Reprimo, Cdk1, and Cyclin B, exhibited no significant differences (Figure 2C). Taken together, the results indicated that p53 deletion was successfully realized in OE-Cdx2 haESCs.

To determine whether p53 mutation could stabilize haploidy in OE-Cdx2 haESCs, we used a serum ESC medium without 2i (Elling et al., 2011) to trigger severe diploidization. We cultured PO1, PO2, PO3, and PO4 separately in this serum ESC medium for five passages without sorting and found that the PO4 cell line maintained haploidy at a very high percentage relative to the other subclones (Figure S3E). In another repeat trial, we cultured PO4 and #1 in 2i/L medium; 40.2% of PO4 cells remained in the 1n peak (haESCs at G0/G1 phase), whereas the 1n peak percentage in #1 was reduced to 16% during the same period (Figure S3F). Given that PO4 with p53 deletion maintained haploidy in a steady manner, we performed conversion with PO4 to derive haiTSCs with Dox and puromycin monitoring. After conversion for approximately 11 days and three subsequent rounds of haploidy purification by FACS (Figure S3G), a haploid cell line was established and showed mostly one set of chromosomes by FACS (Figure 2D) and karyotype analysis (Figure 2E). To enrich authentic haiTSCs from these haploid derivatives, we purified them with the TSC-specific antibody CDCP1 (Rugg-Gunn et al., 2012). The FACS results showed that approximately 39.7% of cells were already CDCP1 positive (Figure 2F); these cells were harvested and further cultured in either feeder or feeder-free culture systems with typical TSC colony morphology (Figure 2G). Immunofluorescence staining results revealed
that haiTSCs expressed the TSC-specific markers Cdx2 and Eomes instead of the ESC marker Oct4 (Figure 2H).

**Molecular Characterization and Differentiation Potentials of haiTSCs**

To assess the purity of haiTSCs after CDCP1-positive cell sorting, we passaged them several times and reanalyzed them with the CDCP1 antibody by FACS. CDCP1-positive cells remained at a high percentage (92.3%) in haiTSCs, which meant that haiTSCs could sustain TE identity during self-renewal (Figure 3A).

The CDCP1-antibody-purified haiTSCs were expanded several times (three passages), and 15.3% of
Figure 3. Identification of haiTSCs Properties and Differentiation Potential

(A) FACS analysis of CDCP1-positive cells among established haiTSCs. haESCs are used as negative control.

(B) The expression levels of TSC-specific marker genes (Cdx2, Eomes, Elf5, and Tfap2c) in haiTSCs, WT-TSCs, haESCs, and WT-ESCs. t test, **p < 0.01, ***p < 0.001. Data are represented as mean ± SEM.

(C) Western blot analysis of Eomes and Cdx2 in WT-ESCs, haESCs, WT-TSCs, and haiTSCs. GAPDH is used as a loading control.

(D) DNA methylation status in the promoter regions of Oct4, Nanog, and Elf5. An haESCs line and a WT-TSCs line were used as controls.

(E) Global gene expression cluster analysis of transcripts in WT-ESCs, haESCs, WT-TSCs, and haiTSCs.

(F) Images of differentiated haiTSCs. Red arrow indicates syncytiotrophoblast cells, and yellow arrow indicates trophoblast giant cells. Scale bar, 100 µm.

(G) Immunofluorescence staining of three trophoblast-lineage-specific markers, Tpbpa (fluorescein isothiocyanate channel), proliferin (Tetramethylrhodamine [TRITC] channel), and GCM1 (TRITC channel), in cells differentiated from haiTSCs in vitro. DNA is stained with DAPI. Scale bar, 25 µm.

(H) Immunohistochemical (IHC) analysis of hemorrhagic lesions derived from haiTSCs. Fixed sample is IHC stained against the endothelial marker CD31 and the trophoblast marker Tfap2c.

(I) Images of BF (top) and GFP (bottom) of a chimeric placenta following blastocyst injection of haiTSCs-eGFP cells. Scale bar, 2 mm.

(J) Immunofluorescence staining of chimeric placenta following sectioning, with trophoblast-lineage-specific antibodies against proliferin and Tpbpa. Scale bar, 25 µm.

See also Figures S4 and S5.
haploid cells remained (Figure S4A), indicating that haiTSCs could maintain haploidy in a TSC-specific manner. qPCR further confirmed that the expression levels of TSC-specific genes (Cdx2, Eomes, Elf5, and Tfpap2c) were high in haiTSCs and WT-TSCs (Figure 3B), whereas expression levels of pluripotent genes (Oct4 and Nanog) in haiTSCs and WT-TSCs were much lower than those in haESCs and WT-ESCs (Figure S4B). Western blot results also suggested that haiTSCs and WT-TSCs were positive for Cdx2 and Eomes relative to ESCs (Figure S3C). DNA-methylation-mediated gene regulation is crucial in the determination of specific transcription factors, which are extensively utilized to judge cell identities (Wu et al., 2011). We analyzed the DNA methylation status of the Oct4, Nanog, and Elf5 promoters in haiTSCs by performing bisulfite sequencing, with haESCs and WT-TSCs as controls. Results revealed that the Elf5 promoter in haiTSCs was hypomethylated, whereas the Oct4 and Nanog promoters in haiTSCs were hypermethylated (Figure 3D), which were consequent with previous report (Ng et al., 2008). To elucidate the properties of haiTSCs on the transcriptome scale, we analyzed the global RNA levels of haiTSCs by random differentiation. To investigate the in vivo differentiation potential of haiTSCs, we cultured the haiTSCs by withdrawing FGF4, heparin, and mouse embryonic fibroblasts in vitro. Obvious polyploid trophoblast giant cells (TGCs) and syncytiotrophoblast (Syt) cells were observed (Figure 3F) and increased gradually (Figure S5A). Immunofluorescence staining also revealed that the spongiosotrophoblast-cell-specific marker Tbpba, the TGC-specific marker proliferin, and the labyrinth-progenitor-specific marker Gcm1 were observed in differentiated cells from haiTSCs (Figure 3G). We analyzed trophoblast-lineage-specific gene expression levels of differentiated cells (4 days) from haiTSCs, iTSCs (diploid), and WT-TSCs by qPCR. Accordingly, differentiated cells from haiTSCs expressed all lineage-specific genes; in particular, Ctsq, Pkl2d1, and Tpbpa increased significantly (Figure S5B). Taken together, the results indicated that haiTSCs could differentiate into diverse trophoblast lineage cells in vitro by random differentiation. To investigate the in vivo differentiation potential of haiTSCs, we transplanted approximately 1 × 10⁶ haiTSCs into the testis of ICR mice, with WT-TSCs and ESCs as parallel controls. After 3 weeks, hemorrhagic lesions formed in the haiTSC and WT-TSC groups, whereas teratomas formed in the ESC group (Figure S5C). Blood vessels with TGC invasion were confirmed in haiTSC-derived hemorrhagic lesions by immunohistochemical staining of Tfpap2c and CD31 (Figure 3H), indicating that haiTSCs could mimic placental development by undergoing differentiation in vivo. Next, we microinjected GFP-labeled haiTSCs into blastocysts to test whether they could contribute to functional placenta. In the reconstructed blastocysts, the GFP-positive haiTSCs integrated into the TE instead of the inner cell mass (ICM), showing their trophoderm nature (Figure S5D). To evaluate further development, the reconstructed embryos were transferred to the uteruses of pseudopregnant mice. On embryonic day 10.5 (E10.5), placentas and embryos were dissected from the pseudopregnant mice. We found GFP-positive cells contributing to the placentas (Figure 3I), and these cells were positive for proliferin and Tpbpa antibody staining according to fluorescence histology analysis (Figure 3J). Our findings revealed that haiTSCs could not only mimic placental development in vivo but also contributed to functional placenta, which are typical features of trophoblast progenitor cells.

piggyBac Transposon-Mediated High-Throughput Gene Trapping in haiTSCs

Haploid cells are easy to use for generating genome-wide homozygous mutant libraries via PB integration or virus infection (Li and Shuai, 2017). To determine the feasibility of using haiTSCs in this process, approximately 1 × 10⁷ haiTSCs were transfected with PB-based gene trap vectors (Leeb and Wutz, 2011) carrying a puromycin resistance (Puro⁺) gene (Figure 4A). After 4 days of puromycin selection, integrated haiTSCs with the Puro⁺ gene survived, and the control group mostly died (Figure 4B). We analyzed the mutant haiTSCs by FACS and found that 38.1% of haploid cells remained, suggesting high efficiency for obtaining homozygous mutant haiTSCs (Figure 4C). We harvested mutant haiTSCs and amplified the integration sites by performing splinkerette PCR. Obvious strands demonstrating integration were visualized (Figure 4D). We linked the products into plasmids and sequenced them by Sanger sequencing. In total, 50 different sites were addressed, of which 24 were located inside the gene body (Figures S6A
and S6B). These data indicate that haiTSCs could undergo gene manipulation and generate homozygous mutant libraries.

To apply haiTSCs in trophoblast lineage-specific genetic screening, we focused on key genes regulating spongiotrophoblast differentiation and utilized a specific antibody against Tpbpa (Latos and Hemberger, 2016) as indicated. Briefly, we performed random differentiation in mutated haiTSCs, as previously described, and nonmutated haiTSCs separately for 3 days. Tpbpa-positive cells demonstrating spongiotrophoblast features were analyzed by immunostaining and FACS (Figure 4E). If the percentage of Tpbpa-positive cells increased significantly in the mutated haiTSCs group relative to the nonmutated haiTSCs group, the mutated Tpbpa-positive cells were harvested with a FACS-assisted antibody for further bioinformatics analysis. We performed this experiment several times and deep-sequenced two repeats with the Tpbpa-positive cells increasing group (Figures 4E, S6C, and S6D). According to deep sequencing, approximately 4 million independent insertions across more than 20,000 genes were identified, of which 49.8% were derived from the sense orientation (Figure 5A). In addition, approximately 57% of the insertions landed in intergenetic regions (coding regions + intron + 5’/3’ UTR), whereas 39% of insertions landed in intragenetic regions. Enrichment analysis with gene ontology databases showed that insertions preferred genes carrying specific functions for epigenetic modifications, such as repressing transcription factor binding and DNA methyltransferase activity (Figure 5C). Ten genes, including Zfp704, Htra1, and Rsf1, were identified (Figure 5D) due to both frequent insertions determined by PB screening and higher transcription activity measured from RNA-seq (Figures 5E and S7). Of the top candidates, Htra1 was chosen for further validation experiments in WT-TSCs (Figure 5E).

**Htra1 As a Blocker for Spongiotrophoblast Differentiation**

To test whether Htra1 is an essential block against spongiotrophoblast differentiation in TSCs, we designed specific sgRNAs to knock out Htra1 in WT-TSCs with the CRISPR/Cas9 system (Figure 6A).
transfected TSCs with a Cas9-sgRNAs vector (targeted), and control TSCs were transfected with an empty vector (mock). We performed random differentiation independently with the targeted TSCs (gene edited) and mock TSCs (non-gene edited) for 3 days. Spongiotrophoblast cells were observed by morphology (Figure 6B) analysis and Tpbpa antibody immunostaining (Figure 6C). According to a FACS analysis, the percentage of Tpbpa-positive cells in the targeted group significantly increased relative to the mock TSCs group (Figure 6D). In a parallel experiment, we analyzed the expression levels of trophoblast-lineage-specific markers in differentiated cells from the targeted and mock groups and found that the expression level of the spongiotrophoblast-specific marker Tpbpa in the targeted group increased significantly relative to the mock group, whereas the other two lineage markers in the targeted group decreased (Figure 6E). We tested the genotypes of targeted TSCs, and the results indicated that the cells carried a gene mutation in \textit{Htra1} (Figure 6F). Thus \textit{Htra1} plays a very important role in spongiotrophoblast specification.

**DISCUSSION**

In this study, we demonstrate that haiTSCs can be generated by the conversion of haESCs in vitro and maintain haploidy well, with the potential to differentiate into placental cells. We utilized a Dox-inducible Tet-On system to introduce \textit{Cdx2} overexpression in haESCs, which was reported to be a key regulator of TE development (Strumpf et al., 2005). Although a severe diploidization phenomenon was observed during ESC-TSC conversion, we adopted two strategies to establish haiTSCs: (1) knock out p53 in the initial OE-Cdx2 haESCs to stabilize the haploid genome and (2) modify the TSC-specific culture system to promote cell fate alternation (including specific growth factors and inhibitors, TSC-specific antibody sorting, etc.).
deletion and guaranteed their potential in a subsequent gene trapping procedure. The results proved that the p53 gene was a vital regulator of diploidization, which was consistent with a previous report (Olbrich et al., 2017). However, the exact mechanism by which p53 triggers diploidization remains unknown. In addition to gene modification, we optimized the TSC culture medium with a serum-free system, named TX medium (Kubaczka et al., 2014). Haploidy was better sustained in TX medium than in traditional TSC (7CM) medium (Tanaka et al., 1998), potentially due to some unknown ingredients in serum (data not shown). In another parallel experiment, we attempted to derive haploid TSCs from haploid blastocysts in TX medium. However, no haploid cell line was established, mainly because embryonic TSCs tended to diploidize more severely. This leaves an open question: can haploid TSCs be established directly from haploid embryos and what is the mechanism underlying the thorough diploidization of haploid TE? We can compare the transcriptome differences between haiTSCs and haploid TEs from blastocysts in a future study.

As our haiTSCs were generated from haESCs in vitro, their purity and TSC properties needed to be confirmed. CDCP1 is a TSC-specific marker that has been widely used in research related to trophoblast cell fate determination (Rugg-Gunn et al., 2012; Nosi et al., 2017). Although our haiTSCs were established from an inducible system, a high percentage of CDCP1-positive cells was obtained (Figure 3A) after two rounds of FACs, and these cells maintained TSC properties. haiTSCs were maintained in a haploid state assisted by Hoechst 33342 purification and possessed differentiation potential to mimic placental...
development (Figures S5C and S3G), even contributing to functional placenta according to a chimeric experiment (Figures 3H and 3I). These features make haiTSCs perfect tools to evaluate the function of recessive genes related to placental development. Recently, the derivation of human TSCs was reported, raising extensive concerns regarding the study of placental abnormality in diseases associated with abortion (Okae et al., 2018). In this study, haiTSCs served as a platform to screen essential placental development genes through reverse genetic screening. Abundant mutations were introduced into haiTSCs by the PB transposon (Figures 4D, S6A and S6B), consistent with a previous report (Wang et al., 2018). Stable maintenance of haploidy by haiTSCs facilitated the production of homozygous mutant libraries (Figure 4C), which are quite valuable for genetic screening. Therefore, in this study, we used haiTSCs in high-throughput screening to identify Htra1 as a crucial gene regulating trophoblast specification. Early cell fate determination induces blastomeres to divide into the TE and ICM (Vogel, 2005; Takaoka and Hamada, 2012); however, the molecular mechanism remains unclear. haiTSCs can also serve as a tool for the discovery of key genes regulating TE and ICM interactions.

In summary, we converted haESCs to haiTSCs by controlling the expression level of Cdx2 with a Dox-inducible system. Although diploidization accompanied the conversion process, deletion of the p53 gene reduced diploidization to some degree and facilitated the derivation of haiTSCs. haiTSCs not only showed standard TSC colonies and expressed TSC-specific genes but also held the potential to differentiate into placental cells. Therefore these haploid extraembryonic progenitor cells show great advantages for trophoblast lineage genetic screening.

Limitations of the Study
In this experiment, we adapted p53 knockout strategy to establish authentic haploid extraembryonic cell line in a haploid nature. Whether the p53 knockout method could also benefit derivation of haploid TSCs from haploid blastocysts remains unknown. Meanwhile, p53 is an essential gene related to many important biological processes, including mitosis (Cross et al., 1995) and cancer (Li et al., 2010). Strikingly, p53 knockout rescued the viability of tetraploid ESCs and enabled the generation of late-stage mouse tetraploid embryos (Hori et al., 2015). Whether p53 knockout genotype would affect some unknown function of haiTSCs warrants more investigations.

METHODS
All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION
Supplemental Information includes Transparent Methods, seven figures, and one table and can be found with this article online at https://doi.org/10.1016/j.isci.2018.12.014.

ACKNOWLEDGMENTS
We acknowledge Dr. Xudong Wu from Tianjin Medical University for cell sorting. This work was supported by the National Key Research and Development Program of China (2018YFC1004101 to L.S.), the National Natural Science Foundation of China (31501186, 31671538, and 31872841 to L.S.), the Natural Science Foundation of Tianjin (15JCZDJC65300 to L.S.), and the Fundamental Research Funds for the Central Universities.

AUTHOR CONTRIBUTIONS
L.S. designed and supervised this project. K.P., X.L., Y.W., J.Z., Q.G., W.Z., and Q.Z. performed the experiments. J.Y. and C.W. analyzed the bioinformatics. L.S., Y.F., and Y.Y. wrote the manuscript with contribution by all authors.

DECLARATION OF INTERESTS
The authors declare no competing interests.
REFERENCES

Benchetrit, H., Herman, S., van Wietmarschen, N., Wu, T., Makedonska, K., Maoz, N., Yom Tov, N., Stave, D., Lasry, R., and Zayat, V. (2015). Extensive nuclear reprogramming underlies lineage conversion into functional trophoblast stem-like cells. Cell Stem Cell 17, 543–556.

Cross, S.M., Sanchez, C.A., Morgan, C.A., Schimke, M.K., Ramel, S., Idzerda, R.L., Raskind, W.H., and Reid, B.J. (1995). A p53-dependent mouse spindle checkpoint. Science 267, 1333–1336.

Elling, U., Taubenschmid, J., Wirsberger, G., O’Malley, R., Demers, S.P., Vanhaelen, O., Shukalyuk, A.I., Schmauss, G., Schramek, D., Schnuettgen, F., and von Melchner, H. (2011). Forward and reverse genetics through derivation of haploid mouse embryonic stem cells. Cell Stem Cell 9, 563–574.

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 panel of Pa6-GFP reporter. Stem Cells Dev. 27, 479–487.

He, W., Zhang, X., Zhang, Y., Zheng, W., Xiong, Z., Hu, X., Wang, M., Zhang, L., Zhao, K., Qiao, Z., and Lai, W. (2018). Reduced self-diploidization and improved survival of semi-cloned mice produced from androgenic haploid embryonic stem cells through overexpression of Dmnt3. Cell Stem Reports 10, 477–493.

Horii, T., Yamamoto, M., Morita, S., Kimura, M., Nagao, Y., and Hatada, I. (2015). p53 suppresses tetraploid development in mice. Sci. Rep. 5, 8907.

Kubaczka, C., Senner, C., Arauzo-Bravo, M.J., Sharma, N., Kuckenberg, P., Becker, A., Zimmer, A., Brustle, O., Peitz, M., Hemberger, M., and Schorle, H. (2014). Derivation and maintenance of murine trophoblast stem cells under defined conditions. Stem Cell Reports 2, 232–242.

Kubaczka, C., Senner, C.E., Cierlitza, M., Arauzo-Bravo, M.J., Kuckenberg, P., Peitz, M., Hemberger, M., and Schorle, H. (2015). Direct induction of trophoblast stem cells from murine fibroblasts. Cell Stem Cell 17, 557–568.

Latos, P.A., and Hemberger, M. (2016). From the stem of the placental tree: trophoblast stem cells and their progeny. Development 143, 3650–3660.

Leeb, M., Dietmann, S., Paramor, M., Niwa, H., and Smith, A. (2014). Genetic exploration of the exit from self-renewal using haploid embryonic stem cells. Cell Stem Cell 14, 385–393.

Leeb, M., and Wutz, A. (2011). Derivation of haploid embryonic stem cells from mouse embryos. Nature 479, 131–134.

Lim, M., Fang, X., Baker, D.J., Guo, L., Gao, X., Wei, Z., Han, S., van Deursen, J.M., and Zhang, P. (2010). The ATM-p53 pathway suppresses aneuploidy-induced tumorigenesis. Proc. Natl. Acad. Sci. U S A 107, 14188–14193.

Li, Y., and Shuai, L. (2017). A versatile genetic tool: haploid stem cells. Stem Cell Res. Ther. 8, 197.

Ng, R.K., Dean, W., Dawson, C., Lucifero, D., Madeja, Z., Reik, W., and Hemberger, M. (2008). Epigenetic restriction of embryonic cell lineage fate by methylation of ElF5. Nat. Cell Biol. 10, 1280–1290.

Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., and Rossant, J. (2005). Interaction between Oct3/4 and Cdx2 determines trophoderm differentiation. Cell 123, 917–929.

Nou, U., Lanner, F., Huang, T., and Cox, B. (2017). Overexpression of trophoblast stem cell-enriched MicroRNAs promotes trophoblast fate in embryonic stem cells. Cell Rep. 19, 1110–1109.

Oda, M., Shiotani, K., and Tanaka, S. (2006). Trophoblast stem cells. Methods Enzymol. 419, 387–400.

Okae, H., Toh, H., Sato, T., Hura, H., Takahashi, S., Shirane, K., Kabayama, Y., Suyama, M., Sasaki, H., and Arima, T. (2018). Derivation of human trophoblast stem cells. Cell Stem Cell 22, 50–63.

Olbrich, T., Mayor-Ruiz, C., Vega-Sendino, M., Gomez, C., Ortega, S., Ruiz, S., and Fernandez-Capetillo, O. (2017). A p53-dependent response limits the viability of mammalian haploid cells. Proc. Natl. Acad. Sci. U S A 114, 9367–9372.

Rugg-Gunn, P.J., Cox, B.J., Lanner, F., Sharma, P., Ignatchenko, V., McDonald, A.C., Garner, J., Gramallo, A.O., Rossant, J., and Kislinger, T. (2012). Cell-surface proteomics identifies lineage-specific markers of embryo-derived stem cells. Dev. Cell 22, 887–901.

Sagi, I., Chia, G., Golan-Lev, T., Peretz, M., Weissbein, U., Sui, L., Sauer, M.V., Yanuka, O., Egli, D., and Benvenisty, N. (2016). Derivation and differentiation of haploid human embryonic stem cells. Nature 532, 107–111.

Shuai, L., Wang, Y., Dong, M., Wang, X., Sang, L., Wang, M., Wan, H., Luo, G., Gu, T., Yuan, Y., et al. (2015). Durable pluripotency and haploidy in epiblast stem cells derived from haploid embryonic stem cells in vitro. J. Mol. Cell Biol. 7, 326–337.

Shuai, L., and Zhou, Q. (2014). Haploid embryonic stem cells serve as a new tool for mammalian genetic study. Stem Cell Res. Ther. 5, 20.

Strumpf, D., Mao, C.A., Yamanaka, T., Balston, A., Chawengsaksophak, K., Beck, F., and Rossant, J. (2003). Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. Development 132, 2093–2102.

Takahashi, S., Lee, J., Kohda, T., Matsuzawa, A., Kawasumi, M., Kanai-Azuma, M., Kaneko-Ishino, T., and Ishino, F. (2014). Induction of the G2/M transition stabilizes haploid embryonic stem cells. Development 141, 3842–3847.

Takahashi, K., Yagi, R., and Rossant, J. (2005). Embryology. Embryologists can now make mouse embryos from cells in the blastocyst. Science 308, 782–783.

Wang, H., Zhang, W., Yu, J., Wu, C., Gao, Q., Li, X., Li, Y., Zhang, J., Tian, Y., Tan, T., et al. (2018). Genetic screening and multipotency in rhesus monkey haploid neural progenitor cells. Development 145, https://doi.org/10.1242/dev.160531.

Wu, T., Wang, H., He, J., Kang, L., Jiang, Y., Liu, J., Zhang, Y., Kou, Z., Liu, L., Zhang, X., and Gao, S. (2011). Reprogramming of trophoblast stem cells into pluripotent stem cells by Oct4. Stem Cells 29, 755–763.

Ying, Q.L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature 453, 519–523.
Supplemental Information

Derivation of Haploid Trophoblast Stem Cells via Conversion *In Vitro*

Keli Peng, Xu Li, Congyu Wu, Yuna Wang, Jian Yu, Jinxin Zhang, Qian Gao, Wenhao Zhang, Qian Zhang, Yong Fan, Yang Yu, and Ling Shuai
Figure S1 related to Figure 1

(A) Schematic overview of the vectors used for overexpression of Cdx2. There are three vectors: vector 1 has the rtTA and Neo selection genes; vector 2 has the Cdx2 coding region and Puromycin resistance gene driven by TRE with a minimal CMV promoter; vector 3 is used as a PBase expression vector.

(B) DNA content analysis of the cells on Day 11 (first sorting) and Day 23 (second sorting) during conversion.

(C) Genotype of the six subclones from OE-Cdx2 haESCs with Cdx2 and rtTA indicated.

(D) Sequencing traits of insertion sites of PB in #1 and #2.
Figure S2 related to Figure 1

(A) DNA content analysis of two stable haploid transgenic OE-Cdx2 cell lines, #1 and #2.

(B) The morphology of iTSCs from subclones #1 and #2. Scale bar, 100 μm.

(C) DNA content analysis of the cells converted from #1 and #2 on Day 23 (second sorting).

(D) Immunofluorescence staining of Oct4 (TRITC channel), Eomes (FITC channel) and Cdx2 (FITC channel) in iTSCs. DNA is stained with DAPI. Scale bar, 50 μm.
Figure S3 related to Figure 2

(A) Percentage of Cas9-GFP-positive cells in #1 by FACS 2 days after transfection.
(B) sgRNA sequences used for P53 knockout.
(C) P53-deleted genotypes in subclones PO1, PO2 and PO3.
(D) Percentages of the 1n peaks of four P53 knockout subclones, PO1, PO2, PO3 and PO4, cultured in serum ES medium for 5 passages.
(E) T7EN1 cleavage analysis of the P53 knockout cell lines PO1, PO2, PO3 and PO4. Cleaved products (red triangle) indicate the presence of mutations.
(F) DNA content analysis of #1 and PO4 cultured in 2i/L medium.
(G) DNA content analysis of cells during conversion on Day 0, Day 11, and Day 23 and final established haiTSCs. The percentages of the 1n (G0/G1) peak were 69.4%, 53.2%, 6.59%, and 70.4%, respectively.
Figure S4 related to Figure 3

(A) DNA content analysis of haiTSCs after CDCP1 antibody sorting and expansion. The percentage of the In (G0/G1) peak was 15.3%.

(B) Expression level analysis of pluripotent marker genes (Oct4 and Nanog) in haiTSCs, WT-TSCs, haESCs and WT-ESCs by qPCR. t test, ***p < 0.001. Data are represented as mean ± SEM.

(C) Heatmap of the top 50 most variable genes across all four cell types. Data are first scaled by row and then represented using hierarchical clustering for both samples and genes.

(D) Immunofluorescence staining of Cdx2 (FITC channel) and H3K27me3 (TRITC channel) in female WT-TSCs, male WT-TSCs and haiTSCs. Female WT-TSCs (XX) and male (XY) WT-TSCs are used as controls. DNA is stained with DAPI. Scale bar, 50 μm.
Figure S5 related to Figure 3

(A) Brightfield images of haiTSCs grown in differentiation medium (TS medium without F4H) on Day 0, Day 2, Day 4 and Day 6. Scale bar, 50 μm.

(B) The expression levels of trophoblast lineage-specific gene markers (Ctsq, Prl3b1, Prl2c2, Prl3d1, Tpbpa and Ascl2) in 4-day differentiated cells from WT-TSCs, iTSCs (diploid) and haiTSCs. t test, *p < 0.05, **p < 0.01. Data are represented as mean ± SEM.

(C) Teratomas derived from ESCs and hemorrhagic lesions derived from WT-TSCs and haiTSCs.

(D) Reconstructed chimeric blastocysts, with separate contributions of GFP-haESCs (left) and GFP-haiTSCs (right). The red dashed line indicates the area of the inner cell mass. Scale bar, 40 μm.
Figure S6 related to Figure 4

(A) Detected insertion sites of randomly picked subclones by Splinkerette PCR.
(B) Summary of the detected integrated sites.
(C) Brightfield images of differentiated cells from mutated (trapping) and nonmutated haiTSCs (control) in a single repeat. Scale bar, 100 μm.
(D) FACS analysis of Tpbp-positive cells in differentiated cells from the trapping and control groups in the same repeat.
Figure S7 related to Figure 5
(A-I) Strand-specific coverage tracks of the genes *Zfp704*, *Rsf1*, *Tead4*, *Tead1*, *Pbx3*, *Esrrb*, *Nfkb1*, *Smad3* and *Tet1* for the selected library (red).
| Genotyping          | Target Name | Sequence 5’-3’                  |
|---------------------|-------------|---------------------------------|
| AD                  | F: TATGCAAGTACGCCTCTAA  |
| AD                  | R: AGTAATCCAGACGCGTTTT  |
| OE                  | F: GGTGACGTGGAGGAGAATCC  |
| OE                  | R: CTCGTAGAGGAGGAGGTGGC  |
| Htra1               | F: GGCCAATGGGTCTAACCCTTG  |
| Htra1               | R: GACTGGTCGGGCTGAGTTTG   |
| Construct vector    |             |                                 |
| rtTA-neo            | F: CCTACTAGTCACCGGATCCAGATGATAAGATACATT |
| rtTA-neo            | R: TCACACGGTGTCACGTTACCTTACTTTAAGATTACCTG |
| Cdx2                | F: GAGCTCAGGTACCAGCCGGATCCCTG |
| Cdx2                | R: ACGCGCTCTGAGGAGTAAGTGGAGGG  |
| Puro                | F: GTCACCCCAGCGGCTGAGCTGAGCTG |
| Puro                | R: GCTGCCACTGTTTCTTATTGAG  |
| qPCR                |             |                                 |
| Gapdh               | F: AGGTGCGGTTGAAACGGATTTG  |
| Gapdh               | R: TGATGACCATGTAGTTAGGAGGTCA  |
| Oct4                | F: GGGATGGCATACCTGAGACCTCACCTG |
| Oct4                | R: TTTTGTTCCTGGGACTCTCTCG  |
| Nanog               | F: CCAGGGCTATCTGGGTTGAAGC  |
| Nanog               | R: CCCGAAGTTATGGGCGGAGA   |
| Rex1                | F: CCCTCGACAGACTGAACCTCAA  |
| Rex1                | R: TCGGGGCTAATCTCAACCTTCA  |
| Cdx2                | F: GTCCCTAGGAAAGACAGAAGGAG  |
| Cdx2                | R: TTGGCTCTGCAGTCAAGAGA  |
| Eomes               | F: GGAAGTGAGCAGAGGGACTGAGT  |
| Eomes               | R: TTGGCCGAAAGGCTTATGAGC  |
| Elf5                | F: TCTGTCAGCGACCAGTACAG    |
| Elf5                | R: GGAGTACAACCTTGGGAGCAGAA  |
| Tfap2C              | F: ATCCCTACACCTCCCTTCTC    |
| Tfap2c              | R: CCAGATGCGAGTAAATGCTCGG  |
| Ctsq                | F: AGGTATGTAGCTCGTGTTGA    |
| Ctsq                | R: GGCACCAGTCACAGGAAAAAG   |
| Prl3b1              | F: CCAGAAAACAGCGGACTGAGT   |
| Prl3b1              | R: CCAGGCTTGTAAAATGAGTGAG  |
| Prl2c2              | F: GCGGGCATTTTCTTCATAA    |
| Prl2c2              | R: TGAGCCCGAGCAGTTAGAA     |
| Gene   | F            | R            |
|--------|--------------|--------------|
| TpbpA  | GCTATAGTCCCTGAAGCGCA | ACTCCACACTGCTTTTATGAGA |
| Prl3d1 | GCCGCAGATGTGATATAGGGA | AGGGGAAGTGTTCTTGCTGT |
| P53    | CATGAACCGCCGACCTATCC | GCAGTTCAGGGCAAAAGGACT |
| P21    | CCCGAGAACGGTGGAACTTT | AGAGTGCAAGACAGCGACAA |
| Reprimo| TCGCAGTCATGTGTCGCTC | GCCTCCCGTCTCCTCACTAGG |
| Cdk1   | ACACACGAGGTAGTGACGC | TCTGAGTGCCGTGGAAGAG |
| Ascl2  | AAGCACACCTTGACTGTACG | AAGTGAGCGTTTGACACCTCA |
| Gcm1   | GCTCCACAGAGGAAGGCGC | GTTGGTGACCGGGAAGGCC |

### Splinkerette PCR

| Adaptor top | GTCCCATGGTACTACTCATATAATACTACGACTCACTATAGG |
|-------------|-------------------------------------------|
| Adaptor bottom | GCCGTGGTCTGCACCTATAGTGACGTATTATAATTTTTTT |
| Adaptor R1 | GTCCCATGGTACTACTCATA |
| PB5’ F1 | GATATACAGACCAGATAAAACACATGCAG |
| PB3’ F1 | GACGAGTTCGCCTTTAGGAAAGAGAG |
| Adaptor R2 | TAATACGACTCACTATAGG |
| PB5’ F2 | AGCGATGTATATTCTAACGGAAGTC |
| PB3’ F2 | CATGCGTCAATTGACGAGCA |

### P53 knockout sgRNA

| sgRNA1-1 | CACCGAGGAGCTCCTGACACTCGGA |
| sgRNA1-2 | AAATCCGAGGTGACGGAGCTCCTC |
| sgRNA2-1 | CACCGACCTGTACCGAGACCCCT |
| sgRNA2-2 | AAACAGGGGTTCGTTGACAGGGTC |

### Htra1 knockout sgRNA

| sgRNA1-1 | CACCGGCCGCGTCGCGAGCCGTG |
| sgRNA1-2 | AAACACGCGGCTGCGGACGCGGC |


Transparent Methods
Animal use and care
Specific pathogen-free (SPF) grade mice were purchased from Beijing Vital River Laboratory Animal Technology Co. Ltd. All animal procedures were performed under the ethical guidelines of Nankai University Animal Center. Female 129Sv/Jae mice were sacrificed to provide oocytes for the establishment of haESCs and haiTSCs.

Cell Culture
HaESCs were cultured as described previously (Li et al., 2012). TSC medium was made according to a previous report (Tanaka et al., 1998) with slight modification. Briefly, we used a combined TSC medium termed 70CM, containing 30% RPMI 1640 (Thermo) and supplemented with 20% FBS (BI), 2 mM L-glutamine (Sigma), 25 ng/ml human recombinant FGF4 (PeproTech) and 1 μg/ml heparin (Millipore), and 70% MEF conditioned medium with the same supplements. To culture WT-TSCs and stable haiTSCs in chemically defined medium, cells were grown on Matrigel (BD)-coated or feeder-coated dishes in TX medium as described previously (Kubaczka et al., 2014). For differentiation experiments, TS medium lacking FGF4 and heparin was used.

Vector construction
For vector 1, the PiggyBac Dual Promoter was digested by BamHI and KpnI (Thermo), and rtTA and Neo were inserted. For vector 2, the PiggyBac Dual promoter was digested by XhoI and EcoRV, and the TRE, Cdx2 coding sequence and Puro fragments were inserted. P53-targeting sgRNAs were created with the CRISPR design tool (www.crispr.mit.edu). The pSpCas9 (BB)-2A-GFP (PX458) vector was digested and dephosphorylated by BbsI and FastAP (Fermentas). Single-strand oligonucleotides were synthesized and phosphorylated by T4PNK (Takara Japan), and a pair of oligonucleotides was annealed and ligated into lined PX458 to generate knockout plasmids. All plasmids were purchased from Addgene (USA). All primers used are listed in Table S1.

Transfection
To obtain OE-Cdx2 cell lines, approximately 1×10^6 haESCs were electroporated with 6 μg of vector 1 or vector 2 and 2 μg of transposase vector by using an electroporator (Invitrogen) at 1,400 V, 10 μs with three pulses. For vector 1 selection, cells were treated with 250 μg/ml G418 (Thermo) for 7 days. Neomycin-resistant cells were purified for haploid cells. For deletion of P53, approximately 1×10^6 cells were transfected with 4 μg of sgRNA-Cas9 plasmid. Cas9-GFP-positive cells were sorted 36 hours after transfection by flow cytometry. For the gene trapping experiment, a combination of 10 μg of PBase plasmid and 30 μg of piggyBac plasmid (designs see Figure 4A) was electroporated into 1×10^7 haiTSCs using the same conditions as described above (Wang et al., 2018). Puromycin-resistant cells were harvested for Splinkerette PCR.

Generation of induced TSCs from haESCs in vitro
OE-Cdx2 haESCs were seeded at a density of 1×10^5 per well of a six-well plate containing feeder cells and cultured in ESC medium for the first 24 hours. To initiate induction, cells were cultured in TSC media with doxycycline (1 μg/ml) and puromycin (1 μg/ml). Puromycin was withdrawn after 5 days of selection. The medium was refreshed every day until Day 11 to enrich haploid cells, as shown by Hoechst 33342 staining. Haploid cells were sorted according to a previous report (Shuai et al., 2014) a second time after another 12 days. Enriched haploid cells were expanded, and FACS was performed with the CDCP1 antibody as
described previously (Rugg-Gunn et al., 2012).

**Immunostaining, AP staining, and karyotype Analysis**

Before staining, cells were fixed with 4% paraformaldehyde. Permeabilization was performed with 0.3% Triton X-100 (Sigma) for 40 minutes, and then the cells were incubated with 1% BSA (Sigma) for 1 hour at room temperature. Samples were incubated with primary antibodies against Oct4 (Abcam), SSEA1 (Santa Cruz), Cdx2 (Abcam), Nanog (Abcam), Eomes (Abcam) and Tpbpa (Abcam) at 4°C overnight. After three wash steps, the cells were stained with fluorescently coupled secondary antibodies at room temperature for 1 hour. After another three times wash steps, the nuclei were stained with DAPI (YEASEN) for 10 minutes at room temperature. For immunofluorescence: sections were permeabilized with 0.5% Triton X-100 in PBS and then blocked with 5% BSA for 1 h at RT. Slides were then incubated in primary antibody diluted in blocking solution at 4°C overnight. The following primary antibodies were used: Alpha (Ab13970, Abcam), Proliferin (sc-271891, Santa Cruz), GCM1 (sc-101173, Santa Cruz), and Anti-GFP (Ab104401, Abcam). Subsequently, sections were washed with PBST three times and incubated for 1 h at RT in the dark with secondary fluorescence antibodies. The nuclei were stained with Hoechst 33342 for 10 min at RT. AP staining of OE-Cdx2 haESCs was performed as described previously (Zhao et al., 2009). AP staining was performed according to the manufacturer’s instructions for the alkaline phosphatase kit (Beyotime). The results were observed under an inverted microscope or laser-scanning confocal microscope. Karyotype analysis was performed following standard instructions.

**Quantitative Real-time PCR**

Total RNA was purified from cells using Trizol reagent (Thermo), while cDNA was obtained using a Prime Script™ RT Reagent Kit with gDNA Eraser (Takara). Quantitative PCRs were performed with an ABI QuantStudio™ 6 Flex machine using FS Universal SYBR Green Master (Roche). Relative expression levels were normalized to Gapdh. Averages and SD values were from three independent experiments. All the primers used are listed in Table S1.

**Western blotting**

For western blotting experiments, protein samples were extracted by RIPA lysis solution (Solarbio) from ESCs and TSCs. Lysates were precleared by centrifugation for 5 minutes. Equal amounts of cell lysates were separated by SDS-PAGE and transferred to a polyvinylidene fluoride (PVDF) membrane (GE). The membranes were blocked in 5% nonfat dry milk for 1 hour, washed three times by TBST, 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. Signals on the membrane were detected by using Western Blotting Detection Reagents (Engreen) and imaged. The primary antibodies used for western blotting included anti-Cdx2 (Abcam), anti-Eomes (Abcam) and anti-Gapdh (Abcam).

**Bisulphite Sequencing**

Genomic DNA of each sample was modified with the EpiTect Bisulfite Kit (Qiagen). Two-round nested PCR was performed to amplify the promoter region of each gene and PCR products were purified from agarose gel by Gel Extraction Kit (OMEGA). Primer sequences were as previous described for Oct4 (Blelloch et al., 2006), Nanog(Takahashi and Yamanaka, 2006) and Elf5(Ng et al., 2008).

**Splinkerette PCR**

Splinkerette PCR was used to find the insertion sites after piggyBac transposon transfection as previously described (Uren et al., 2009). Briefly, the genomic DNA purified from cells was digested by BspI43I
and an adenine was added to the tails of fragments by Taq DNA Polymerase (Thermo). Purified products were linked to Splinkerette adaptors for nest PCR. Nest PCR products were linked into pEASY Blunt simple vectors (Transgene) for Sanger sequencing. All primers and Splinkerette adaptors used are listed in Table S1.

**T7EN1 assay**

Following the manufacturer’s recommended protocol (NEB), target DNA fragments containing sgRNA target sites were amplified from transfected cells and wild-type cell genomic DNA by Q5® Hot Start High-Fidelity DNA Polymerase (NEB). The two kinds of fragments were mixed together and annealed in Buffer 2.0 (NEB). T7 Endonuclease I was added into the mixture, and digestion was carried out for 1 hour at 37°C. The products were run on a 2% TAE gel (Biowest), stained with SYBR® Safe DNA gel stain (Thermo) and imaged on a gel imager (Bio-Rad).

**TSC Transplantation and immunohistochemistry staining**

TSC transplantations were performed as published (Kubaczka et al., 2014). Briefly, 1×10⁶ haiTSCs were resuspended in 200 μl of 70CM containing FGF4/heparin and injected into the testis of 8-week-old male ICR mice. Approximately 21 days postinjection, hemorrhagic lesions were dissected from the testis, fixed in 4% paraformaldehyde overnight, embedded in paraffin, and sectioned. Sections were rehydrated and treated for antigen retrieval in sodium citrate buffer at a sub-boiling temperature. Next, endogenous peroxidase activity was quenched with 3% hydrogen peroxidase at room temperature (RT). For immunohistochemistry: sections were blocked with 5% BSA for 1 hour at RT. The following primary antibodies were used: Tfap2c (sc-12762; Santa Cruz) and CD31 (MA5-13188, Thermo Fisher). For immunofluorescence: sections were permeabilized with 0.5% Triton X-100 in PBS and then blocked with 5% BSA for 1 h at RT. All primary antibodies against Alpha (Ab13970, Abcam), Proliferin (sc-271891, Santa Cruz), GCM1 (sc-101173, Santa Cruz), and Anti-GFP (Ab104401, Abcam) were incubated with samples overnight at 4°C.

Before staining, cells were fixed with 4% paraformaldehyde. Permeabilization was performed with 0.3% Triton X-100 (Sigma) for 40 minutes and then incubated with 1% BSA (Sigma) for 1 hour at RT. Samples were incubated with primary antibodies against Oct4 (Abcam), SSEA1 (Santa Cruz), Cdx2 (Abcam), Nanog (Abcam), Eomes (Abcam) and Tpbpa (Abcam) at 4°C overnight. After three times washing steps, the samples were stained with fluorescently coupled secondary antibodies at room temperature for 1 hour.

**Diploid chimeric assay**

Chimeric embryos were generated by injecting haiTSC-GFP cells into CD-1 mouse recipient 4-cell embryos according to a previous protocol (Li et al., 2012), using ES-GFP cells as a control treatment. Diploid 4-cell embryos were collected from the uterus of a 2.5 dpc female CD-1 mouse. Approximately 10-15 ES-GFP cells and FACS-purified G0/G1 phase haiTS-GFP cells were injected into each blastocyst. Then, chimeric embryos were cultured to the blastocyst stage in KSOM (Millipore) for subsequent analysis.

**Genetic screening in haiTSCs**

HaiTSCs were mutated with trapping vectors (Figure 4A) and treated with puromycin (1 μg/ml) for 1 day. Puromycin-resistant haiTSCs and nontransfected haiTSCs were differentiated in standard TS medium without heparin and FGF4 for 3 days. Cells positive for Tpbpa (Abcam, ab104401) among the differentiated cells were harvested by antibody staining and sorting according to the manufacturer’s
instructions. In each repeat, we selected Tpbpa-positive cells in the significantly increasing group relative to the nontransfected group to perform high-throughput sequencing.

**Analysis of RNA-seq data**

Raw data were trimmed using customized scripts to remove low-quality bases (quality score < 5) or reads (low-quality bases > 50% of the read size). Then, the abundance of each transcript was counted using Kallisto (Bray et al., 2016) with Gencode M18 (Harrow et al., 2006) and further summarized for each gene using the R package tximport (Soneson et al., 2015). Genes covered by less than 2 reads across all samples or with more than 2 samples possessing zero reads were filtered out. Genes were normalized using relative-log-expression (RLE) from DESeq2 (Love et al., 2014). Correlations among samples were measured based on the top 50 most variable genes using cosine distance. To generate the MDS plot of all samples, data were first transformed using the function variance Stabilizing() and then the function plotPCA() from DESeq2 (Love et al., 2014). Plots were visualized by ggplot2 (Ginestet, 2011).

**Analysis of piggyBac screening data**

Reads were mapped to the genome assembly from UCSC (mm10) (Zimin et al., 2014) using bowtie2 (v2.2.3) (Langmead and Salzberg, 2012) under sensitive-local mode. Read distributions across different regions were summarized using RSeQC (v2.6.1) (Wang et al., 2012). VISITs (v0.22) (Yu and Ciaudo, 2017) was used to estimate the number of insertions for each gene using the annotation from Gencode (Harrow et al., 2006), excluding duplicated and multiple-hit reads.

Data were first filtered by removing genes covered by < 2 reads, then normalized using relative-log-expression (RLE) from DESeq2 (v1.10.1) (Love et al., 2014). Log2-transformed fold changes were further calculated based on normalized data. The coverage tracks were generated using R package GenomeGraphs (v1.30) (Bullard, 2016). For enrichment analysis, gene sets were retrieved from the Gene Ontology (Ashburner et al., 2000) and KEGG databases (Kanehisa and Goto, 2000). The top 200 genes with the most abundant insertions were used. Fisher’s exact test followed by the Benjamini-Hochberg correction (Benjamini and Hochberg, 1995) was used to generate the FDR.

**Supplemental Reference**

ASHBURNER, M., BALL, C. A., BLAKE, J. A., BOTSTEIN, D., BUTLER, H., CHERRY, J. M., DAVIS, A. P., DOLINSKI, K., DWIGHT, S. S., EPPIG, J. T., HARRIS, M. A., HILL, D. P., ISSEL‐TARVER, L., KASARSKIS, A., LEWIS, S., MATESE, J. C., RICHARDSON, J. E., RINGWALD, M., RUBIN, G. M. & SHERLOCK, G. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*, 25, 25-9.

BENJAMINI, Y. & HOCHBERG, Y. 1995. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Methodological*, 57, 289-300.

BLELLOCH, R., WANG, Z., MEISSNER, A., POLLARD, S., SMITH, A. & JAENISCH, R. 2006. Reprogramming efficiency following somatic cell nuclear transfer is influenced by the differentiation and methylation state of the donor nucleus. *Stem Cells*, 24, 2007-13.

BRAY, N. L., PIMENTEL, H., MELSTED, P. & PACHTER, L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol*, 34, 525-7.

BULLARD, S. D. A. J. 2016. GenomeGraphs: Plotting genomic information from Ensembl. R package version 1.32.0. *Bioconductor*. 
GINESTET, C. 2011. ggplot2: elegant graphics for data analysis. Journal of the Royal Statistical Society: Series A (Statistics in Society).

HARROW, J., DENOUD, F., FRANKISH, A., REYMOND, A., CHEN, C. K., CHRAST, J., LAGARDE, J., GILBERT, J. G., STOREY, R., SWARRECK, D., ROSSIER, C., UCLA, C., HUBBARD, T., ANTONARAKIS, S. E. & GUIGO, R. 2006. GENCODE: producing a reference annotation for ENCODE. Genome Biol, 7 Suppl 1, S4 1-9.

KANEHISA, M. & GOTO, S. 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res, 28, 27-30.

KUBACZKA, C., SENNER, C., ARAUZO-BRAVO, M. J., SHARMA, N., KUCKENBERG, P., BECKER, A., ZIMMER, A., BRUSTLE, O., PEITZ, M., HEMBERGER, M. & SCHORLE, H. 2014. Derivation and maintenance of murine trophoblast stem cells under defined conditions. Stem Cell Reports, 2, 232-42.

LANGMEAD, B. & SALZBERG, S. L. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods, 9, 357-9.

LI, W., SHUAI, L., WAN, H., DONG, M., WANG, M., SANG, L., FENG, C., LUO, G. Z., LI, T., LI, X., WANG, L., ZHENG, Q. Y., SHENG, C., WU, H. J., LIU, Z., LIU, L., WANG, L., WANG, X. J., ZHAO, X. Y. & ZHOU, Q. 2012. Androgenetic haploid embryonic stem cells produce live transgenic mice. Nature, 490, 407-11.

LOVE, M. I., HUBER, W. & ANDERS, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol, 15, 550.

NG, R. K., DEAN, W., DAWSON, C., LUCIFERO, D., MADEJA, Z., REIK, W. & HEMBERGER, M. 2008. Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5. Nat Cell Biol, 10, 1280-90.

RUGG-GUNN, P. J., COX, B. J., LANNER, F., SHARMA, P., IGNATCHENKO, V., MCDONALD, A. C., GARNER, J., GRAMOLINI, A. O., ROSSANT, J. & KISLINGER, T. 2012. Cell-surface proteomics identifies lineage-specific markers of embryo-derived stem cells. Dev Cell, 22, 887-901.

SHUAI, L., LI, W., WAN, H., ZHAO, X. Y., WANG, L. & ZHOU, Q. 2014. Generation of Mammalian offspring by haploid embryonic stem cells microinjection. Curr Protoc Stem Cell Biol, 31, 1A 6 1-15.

SONESON, C., LOVE, M. I. & ROBINSON, M. D. 2015. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res, 4, 1521.

TAKAHASHI, K. & YAMANAKA, S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell, 126, 663-76.

TANAKA, S., KUNATH, T., HADJANTONAKIS, A. K., NAGY, A. & ROSSANT, J. 1998. Promotion of trophoblast stem cell proliferation by FGF4. Science, 282, 2072-5.

UREN, A. G., MIKKERS, H., KOOI, J., VAN DER WEYDEN, L., LUND, A. H., WILSON, C. H., RANCE, R., JONKERS, J., VAN LOHUIZEN, M., BERNS, A. & ADAMS, D. J. 2009. A high-throughput splinkerette-PCR method for the isolation and sequencing of retroviral insertion sites. Nat Protoc, 4, 789-98.

WANG, H., ZHANG, W., YU, J., WU, C., GAO, Q., LI, X., LI, Y., ZHANG, J., TIAN, Y., TAN, T., JI, W., LI, L., YU, Y. & SHUAI, L. 2018. Genetic screening and multipotency in rhesus monkey haploid neural progenitor cells. Development, 145.

WANG, L., WANG, S. & LI, W. 2012. RSeQC: quality control of RNA-seq experiments. Bioinformatics, 28, 2184-5.

YU, J. & CIAUDO, C. 2017. Vector Integration Sites Identification for Gene-Trap Screening in Mammalian Haploid Cells. Scientific reports, 7, 44736.
ZHAO, X. Y., LI, W., LV, Z., LIU, L., TONG, M., HAI, T., HAO, J., GUO, C. L., MA, Q. W., WANG, L., ZENG, F. Y. & ZHOU, Q. 2009. iPS cells produce viable mice through tetraploid complementation. *Nature*, 461, 86-U88.

ZIMIN, A. V., CORNISH, A. S., MAUDHOO, M. D., GIBBS, R. M., ZHANG, X., PANDEY, S., MEEHAN, D. T., WIPFLER, K., BOSINGER, S. E., JOHNSON, Z. P., THARP, G. K., MARCAIS, G., ROBERTS, M., FERGUSON, B., FOX, H. S., TREANGEN, T., SALZBERG, S. L., YORKE, J. A. & NORGNEN, R. B., JR. 2014. A new rhesus macaque assembly and annotation for next-generation sequencing analyses. *Biol Direct*, 9, 20.