STEM CELLS

Haploid-genetic screening of trophectoderm specification identifies *Dyrk1a* as a repressor of totipotent-like status

Wenhao Zhang[1]†, Shengyi Sun[1]†, Qing Wang[1]†, Xu Li[1,2]†, Mei Xu[1]†, Qian Li[3], Yiding Zhao[1], Keli Peng[1], Chunmeng Yao[1], Yuna Wang[1], Ying Chang[1], Yan Liu[4], Xudong Wu[3]*, Qian Gao[1]*, Ling Shuai[1,5]*

Trophectoderm (TE) and the inner cell mass are the first two lineages in murine embryogenesis and cannot naturally transit to each other. The barriers between them are unclear and fascinating. Embryonic stem cells (ESCs) and trophoblast stem cells (TSCs) retain the identities of inner cell mass and TE, respectively, and, thus, are ideal platforms to investigate these lineages in vitro. Here, we develop a loss-of-function genetic screening in haploid ESCs and reveal many mutations involved in the conversion of TSCs. The disruption of either *Catip* or *Dyrk1a* (candidates) in ESCs facilitates the conversion of TSCs. According to transcriptome analysis, we find that the repression of *Dyrk1a* activates totipotency, which is a possible reason for TE specification. *Dyrk1a*-null ESCs can contribute to embryonic and extraembryonic tissues in chimeras and can efficiently form blastocyst-like structures, indicating their totipotent developmental abilities. These findings provide insights into the mechanisms underlying cell fate alternation in embryogenesis.

INTRODUCTION

The first cell fate decision in mice occurs at the time between the 8-cell embryo stage and the 16-cell embryo stage, which leads to the first two different cell lineages. In this period, identical blastomeres segregate to the inner cell mass (ICM) (the origin of the epiblast, which further develops into an offspring) and the primitive endoderm and trophectoderm (TE) (the origin of the placenta) (*1*). There are definite boundaries between the ICM and TE; thus, they are not able to naturally transition to the other in murine species. The exact mechanisms underlying cell fate segregation remain unknown and raise extensive concerns. ICM-derived embryonic stem cells (ESCs) and TE-derived trophoblast stem cells (TSCs) are excellent platforms to study the relationships and differences between ICM and TE in vitro *(2)*. Previous evidence showed that the suppression of *Oct4* (*3*) or overexpression of *Cdx2* (*4*) allowed ESCs to differentiate into the TE lineage. Likewise, the overexpression of *Oct4* facilitated TSC reprogramming to ESCs to regain pluripotency (*5*). Another transcription factor (TF), *Tead4*, is also very important for TE specification in preimplantation development (*6*), mainly through regulation of the Hippo-Yap pathway in this process (*7*, *8*). Moreover, Notch is involved in TE specification by targeting many TE-specific genes (*9*). A combination of several TFs, including *Gata3*, *Eomes*, *Tfap2c*, and *Ets2*, also promoted TE specification from fibroblasts (*10*, *11*) and ESCs (*12*). In addition, microRNAs were found to be sufficient to drive the conversion of the extraembryonic lineage from ESCs (*13*). However, whether there are other genes and pathways regulating TE specification is still unclear.

Recently established mouse haploid ESCs (haESCs) are premium tools for finding functionally unknown genes due to their hemizygous and pluripotent nature (*14*). Many attempts to find target genes of important biological processes have been realized through high-throughput genetic screening in haESCs combined with genome-wide random mutations (*15*, *16*). *PiggyBac* (PB) is an efficient transposon that randomly induces abundant mutations in the mammalian genome and is widely used in generating mutation libraries. Combined with the PB system, haESCs can easily produce genome-wide homozygous mutations (*17*, *18*), which is useful for the selection for key genes regulating any desired approach. Whether such a powerful genetic screening system is suitable for revealing key restrictions of TE specification needs further investigation.

Here, we describe the derivation of genome-scale mutant libraries in haESCs assisted with the PB system. The mutant ESCs can develop TSC properties after a period of culture in TSC medium, which are sequenced and analyzed to identify the insertions. Thereafter, we performed validation experiments of candidate genes to prove their essential roles in TE specification and discuss the potential mechanism underlying this transition.

[1]State Key Laboratory of Medicinal Chemical Biology and College of Pharmacy and Tianjin Central Hospital of Gynecology Obstetrics/Tianjin Key Laboratory of Human Development and Reproductive Regulation, Nankai University, Tianjin 300350, China. [2]Department of Molecular Biology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA. [3]State Key Laboratory of Experimental Hematology, The Province and Ministry Co-sponsored Collaborative Innovation Center for Medical Epigenetics, Key Laboratory of Immune Microenvironment and Disease (Ministry of Education), Tianjin Key Laboratory of Medical Epigenetics, Department of Cell Biology, Tianjin Medical University, Tianjin 300070, China. [4]Department of Obstetrics, Tianjin First Central Hospital, Nankai University, Tianjin 300192, China. [5]National Clinical Research Center for Obstetrics and Gynecology, Peking University Third Hospital, Beijing 100191, China.
*Corresponding author. Email: wuxudong@tmu.edu.cn (X.W.); gaoqian@nankai.edu.cn (Q.G.); lshuai@nankai.edu.cn (L.S.)
†These authors contributed equally to this work.

RESULTS

Genetic screening related to TE conversion in mutant haESCs

To construct mutant cell libraries, we used a mouse haESC line with a 129Sv/Jae background established in our previous work (*19*). The haESCs grew well and had a high proportion of haploid cells before mutation (fig. S1A). PB transposons had a high stability and were easy to modify, and they were also easily located by inverse polymerase chain reaction (PCR) (*20*). Therefore, we designed a PB-splice acceptor (SA)–based plasmid carrying a puromycin resistance gene (*Puro*^{r}) to introduce mutations into the genome (fig. S1B). For genetic screening of TE specification, we chose a well-developed cell culture system to guide this process (Fig. 1A). Briefly, a combination of 24 μg of PB and 8 μg of PBase vectors was electroporated into approximately 1×10^7 haESCs. When cells recovered from electroporation for 2 days, puromycin was used to select for mutated ESCs with PB insertions. Mutant cells with puromycin resistance [incubated with puromycin (1 μg/ml) for 2 days] survived, whereas almost all cells in the control group died (fig. S1C). When the mutant ESCs were cultured in TSC medium [70% conditioned medium (CM) with fibroblast growth factor 4 (FGF4) and heparin (F4H)] for 10 days, the cell cultures gradually showed TSC-like colonies [termed induced TSCs (iTSCs)]. For enrichment of these iTSCs in the cell cultures, we chose a TSC-specific flow cytometric antibody, CDCP1 (*21*), to purify them. Approximately 36.6% of the cells were CDCP1-positive and were harvested for further culture (Fig. 1B). The morphology of the CDCP1-enriched iTSCs [termed screened out iTSCs (S-iTSCs)] was similar to that of the wild-type (WT) TSCs (Fig. 1C). In contrast, WT-ESCs were also induced using the same strategy (Fig. 1A). The immunofluorescence results showed that only the S-iTSCs presented CDX2 and EOMES positivity, whereas the cell cultures of WT-ESCs failed to achieve iTSCs in the same culture system (Fig. 1D). After expansion, the S-iTSCs showed high positive proportions of two TE-specific markers (CDCP1, 92.0%; CD40, 87.3%) during culture, as shown by flow cytometric analysis (Fig. 1, E and F), and expressed TSC-specific markers (CDX2 and EOMES), as shown by Western blotting (WB) (fig. S1D). In addition, the S-iTSCs could further differentiate into terminal trophoblast lineages, including trophoblast giant cells (PROLIFERIN-positive) and spongiotrophoblast cells (TPBPA-positive), as indicated by immunostaining (fig. S1E).

Analysis of insertions in the S-iTSCs

Thereafter, we expanded the S-iTSCs and prepared them for next-generation sequencing, with original mutated ESCs without conversion as controls. Next, we performed splinkerette PCR (*22*) with the total genomic DNA of the S-iTSCs and controls, the PCR product of which was sent to a local company for deep sequencing. According to the analysis, we identified nearly 129 thousand independent hits across more than 10,000 genes (fig. S2A). Among these hits, approximately 50.38% of the insertions were mapped to gene bodies (promoter, 15.13%; untranslated region, 1.04%; exon, 2.09%; intron, 32.12%), whereas the rest of the insertions were located in intergenic areas (Fig. 2A). The Kyoto Encyclopedia of Genes and Genomes

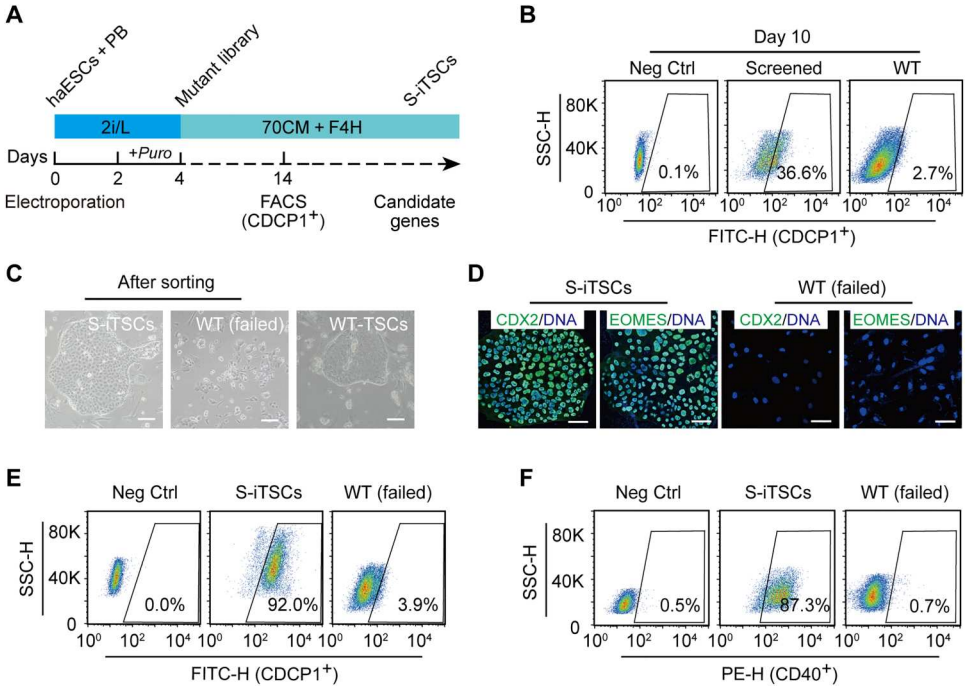

Fig. 1. A genetic screening of TE specification with mutant haESCs. (**A**) Schematic overview of forward genetic screening using mutant haESCs to conduct TE specification. (**B**) Proportion of CDCP1-positive cells in the cell cultures of mutant haESCs cultured in TSC medium for 10 days. WT-haESCs were used as a negative control. SSC-H, side scatter-height. (**C**) Phase-contrast images of S-iTSCs, WT-ESCs in the same conversion culture system (failed), and WT-TSCs. Scale bars, 100 μm. (**D**) Immunofluorescence of TSC-specific markers CDX2 (green) and EOMES (green) in the S-iTSC– and WT-ESC–derived cell cultures (failed). DNA was stained with Hoechst 33342 (blue). Scale bars, 50 μm. (**E**) Fluorescence-activated cell sorting (FACS) analysis of TE-specific marker–positive (CDCP1$^+$) cells in the S-iTSCs (92.0%) during culture. WT-ESCs and WT-ESC–derived cell cultures (failed) were used as controls. (**F**) FACS analysis of another TE-specific marker–positive (CD40$^+$) cells in the S-iTSCs (87.3%) during culture. WT-ESCs and WT-ESC–derived cell cultures (failed) were used as controls. PE, phycoerythrin.

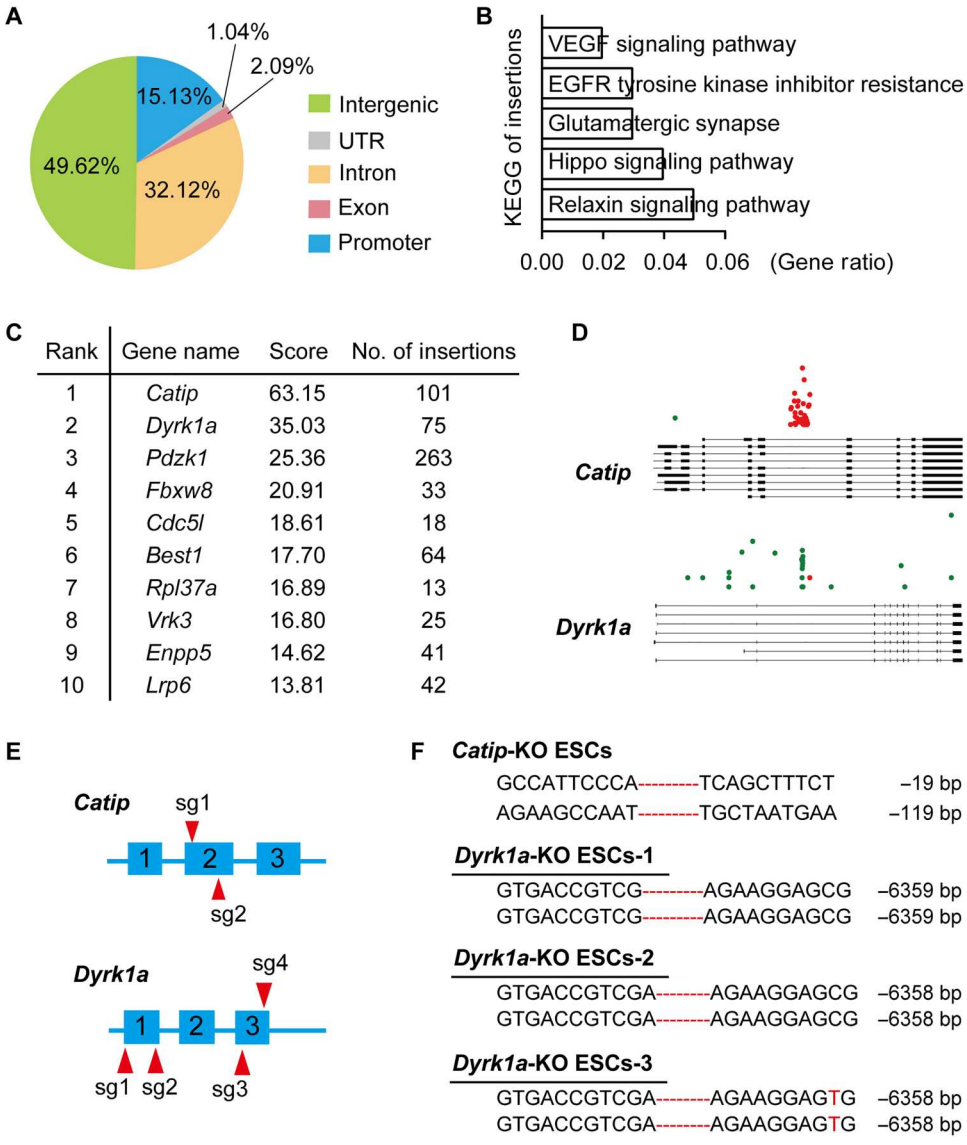

Fig. 2. Bioinformatic analysis of insertions in the S-iTSCs. (**A**) Proportion of PB integration sites across the genome: promoters (15.13%), untranslated region (UTR) (1.04%), exons (2.09%), introns (32.12%), and intergenic regions (49.64%). (**B**) Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of the top 200 genes with the most frequent insertions from the S-iTSCs. VEGF, vascular endothelial growth factor; EGFR, epidermal growth factor receptor. (**C**) The top 10 candidate genes among the inserted genes from the S-iTSCs. (**D**) Sense (red) and antisense (green) insertions of PB in *Catip* and *Dyrk1a* gene bodies from the S-iTSCs. The rectangles indicate the exons. (**E**) Schematic diagram of the strategy to *Catip*-KO and *Dyrk1a*-KO in WT-ESCs by using CRISPR-Cas9 system. (**F**) Sequences of genotypes in the *Catip*-KO and *Dyrk1a*-KO ESCs.

analysis showed that most insertions were related to the Relaxin signaling pathway, Hippo signaling pathway, and other biological activities (Fig. 2B). We further analyzed the top 2000 inserted genes [listed according to the score evaluation by HaSAPPy methodology (*23*)] and found that many insertions in some critical pathways (Hippo/Yap, LIF/STATA3, and others) were related to the first cell fate determination (fig. S2B). The top 10 genes were ranked according to the score and insertion frequency (Fig. 2C), from which the top 2 candidates (*Catip* and *Dyrk1a*) showed enrichment in many hit loci (Fig. 2D). Overall, our genome-wide screening of TE specification in mouse haESCs revealed numerous key genes and pathways involved in the ESC-to-TSC transition. For a proof-of-concept study, we chose the top 2 candidate genes (*Catip* and *Dyrk1a*) to perform the validation experiments. We investigated whether the absence of *Catip* or *Dyrk1a* promoted the conversion from ESCs to TSCs. We designed Cas9–green fluorescent protein (GFP) knockout (KO) vectors to induce *Catip*-KO or *Dyrk1a*-KO and transfected WT-ESCs with these vectors by electroporation (Fig. 2E). Next, the selected KO subclones were genotyped and confirmed by sequencing (Fig. 2F). We chose one *Catip*-KO ESC line and three *Dyrk1a*-KO ESC lines for further investigations.

Pluripotency and genome integrity of *Catip*-KO and *Dyrk1a*-KO ESCs

To examine whether KO of *Catip* or *Dyrk1a* affects genome integrity and pluripotency, we performed subsequent experiments. We assessed the cell proliferation and karyotypes of the two KO ESC lines and found no notable change in proliferation, and there was mere karyotype abnormality between the KO ESCs (*Catip*-KO and *Dyrk1a*-KO) and WT-ESCs (Fig. 3, A to C). To address the differentiation potentials of the KO ESCs, we injected approximately 1 $\times 10^7$ KO ESCs into each group separately under the skin of severe combined immunodeficiency mice, with WT-ESCs as a control. Both the *Catip*-KO ESCs and the *Dyrk1a*-KO ESCs could form standard teratomas 3 weeks later (fig. S3A), with three germ layer representative tissues according to the histological analysis (Fig. 3D). Furthermore, the global transcriptome analysis showed that both the *Catip*-KO ESCs (R = 0.96) and the *Dyrk1a*-KO ESCs (R = 0.97) had very high correlations with WT-ESCs (fig. S3B). In addition, the expression levels of pluripotent genes (*Oct4*, *Nanog*, *Klf4*, and *Rex1*) and TE-specific genes (*Cdx2*, *Eomes*, and *Tead4*) in the KO ESCs were not changed compared with those in the WT-ESCs according to RNA sequencing (RNA-seq) analysis (fig. S3C). The differentially expressed genes (DEGs) between *Catip*-KO ESCs and WT-ESCs were enriched in protein binding and other functions (Fig. 3E), being similar to the DEGs between *Dyrk1a*-KO ESCs and WT-ESCs (Fig. 3F), as shown by Gene Ontology (GO) analysis. Thus, the KO of *Catip* or *Dyrk1a* did not affect

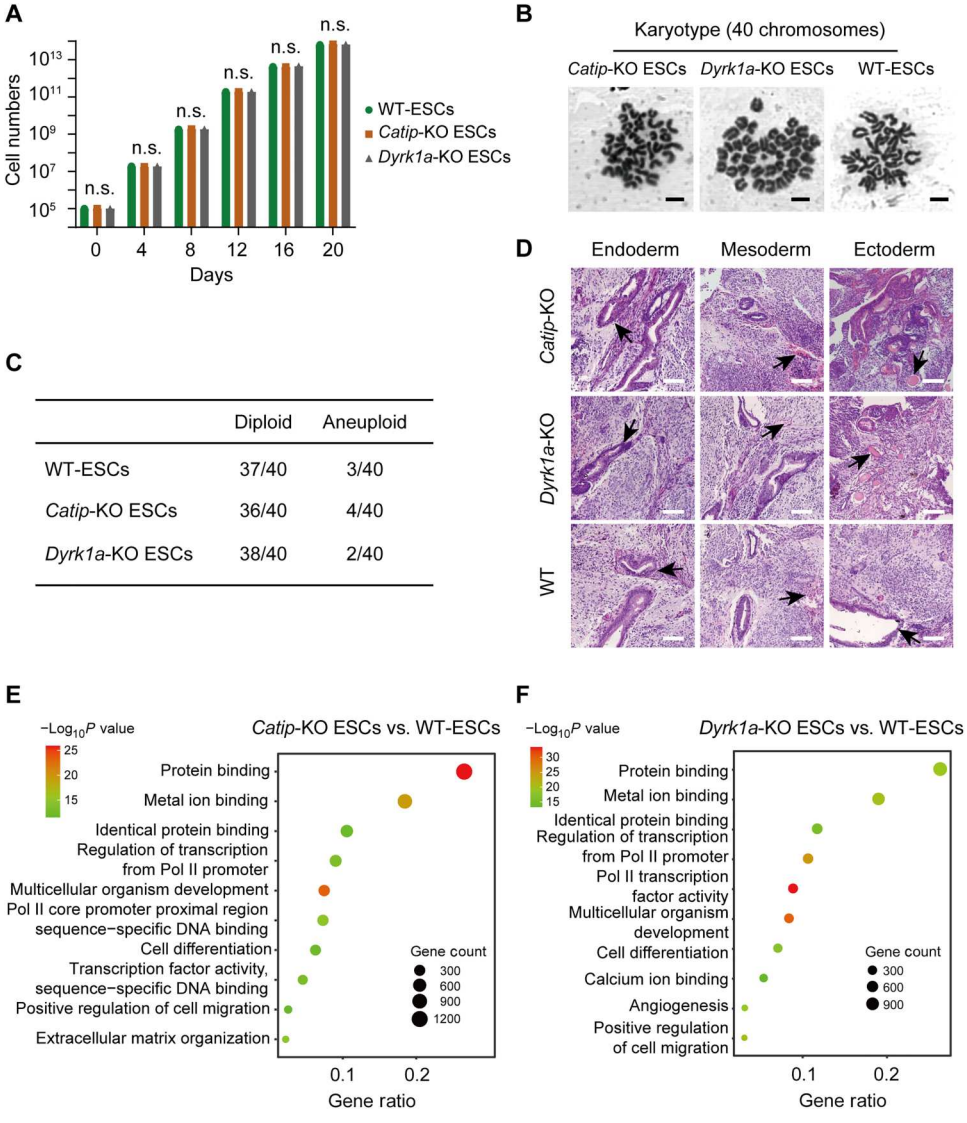

Fig. 3. Pluripotency and genome integrity analysis of *Catip*-KO ESCs and *Dyrk1a*-KO ESCs. (**A**) The live cell numbers of the *Catip*-KO ESCs, *Dyrk1a*-KO ESCs, and WT-ESCs during daily culture. *Catip*-KO and *Dyrk1a*-KO did not affect the proliferation of ESCs. n.s., not significant. (**B**) Chromosome spread analysis of *Catip*-KO ESCs, *Dyrk1a*-KO ESCs, and WT-ESCs (diploid). Scale bars, 10 μm. (**C**) Statistical analysis of chromosome numbers in the *Catip*-KO ESCs, *Dyrk1a*-KO ESCs, and WT-ESCs. (**D**) Teratomas formed from the *Catip*-KO ESCs, *Dyrk1a*-KO ESCs, and WT-ESCs were identified by hematoxylin and eosin staining. Scale bars, 100 μm. The tissues shown were glands (endoderm), blood (mesoderm), and neural tissue (ectoderm). (**E**) GO analysis of DEGs between *Catip*-KO ESCs and WT-ESCs. (**F**) GO analysis of DEGs between *Dyrk1a*-KO ESCs and WT-ESCs. Pol II, polymerase II.

the main properties of ESCs. Further on, we investigated whether KO of the two genes would promote the ESC-to-TSC transition.

The ESC-to-TSC transition with *Catip*-KO and *Dyrk1a*-KO

Next, we cultured *Catip*-KO ESCs and *Dyrk1a*-KO ESCs in TSC medium to induce TE specification. After 10 days, many TSC-like colonies appeared in the cell cultures of the *Catip*-KO and *Dyrk1a*-KO groups. Then, we performed fluorescence-activated cell sorting (FACS) to enrich the iTSCs from the cell cultures on day 10 with a CDCP1–flow cytometric antibody. These CDCP1-positive cells were further cultured in TSC medium to expand and showed the typical morphology of TSCs (Fig. 4A). The *Catip*-KO and *Dyrk1a*-KO iTSCs maintained a high proportion of CDCP1-positive (92.6 and 96.3%, respectively) and CD40-positive (82.1 and 84.6%, respectively) cells in long-term culture (fig. S4A). The immunofluorescence results showed that these iTSCs were CDX2- and EOMES-positive (Fig. 4B). For the random differentiation of the *Catip*-KO and *Dyrk1a*-KO iTSCs, we cultured them in TSC medium without F4H. DNA content analysis of the differentiated cells on days 0, 2, 4, and 6 showed that 8n polyploidy peaked with

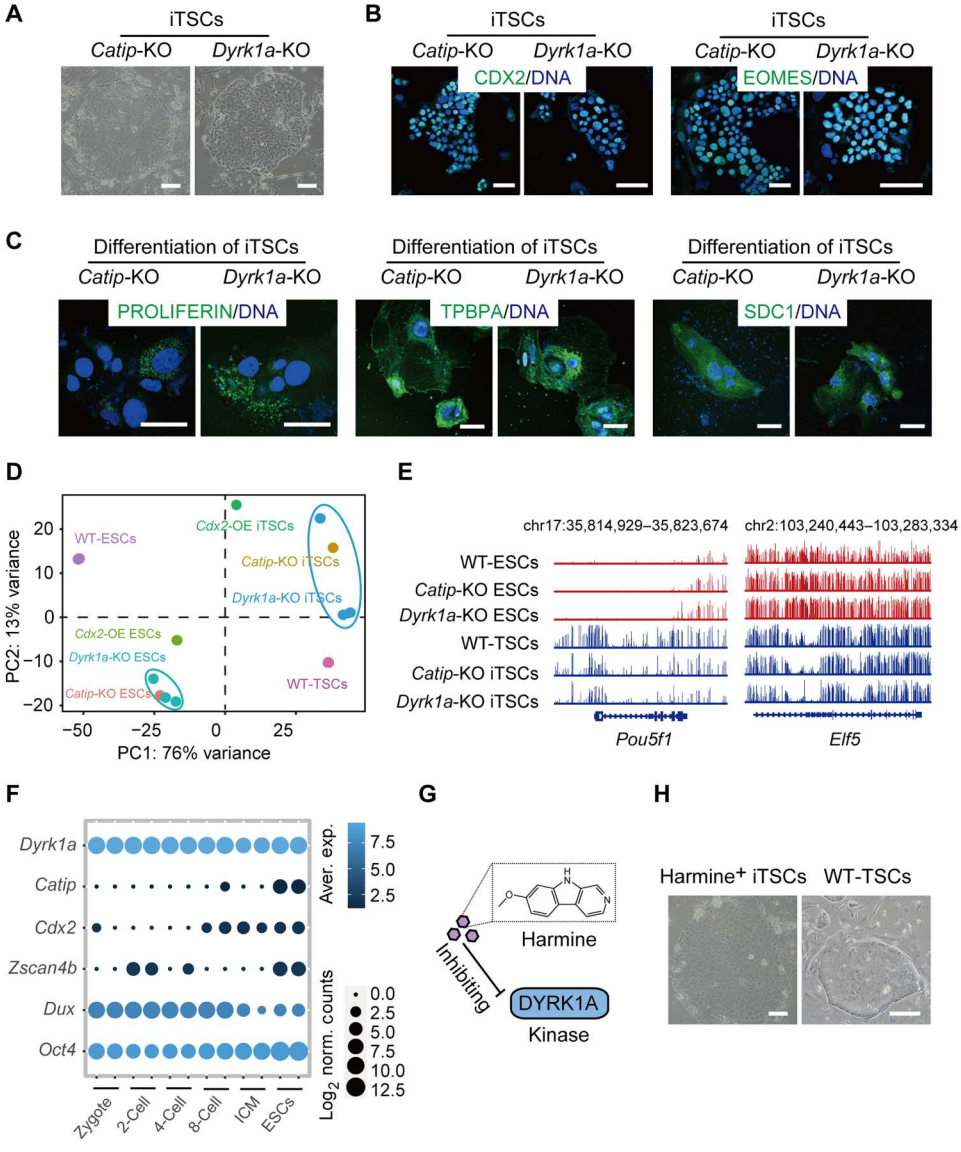

Fig. 4. Derivation of *Catip*-KO and *Dyrk1a*-KO iTSCs. (**A**) Phase-contrast images of established *Catip*-KO iTSCs and *Dyrk1a*-KO iTSCs. Scale bars, 100 μm. (**B**) Immunofluorescence of TSC-specific markers CDX2 (green) and EOMES (green) in the *Catip*-KO iTSCs and *Dyrk1a*-KO iTSCs. DNA was stained with Hoechst 33342 (blue). Scale bars, 50 μm. (**C**) Immunofluorescence of the trophoblast lineage–specific markers PROLIFERIN (green), TPBPA (green), and SDC1 (green) in differentiated cells derived from the *Catip*-KO iTSCs and *Dyrk1a*-KO iTSCs. DNA was stained with Hoechst 33342 (blue). Scale bars, 50 μm. (**D**) PCA of RNA-seq datasets from the WT-ESCs, *Catip*-KO ESCs, *Dyrk1a*-KO ESCs, *Cdx2*-OE ESCs, WT-TSCs, *Catip*-KO iTSCs, *Dyrk1a*-KO iTSCs, and *Cdx2*-OE iTSCs. (**E**) Genome browser tracks showing CpG methylation status in *Pou5f1* and *Elf5* loci of WT-ESCs, *Catip*-KO ESCs, *Dyrk1a*-KO ESCs, WT-TSCs, *Catip*-KO iTSCs, and *Dyrk1a*-KO iTSCs. (**F**) Bubble plot representing the expression of *Catip*, *Dyrk1a*, *Cdx2*, *Zscan4b*, *Dux*, and *Oct4* in early embryonic developing stages and ESCs. (**G**) Schematic diagram of the DYRK1A inhibitor–harmine. (**H**) Phase-contrast images of established harmine^{+} iTSCs and WT-TSCs. Scale bars, 100 μm.

the increase in days, indicating the emergence of differentiating giant cells (fig. S4B). The immunofluorescence staining positive for PROLIFERIN, TPBPA, and SDC1 (syntrophoblast-specific marker) in differentiated cells also confirmed their differentiation potentials (Fig. 4C). We further analyzed the RNA-seq data of the *Catip*-KO ESCs, *Dyrk1a*-KO ESCs, *Cdx2*-overexpression (OE) ESCs, WT-ESCs, *Catip*-KO iTSCs, *Dyrk1a*-KO iTSCs, *Cdx2*-OE iTSCs, and WT-TSCs to investigate their properties. Principal components analysis (PCA) revealed that the *Catip*-KO iTSCs and *Dyrk1a*-KO iTSCs were closer to the WT-TSCs and *Cdx2*-OE iTSCs than to the other cells (Fig. 4D). The whole-genome bisulfite sequencing (WGBS) analysis showed that *Catip*-KO iTSCs and *Dyrk1a*-KO iTSCs had similar methylation status in the loci of *Pou5f1* and *Elf5* to WT-TSCs, being different from WT-ESCs (Fig. 4E). In total, our findings validated that either *Catip*-KO or *Dyrk1a*-KO ensured that ESCs were converted to TSCs, along with differentiation potentials into terminal trophoblast lineages. To further investigate the roles of *Catip* and *Dyrk1a* involved in the first cell fate decision, we checked the expression levels of them in early embryo developing stages (from zygote to blastocyst). According to the RNA-seq data (download from the public database, PMID: 27309802) (*24*), *Dyrk1a* was highly expressed in all stages before implantation, whereas *Catip* was barely expressed during early embryogenesis (Fig. 4F). Thus, we hypothesized that *Dyrk1a* played an important role in early embryonic development. *Dyrk1a* (dual specificity tyrosine phosphorylation–regulated kinase 1A) is a gene encoding a member of the dual-specificity tyrosine phosphorylation–regulated kinase family, which is expressed in developing and mature nervous systems (*25*). It raised a lot of concerns for relating to learning disorders caused by Down syndrome (*26*) and RNA polymerase II (*27*, *28*). However, *Dyrk1a* is rarely studied in TE specification or placental development. To confirm the role of *Dyrk1a* in the ESC-to-TSC transition, we introduced an inhibitor of DYRK1A (harmine) to this process (Fig. 4G). Here, we cultured WT-ESCs in TSC medium with 10 μM harmine to induce the ESC-to-TSC transition. In addition, *Dyrk1a*-inhibited ESCs were able to achieve iTSCs on day 10 by CDCP1-positive enrichment (fig. S4C). The harmine-induced iTSCs presented typical TSC colonies (Fig. 4H) in further culture and were positive for CDX2 and EOMES (fig. S4D). Therefore, inhibition of DYRK1A by harmine could also promote the conversion of TSCs from ESCs; thus, it was interesting to clarify the potential mechanism underlying why *Dyrk1a* deletion promoted the conversion from ESCs to TSCs.

A totipotent-like state of *Dyrk1a*-null ESCs

Since the Notch signaling pathway (*29*), the Hippo/Yap pathway (*30*), and totipotent genes (*31*–*34*) were previously reported to be related to TE specification, we investigated whether the DEGs between the *Dyrk1a*-KO ESCs and the WT-ESCs mapped to the three pathways. The heatmap results showed that some representative totipotent genes were included in the DEGs of *Dyrk1a*-KO ESCs versus the WT-ESCs, while the representative genes of the other two pathways did not change substantially (Fig. 5A). Therefore, we compared our *Dyrk1a*-KO ESCs with the already reported totipotent-like stem cells (TLSCs) globally at transcriptome level (*32*–*34*). The PCA results revealed that *Dyrk1a*-KO ESCs were close to totipotent blastomere–like cells (TBLCs) (*32*), being far from TLSCs (*33*), chemically induced totipotent stem cells (ciTotiSCs) (*34*), and WT-ESCs (Fig. 5B). Consistent with the quantitative analysis, gene set enrichment analysis (GSEA) demonstrated that the DEGs identified in the *Dyrk1a*-KO ESCs versus WT-ESCs were also strongly related to previously reported totipotent genes of TBLCs (Fig. 5C). We further analyzed the overlapping DEGs of *Dyrk1a*-KO ESCs and TBLCs, versus WT-ESCs. There were 2369 co–up-regulated DEGs between *Dyrk1a*-KO ESCs and TBLCs (Fig. 5, D and E), which showed functions of multicellular organism development and others by GO analysis indicating (Fig. 5F). Meanwhile, there were 683 co–down-regulated genes between them (fig. S5, A and B), mapping to regulation of transcription from polymerase II promoter function and other ones (fig. S5C). Recent studies have shown that double homeobox (*Dux* in mice and *DUX4* in humans) can activate the mammalian embryonic genome chromatin landscape and regulate the expression of two-cell embryo (2C) (totipotent-like state cells)–specific genes to govern totipotency (*35*, *36*). DUX or DUX4 can also activate hundreds of endogenous genes, including *Zscan4* family genes, which are precursors of 2C genes (*37*, *38*). Our results showed that *Dux* was up-regulated obviously in the *Dyrk1a*-KO ESCs (fig. S5D); nevertheless, *Zscan4* family genes were almost unchanged compared with those in the WT-ESCs (Fig. 5A). Given that the *Dyrk1a*-KO ESCs did not activate the *Zsacan4* family, it was necessary to determine whether *Dyrk1a*-KO activated another 2C-specific marker, MuERV-L (MuERV-L driving tdTomato), in the cell cultures (normally low). We found that the proportion of MuERV-L$^+$ cells in the *Dyrk1a*-KO ESCs (4.2%) increased compared with that in the WT-ESCs (2.1%) (fig. S5E). Besides, inhibition of DYRK1A in WT-ESCs by 10 μM harmine could also increase that of MuERV-L$^+$ cells (from 2.0 to 7.7%) (fig. S5F), demonstrating that deficiency of *Dyrk1a*-KO could activate MuERV-L. In total, the repression of *Dyrk1a* could shape the partial transcriptome of totipotency including *Dux* and MuERV-L.

Totipotent developmental potentials of *Dyrk1a*-null ESCs

It was important and interesting to address whether *Dyrk1a*-KO ESCs have embryonic and extraembryonic developmental potential as totipotent cells. To answer this question, we conducted a more rigorous experiment by chimeric analysis. We microinjected GFP-labeled *Dyrk1a*-KO ESCs and WT-ESCs (fig. S6A) into mouse four-cell embryos to construct chimeric embryos. In embryonic day 3.5 (E3.5) chimeric blastocysts, the immunofluorescence staining results showed that the *Dyrk1a*-KO ESCs could contribute to both the ICM and TE, while GFP-labeled WT-ESC only integrated into the ICM (Fig. 6A). We further measured the differentiation potentials of the *Dyrk1a*-KO ESCs in E12.5 chimeric embryos in vivo. About 125 WT four-cell embryos were injected with GFP-labeled *Dyrk1a*-KO ESCs or WT-ESCs, separately. There were about 30% embryos that could develop to E12.5. In the *Dyrk1a*-KO group, there were 10 chimeric fetuses, 6 chimeric placentas, and 4 chimeric yolk sacs with contribution of GFP cells (Fig. 6B and fig. S6B), while the WT-ESC group could only generate chimeric fetus (fig. S6B). FACS analysis of these chimeric embryos showed that the *Dyrk1a*-KO GFP$^+$ cells contributed to 43.6% of the fetus, 14.9% of the placenta, and 54.4% of the yolk sac (Fig. 6C). In another parallel experiment, we analyzed the *Dyrk1a*-KO GFP$^+$ ESC–derived chimeric placenta with uninjected placenta as a control (fig. S6C) by FACS analysis. The results showed that there were 4.3% positive cells for TFAP2C (a placental specific marker) among GFP$^+$ cells

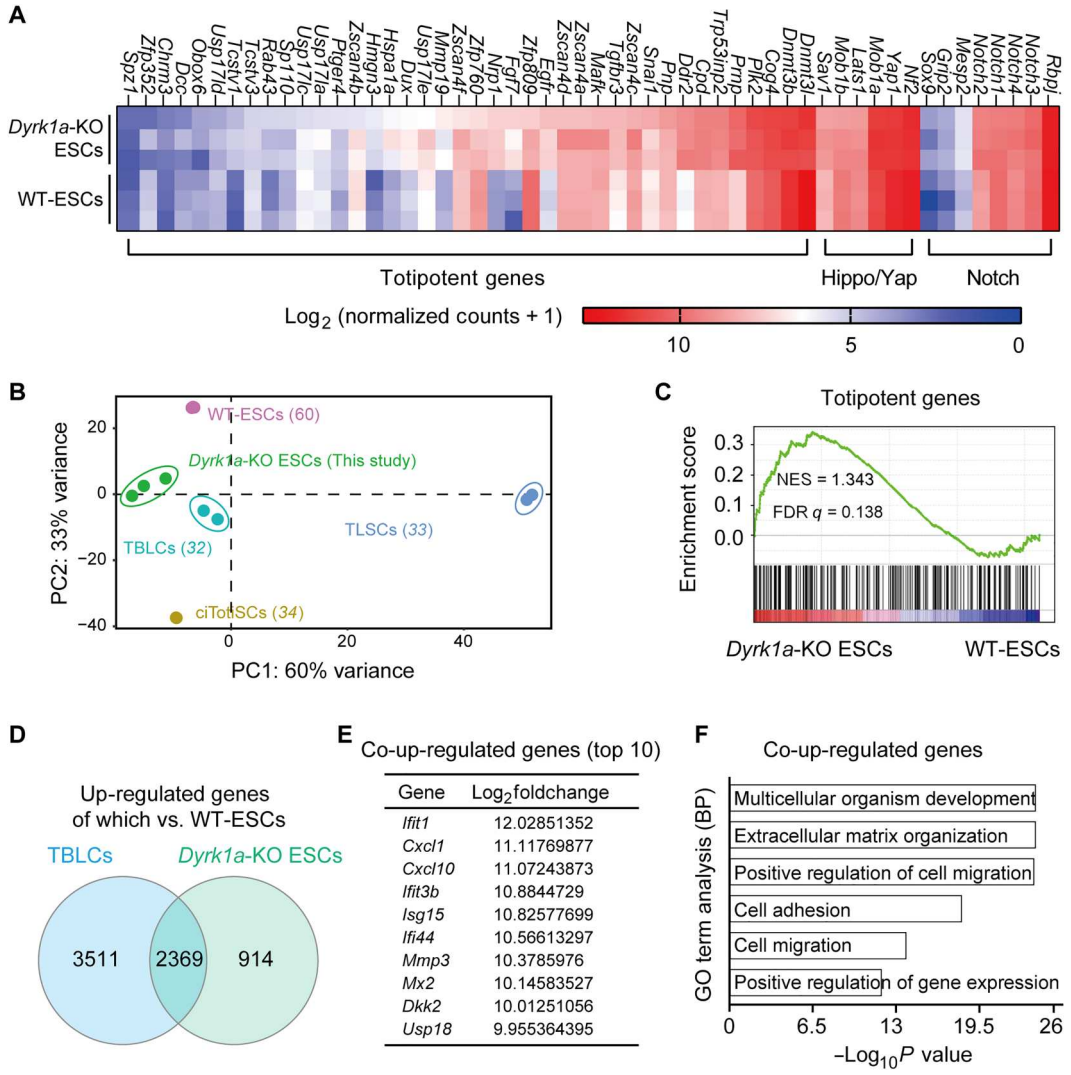

Fig. 5. Bioinformatic analysis of *Dyrk1a*-KO ESCs. (**A**) Heatmap showing some DEGs between *Dyrk1a*-KO ESCs and WT-ESCs mapping to Notch pathway genes, Hippo/Yap pathway genes, and totipotent genes. (**B**) PCA results of RNA-seq datasets from the WT-ESCs, *Dyrk1a*-KO ESCs, TBLCs, TLSCs, and ciTotiSCs. (**C**) GSEA of DEGs associated with totipotent genes in *Dyrk1a*-KO ESCs. NES, normalized enrichment score; FDR, false discovery rate. (**D**) The Venn diagram of the up-regulated DEGs between *Dyrk1a*-KO ESCs and TBLCs, versus WT-ESCs. (**E**) Top 10 genes in the co–up-regulated DEGs between *Dyrk1a*-KO ESCs and TBLCs, versus WT-ESCs. (**F**) GO analysis of co–up-regulated DEGs between *Dyrk1a*-KO ESCs and TBLCs, versus WT-ESCs.

from *Dyrk1a*-KO GFP$^+$ ESC–derived chimeric placenta, suggesting the contribution of *Dyrk1a*-KO GFP$^+$ cells into the placental tissues (fig. S6D). Next, immunofluorescence analysis of the *Dyrk1a*-KO ESC–derived E12.5 chimeric placenta further confirmed that *Dyrk1a*-KO ESCs had developmental potential in the functional placenta (TPBPA-, PROLIFERIN-, TFAP2C-, and GCM1-positive) (Fig. 6D).

Given that the cells themselves could form standard blastocyst-like structures containing ICM and TE and were generally in a totipotent-like state (*32*, *33*, *39*), we cultured *Dyrk1a*-KO ESCs in blastoid medium (*40*) for 6 to 7 days to form blastocyst-like structures, with WT-ESCs and WT-TSCs as controls. Besides, we performed a reoverexpression of *Dyrk1a* in the *Dyrk1a*-KO ESCs as a rescue control (*Dyrk1a*-KO-OE ESCs) (fig. S7A). The *Dyrk1a*-KO ESCs could self-organize into cavitation with a high efficiency to form blastoids with good quality compared with WT blastocyst (Fig. 6E), whereas the *Dyrk1a*-KO-OE ESCs, WT-ESCs, and WT-TSCs could not cavitate to form blastocyst-like structures (fig. S7, B and C). The immunofluorescence results showed that all the *Dyrk1a*-KO blastoids contained OCT4-, CDX2-, and GATA6-positive cells, demonstrating that *Dyrk1a*-KO blastoid structure contained all lineages of blastocyst including ICM, TE, and primitive endoderm (Fig. 6, F and G). We further evaluated the in vivo developmental potentials of the *Dyrk1a*-KO blastoids by embryo transfer into oviducts of pseudo-pregnant mice at 0.5-day postcoitum. At 6.5 days postcoitum, decidualization instead of normal embryos formed from the *Dyrk1a*-KO blastoids (fig. S7D). Totally, all the above evidences proved that the repression of *Dyrk1a* could induce a totipotent-like state and enable ESCs to efficiently form blastoids.

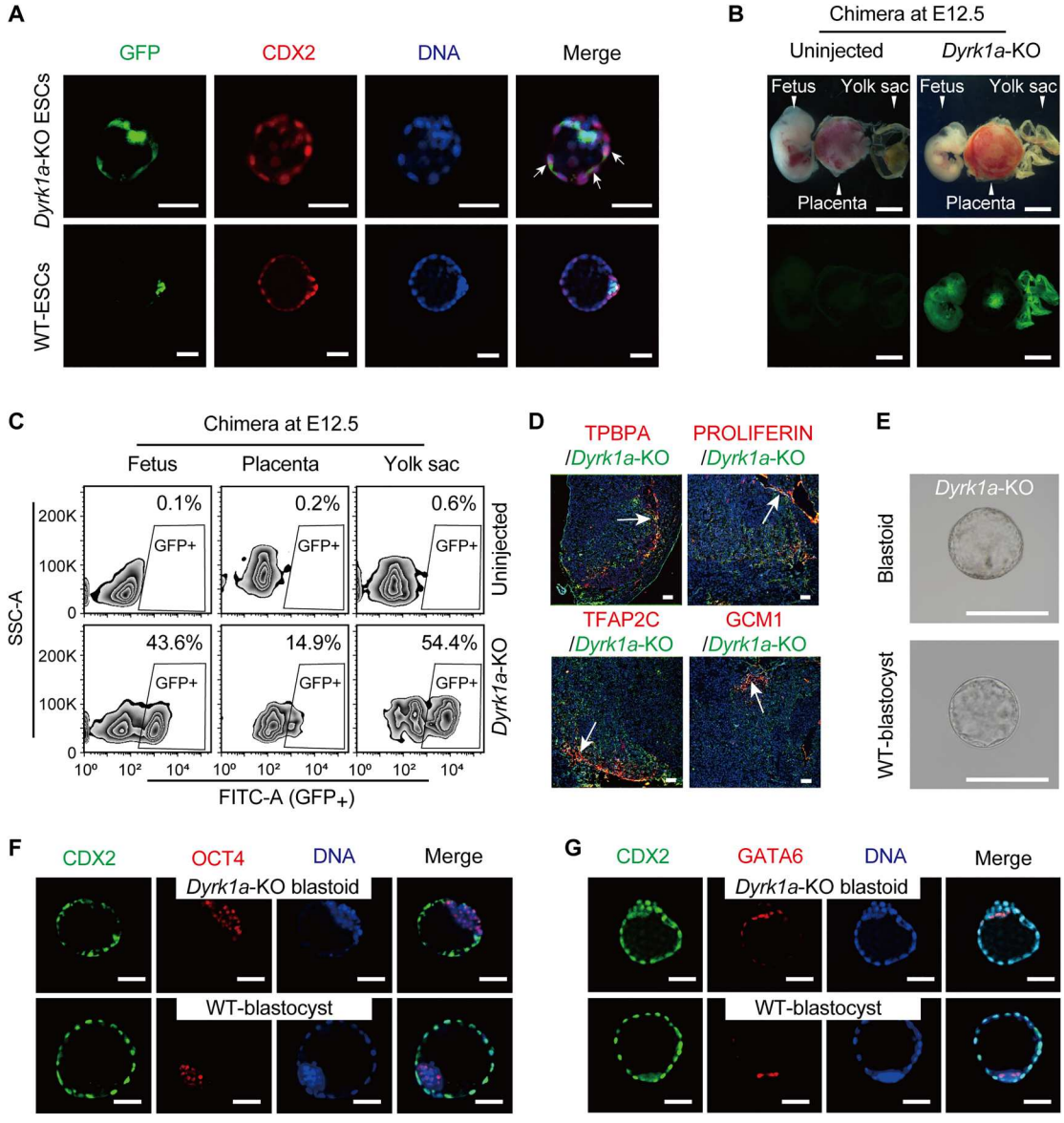

Fig. 6. The totipotency of *Dyrk1a*-KO ESCs. (**A**) Immunofluorescence of CDX2 (red) in chimeric blastocyst from GFP-labeled *Dyrk1a*-KO ESCs and WT-ESCs. Donor cells were GFP$^+$ (green), and DNA was stained with Hoechst 33342 (blue). Scale bars, 50 μm. (**B**) Images of E12.5 chimeras (including fetus, placenta, and yolk sac) derived from GFP-labeled *Dyrk1a*-KO ESCs, with uninjected embryo as a control. Scale bars, 5 mm. (**C**) FACS analysis of GFP$^+$ cells in the fetus, placenta, and yolk sac in E12.5 chimeras derived from the GFP-labeled *Dyrk1a*-KO ESCs. The percentages of GFP$^+$ cells in the fetus, placenta, and yolk sac of *Dyrk1a*-KO chimera were 43.6, 14.9, and 54.4%, respectively; uninjected ones (fetus, placenta, and yolk sac) at E12.5 were as a control. (**D**) Immunofluorescence staining of placental-specific markers (TPBPA, PROLIFERIN, TFAP2C, and GCM1, each in red) in the GFP-labeled *Dyrk1a*-KO ESC–derived E12.5 placentas. The white arrows indicated that GFP-positive cells could contribute to the tissues expressing specific markers. Scale bars, 75 μm. (**E**) Phase-contrast images of blastoid derived from the *Dyrk1a*-KO ESCs. WT-blastocyst was used as a control. Scale bars, 10 μm. (**F**) Immunofluorescence of CDX2 (green) and OCT4 (red) in blastoid derived from the *Dyrk1a*-KO ESCs. WT-blastocyst was used as a control. DNA was stained with Hoechst 33342 (blue). Scale bars, 50 μm. (**G**) Immunofluorescence of CDX2 (green) and GATA6 (red) in blastoid derived from the *Dyrk1a*-KO ESCs. WT-blastocyst was used as a control. DNA was stained with Hoechst 33342 (blue). Scale bars, 50 μm.

DISCUSSION

The first cell fate decision is an important event in preimplantation development, but the underlying mechanisms are unknown. In vitro–established ESCs and TSCs provide convenient platforms to study cell fate conversion, which is very useful for understanding the first cell fate decision. Many studies have shown that the overexpression of TE-specific TFs, including *Cdx2*, *Elf5*, and *Gata3*, can induce TE cell fate in ESCs (*4*, *41*, *42*). In addition to TSC-specific TFs, some epigenetic regulators in ESCs were found to play important roles in the initiation of TE specification. *Tet1*-deleted ESCs can activate *Elf5* and exhibit TE cell properties (*43*). Another study demonstrated that a histone H3K9 methyltransferase, SETDB1, was also essential for TE specification from "ground-state" ESCs (*44*). All of the above results proved that ESCs are good choices for investigating the key modules of TE specification. Haploid stem cells are powerful tools for genetic studies, mainly owing to their single-genome and

self-renewal features. Many attempts at genetic screening in mutant haESCs have been achieved, such as determining key exits of the ground state (*45*) or "formative state" in ESCs (*46*). Hence, we chose mouse haESCs as a powerful tool to select crucial genes regulating TE specification, not only because of their pluripotency and self-renewal but also because of their high efficiency in obtaining homozygous mutants. Although a similar genetic screening was reported recently, only a single gene (*Zfp281*) was found to be related to TE specification, which did not show the advances of the haploid system (*47*). Nevertheless, we present numerous useful genes related to TE specification using the haploid platform (Fig. 2C and fig. S2B). In addition, most other studies focus on overexpression of some specific genes to obtain desired phenotypes, whereas our approach is a loss-of-function genetic screening to uncover many potential trapped genes. *Rif1*, also found in our list, had been proved played critical role in the conversion of TSCs and regulated the pluripotency-to-totipotency transition (*48*).

Whether TSCs derived by overexpressing a single factor present a stable TSC state is controversial (*49*). Although the cell cultures showed TSC-like morphology after overexpression of *Cdx2* in WT-ESCs, the cell fate conversion was still incomplete (*4*). Similarly, *Elf5*-overexpressing ESCs were able to enter the TE lineage but could not maintain an undifferentiated TSC state (*42*). In addition, similar phenomenon was observed in *Gata3*-overexpressing iTSCs, which were prone to differentiate into other lineages (*41*). Therefore, the accurate conversion of TSCs from ESCs with self-renewal is important for TE specification studies. In our research, iTSCs derived from *Catip*-KO or *Dyrk1a*-KO ESCs morphologically resembled WT-TSCs (Fig. 4A) and could be expanded for more than 30 passages without differentiation. Although iTSCs generated from differentiated ESCs show heterogenicity, the TSC-specific antibodies CDCP1 and CD40 were chosen to enrich them. For enrichment of authentic iTSCs, both *Catip*-KO and *Dyrk1a*-KO iTSCs were purified two to three times for CDCP1 and CD40 positivity. The achieved *Catip*-KO and *Dyrk1a*-KO iTSCs stably showed high percentages of CDCP1 and CD40 double-positive cells (fig. S4A). These iTSCs have the potential to further differentiate into terminal trophoblast lineages (Fig. 4C and fig. S4B), suggesting that iTSCs have differentiation potentials similar to those of WT-TSCs. To investigate whether $Dyrk1a^{+/-}$ ESCs could also be converted into TSCs, we cultured $Dyrk1a^{+/-}$ ESCs and $Dyrk1a^{-/-}$ ESCs in TSC medium to induce TE specification, with WT-ESCs as a control. After 10 days of induction, both $Dyrk1a^{-/-}$ and $Dyrk1a^{+/-}$ groups could generate TSC-like colonies, whereas the WT-ESCs group could not (fig. S8A). To evaluate the efficiency of induction, we performed FACS by incubated with CD40 cytometric antibody. The results showed that there were 76.5% of cells that were CD40-positive in homozygous mutant ($Dyrk1a^{-/-}$) group, while 15.3% in heterozygous mutant ($Dyrk1a^{+/-}$) group (fig. S8B). Together, $Dyrk1a^{+/-}$ ESCs could also promote the ESC-to-TSC transition, although slower than $Dyrk1a^{-/-}$ ESCs.

Naturally, murine ESCs cannot transition to TSC properties only if they are genetically edited. Determination of why *Dyrk1a*-KO ESCs tend to generate functional iTSCs is very important. Many groups have obtained extended pluripotent stem cells by optimizing the culture medium, and these cells can further differentiate into ESCs and TSCs (*50*–*52*). Moreover, 2C-like cells among ESCs were able to develop into the ICM and TE (*38*, *53*). All of the above cells present a totipotent-like state, which is logical if repression of *Dyrk1a* can trigger a pluripotency-to-totipotency transition to achieve cell fate alternation. Our results demonstrate that *Dyrk1a*-KO activates some totipotent genes including the 2C-like reporter (MuERV-L) and *Dux* (Fig. 5, A to C, and fig. S5, D and E). *Dyrk1a*-KO ESCs could contribute to fetus, placenta, and yolk sac in vivo in chimera production experiment (Fig. 6, B and C, and fig. S6C), even self-organizing to form blastocyst-like structure in vitro (Fig. 6, E to G, and fig. S7). All the above evidences demonstrated that *Dyrk1a*-KO ESCs presented a totipotent-like state with developmental potentials. Whether *Catip*-KO ESCs showed similar developmental potentials to *Dyrk1a*-KO ESCs in these assays warranted more investigations.

To conclude, our study provides a genome-scale screening strategy for TE specification with the haploid system. Thousands of mutations involved in the conversion from ESCs to TSCs were uncovered, including *Catip* and *Dyrk1a*, by validation experiments. Moreover, repression of *Dyrk1a* activates a totipotent-like state, which might be the reason for the efficient conversion of TSCs. All the findings shed light on the mechanisms underlying cell fate alternation between ESCs and TSCs.

MATERIALS AND METHODS

Mice

All mice were purchased from Beijing Vital River Laboratory Animal Technology Co. Ltd. in specific pathogen-free grade. They were used according to the guidelines of Animal Care and Use Committee of Nankai University.

Mouse haESC culture and cytometric sorting

Mouse haESCs were cultured in t2i/L ESC medium on feeder cells as described previously (*54*). For enrichment of haploid cells, haESCs were incubated with Hoechst 33342 (3 μg/ml; Invitrogen, H3570) for 25 min at 37°C and filtered with a 40-μm cell strainer (BD Biosciences, 352340). The well-prepared samples were sorted with a diploid control on a MoFlo Astrios EQ sorter (Beckman) every period.

Genomic mutation

For the gene-trapping experiment, a PB vector (SBI, PB513B-1) with slight modification was used to construct the PB-dual-SA vector. Briefly, two copies of alternatively spliced sequences (SA) (*55*) were inserted into the PB vector with two orientations. For introduction of mutations into the genome of haESCs, PB-dual-SA and PBase (SBI, PB210PA-1) vectors were electroporated into them with an electroporator (Invitrogen, NEON) at 1400 V for 10 ms with three pulses. The transfected cells were treated with puromycin (1 μg/ml; Gibco, A1113802) to enrich mutated cells, with nontransfected haESCs as a negative control.

Genetic screening of TE specification

Approximately 1×10^7 mutant haESCs were seeded in gelatin-precoated (Sigma-Aldrich, V900863) dishes and cultured in TSC medium {70 CM medium with F4H [human recombinant FGF4 (25 ng/ml; PeproTech, AF-100-31) and heparin (1 mg/ml; Millipore, 375095-100KU)]} (*2*) to induce the transition from ESCs to TSCs. After 10 days, CDCP1 (R&D, AF4515)–positive cells were harvested from the cell cultures according to the instructions of the antibody. The CDCP1-positive cells were further cultured in

TSC medium to establish iTSCs. The total DNA of screened mutant iTSCs was amplified by splinkerette PCR (*56*), the product of which was sent to a local company for next-generation sequencing. All the primers are listed in table S1. For analysis of the insertion sites of PB in screened iTSCs, reads were mapped to the reference genome (mm39) and analyzed using HaSAPPy as described previously (*23*).

WB and immunostaining

The WB experiments were performed according to our previous work (*57*). For immunostaining, chimeric embryos were treated with Tyrode′s solution (Sigma-Aldrich, T1788) for 20 s to remove the zona pellucida and washed three times with M2 medium (Sigma-Aldrich, M7167). Cells or embryos were fixed with 4% paraformaldehyde at room temperature for 30 min. After being washed three times in phosphate-buffered saline, the samples were blocked and permeabilized in 2% bovine serum albumin (diluted with 0.5% Triton X-100) at room temperature for 1 hour (embryos need 4 hour in this step). Then, the samples were incubated with primary antibodies at 4°C overnight. Next, the phosphate-buffered saline–washed samples were incubated with the secondary antibodies at room temperature for 1 hour. Hoechst 33342 (Thermo Fisher Scientific, H3570) was used for nuclear staining. The primary antibodies used for WB and immunostaining were as follows: CDX2 (Biogenes, MU392A), OCT4 (Abcam, ab181557), α-tubulin (Sungene, KM9007), EOMES (Abcam, ab23345), PROLIFERIN (Santa Cruz Biotechnology, SC271891), TPBPA (Abcam, 104401), GATA6 (Cell Signaling Technology, 5851S), GCM1 (Santa Cruz Biotechnology, sc-101173), SDC1 (Abclonal, A4174), and TFAP2C (also known as AP-2γ) (Santa Cruz Biotechnology, sc-12762). Secondary antibodies included fluorescein isothiocyanate (FITC) goat anti-rabbit immunoglobulin G (IgG) (H + L) (Abclonal, AS011), FITC goat anti-mouse IgG (H + L) (Abclonal, AS001), goat anti-rabbit IgG(H + L)–horseradish peroxidase (Sungene, LK2001), and Cy3 goat anti-rabbit IgG (H + L) (Abclonal, AS007). Immunofluorescence images were observed on a confocal laser scanning microscope (Leica, TCS SP8).

Transition from candidate gene-KO ESCs to TSCs

For *Catip* and *Dyrk1a* KO, single-guide RNAs (sgRNAs) were designed using a website (http://crispor.tefor.net/). KO plasmids were constructed according to a previous report (*57*). For *Catip*-KO or *Dyrk1a*-KO ESCs, approximately 2×10^6 WT-ESCs were electroporated with 8 μg of sgRNA-Cas9 plasmids. GFP-positive cells were sorted by FACS 48 hours after electroporation and plated back in ESC medium for further culture. Subclones were randomly picked, and genotyping was performed. The primers for genotyping are listed in table S1. For inhibition of DYRK1A, different concentrations of harmine (MCE, HY-N0737A) were used as indicated.

Thereafter, the *Catip*-KO ESCs, *Dyrk1a*-KO ESCs, harmine-treated ESCs (10 μM), and WT-ESCs were cultured in TSC medium for the conversion of TSCs as described, which were enriched for CDCP1-positive cells on day 10 of transition.

Differentiation of iTSCs

For determination of the differentiation potentials of iTSCs, cells were plated in gelatin precoated dishes and cultured in TSC medium without F4H. Samples were fixed for immunostaining or were collected on days 0, 2, 4, and 6 for further analysis. For DNA content analysis, cells were dissociated into single cells with 0.25% trypsin-EDTA (Thermo Fisher Scientific, 25200072), fixed with 75% ethanol overnight, and then incubated with propidium iodide (50 μg/ml; Sigma-Aldrich, P4170) for 30 min at 37°C. DNA content analysis was performed on an analyzing flow cytometer (BD Biosciences, LSRFortessa).

Teratoma formation and chimeric assay

For teratoma analysis, approximately 1×10^7 *Catip*-KO ESCs, *Dyrk1a*-KO ESCs, and WT-ESCs were injected subcutaneously into the limbs of 8-week-old male severe combined immunodeficiency mice independently. Fully formed teratomas were dissected 3 weeks later, fixed in 4% paraformaldehyde, embedded in paraffin, sectioned, and stained with hematoxylin and eosin for further analysis.

To obtain GFP-labeled cell lines, we constructed a β-*actin*–*GFP* reporter system. Briefly, the homologous arms of a *pax6-GFP* donor vector (*19*) were replaced with those of β-*actin*. The β-*actin*–sgRNAs were designed using a website (http://crispor.tefor.net/) and cloned into the pSpCas9n (BB)–2A-GFP backbone (Addgene, 48140). Then, a combination of 3 μg of β-*actin*–*GFP* donor vector and 2 μg of β-*actin*–sgRNA plasmids was independently cotransfected into 2×10^6 WT-ESCs and *Dyrk1a*-KO ESCs. E3.5 and E12.5 chimeric embryos were generated by microinjecting GFP-labeled WT-ESCs or *Dyrk1a*-KO ESCs into CD-1 background four-cell embryos. In E12.5 embryos, the placentas, yolk sacs, and fetuses were digested into single cells to analyze the percentages of GFP-positive cells on an analyzing flow cytometer (BD Biosciences, LSRFortessa).

Generation of blastoids

Blastoids were generated as described previously (*40*). Briefly, *Dyrk1a*-KO, WT-ESCs, and WT-TSCs were cultured in blastoid medium composed of 25% TSC basal medium (TSC medium without F4H), 25% N2B27 basal medium (N2B27 medium without 2i and LIF) (*58*), and 50% K+ Simplex Optimised Medium (KSOM) (Millipore, MR-020P-5F) with 2 μM Y-27632 (MCE, HY-10071), FGF4 (12.5 ng/ml; PeproTech, 10031), heparin (0.5 mg/ml; Millipore, 375095), 3 μM CHIR99021 (MCE, HY-10182), bone morphogenetic protein 4 (BMP4) (5 ng/ml; PeproTech, 123005ET), and 0.5 μM A83-01 (MCE, HY-10432). A total of 300 to 500 cells were seeded into one well of a 96-well ultra-low attachment plate (Corning, 7007) (*59*). After 5 days, an aggregating structure appeared in each well with many blastocyst-like structures attached to it (fig. S8C). Thereafter, these blastocyst-like structures were separated from the original aggregate gently by a mouth pipette under a stereomicroscope to another well, each of which would further self-organize to form blastoids. After 1 to 2 days, dozens of blastoids could be observed in a single well (fig. S8D) and then be manually collected using a mouth pipette for subsequent experiments.

Analysis of RNA-seq data and WGBS

Total RNA was separately extracted from *Catip*-KO ESCs, *Dyrk1a*-KO ESCs, WT-ESCs (*60*), *Catip*-KO iTSCs, *Dyrk1a*-KO iTSCs, and WT-TSCs and sequenced by a local company (Novogene) to generate raw data based on PE150. The clean data were used for downstream analysis. The reference genome (mm39) and gene model annotation GTF files were directly downloaded from the USCS and GENCODE websites. The index of the reference genome was

built, and the paired-end clean reads were aligned to the reference genome using STAR software (v2.5.3). The read counts were calculated and output with HTSeq for subsequent analysis. DEG analysis was performed using the DEseq2 R package. The correlation coefficients were calculated on the basis of the normalized count values of all genes in each sample. Genes with expression fold changes of >2 and adjusted $P < 0.05$ according to the DEseq2 results were used for GO analysis and volcano plot generation. GSEA (v4.2.4) was used to determine whether the gene set of 2C, and totipotency genes was statistically enriched in *Dyrk1a*-KO ESCs versus WT-ESCs by RNA-seq data. The normalized enrichment score and false discovery rate *q* value were indicated for the relative gene set.

For WGBS analysis, genomic DNA was separately extracted from *Catip*-KO ESCs, *Dyrk1a*-KO ESCs, WT-ESCs, *Catip*-KO iTSCs, *Dyrk1a*-KO iTSCs, and WT-TSCs and sequenced by a local company (Novogene) to generate raw data based on Illumina Novaseq platform. The clean data were aligned to the mm39 genome using bismark. CpG sites with a read depth of <5 were filtered. The resultant bedGraph file was visualized in Itegrative Genomics Viewer (IGV).

Supplementary Materials

This PDF file includes:

Supplementary Text
Figs. S1 to S8
Table S1

REFERENCES AND NOTES

1. C. Yao, W. Zhang, L. Shuai, The first cell fate decision in pre-implantation mouse embryos. *Cell Regen.* **8**, 51–57 (2019).
2. S. Tanaka, T. Kunath, A. K. Hadjantonakis, A. Nagy, J. Rossant, Promotion of trophoblast stem cell proliferation by FGF4. *Science* **282**, 2072–2075 (1998).
3. H. Niwa, J. Miyazaki, A. G. Smith, Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat. Genet.* **24**, 372–376 (2000).
4. H. Niwa, Y. Toyooka, D. Shimosato, D. Strumpf, K. Takahashi, R. Yagi, J. Rossant, Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell* **123**, 917–929 (2005).
5. T. Wu, H. Wang, J. He, L. Kang, Y. Jiang, J. Liu, Y. Zhang, Z. Kou, L. Liu, X. Zhang, S. Gao, Reprogramming of trophoblast stem cells into pluripotent stem cells by Oct4. *Stem Cells* **29**, 755–763 (2011).
6. R. Yagi, M. J. Kohn, I. Karavanova, K. J. Kaneko, D. Vullhorst, M. L. DePamphilis, A. Buonanno, Transcription factor TEAD4 specifies the trophectoderm lineage at the beginning of mammalian development. *Development* **134**, 3827–3836 (2007).
7. Y. Hirate, K. Cockburn, J. Rossant, H. Sasaki, Tead4 is constitutively nuclear, while nuclear vs. cytoplasmic Yap distribution is regulated in preimplantation mouse embryos. *Proc. Natl. Acad. Sci. U.S.A.* **109**, E3389–E3390 (2012).
8. N. Nishioka, K. I. Inoue, K. Adachi, H. Kiyonari, M. Ota, A. Ralston, N. Yabuta, S. Hirahara, R. O. Stephenson, N. Ogonuki, R. Makita, H. Kurihara, E. M. Morin-Kensicki, H. Nojima, J. Rossant, K. Nakao, H. Niwa, H. Sasaki, The Hippo signaling pathway components Lats and Yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass. *Dev. Cell* **16**, 398–410 (2009).
9. H. Wang, C. Zang, X. S. Liu, J. C. Aster, The role of Notch receptors in transcriptional regulation. *J. Cell. Physiol.* **230**, 982–988 (2015).
10. H. Benchetrit, S. Herman, N. van Wietmarschen, T. Wu, K. Makedonski, N. Maoz, N. Yom Tov, D. Stave, R. Lasry, V. Zayat, A. Xiao, P. M. Lansdorp, S. Sebban, Y. Buganim, Extensive nuclear reprogramming underlies lineage conversion into functional trophoblast stem-like cells. *Cell Stem Cell* **17**, 543–556 (2015).
11. C. Kubaczka, C. E. Senner, M. Cierlitza, M. J. Araúzo-Bravo, P. Kuckenberg, M. Peitz, M. Hemberger, H. Schorle, Direct induction of trophoblast stem cells from murine fibroblasts. *Cell Stem Cell* **17**, 557–568 (2015).
12. F. Kaiser, C. Kubaczka, M. Graf, N. Langer, J. Langkabel, L. Arévalo, H. Schorle, Choice of factors and medium impinge on success of ESC to TSC conversion. *Placenta* **90**, 128–137 (2020).
13. U. Nosi, F. Lanner, T. Huang, B. Cox, Overexpression of trophoblast stem cell-enriched microRNAs promotes trophoblast fate in embryonic stem cells. *Cell Rep.* **19**, 1101–1109 (2017).
14. M. Leeb, A. Wutz, Derivation of haploid embryonic stem cells from mouse embryos. *Nature* **479**, 131–134 (2011).
15. U. Elling, J. Taubenschmid, G. Wirnsberger, R. O'Malley, S. P. Demers, Q. Vanhaelen, A. I. Shukalyuk, G. Schmauss, D. Schramek, F. Schnuetgen, H. von Melchner, J. R. Ecker, W. L. Stanford, J. Zuber, A. Stark, J. M. Penninger, Forward and reverse genetics through derivation of haploid mouse embryonic stem cells. *Cell Stem Cell* **9**, 563–574 (2011).
16. A. Monfort, G. di Minin, A. Postlmayr, R. Freimann, F. Arieti, S. Thore, A. Wutz, Identification of spen as a crucial factor for xist function through forward genetic screening in haploid embryonic stem cells. *Cell Rep.* **12**, 554–561 (2015).
17. M. Xu, Y. Zhao, W. Zhang, M. Geng, Q. Liu, Q. Gao, L. Shuai, Genome-scale screening in a rat haploid system identifies Thop1 as a modulator of pluripotency exit. *Cell Prolif.* **55**, e13209 (2022).
18. G. Liu, X. Wang, Y. Liu, M. Zhang, T. Cai, Z. Shen, Y. Jia, Y. Huang, Arrayed mutant haploid embryonic stem cell libraries facilitate phenotype-driven genetic screens. *Nucleic Acids Res.* **45**, e180 (2017).
19. Q. Gao, W. Zhang, L. Ma, X. Li, H. Wang, Y. Li, R. Freimann, Y. Yu, L. Shuai, A. Wutz, Derivation of haploid neural stem cell lines by selection for a Pax6-GFP reporter. *Stem Cells Dev.* **27**, 479–487 (2018).
20. S. Zhao, E. Jiang, S. Chen, Y. Gu, A. J. Shangguan, T. Lv, L. Luo, Z. Yu, PiggyBac transposon vectors: The tools of the human gene encoding. *Transl. Lung Cancer Res.* **5**, 120–125 (2016).
21. P. J. Rugg-Gunn, B. J. Cox, F. Lanner, P. Sharma, V. Ignatchenko, A. C. H. McDonald, J. Garner, A. O. Gramolini, J. Rossant, T. Kislinger, Cell-surface proteomics identifies lineage-specific markers of embryo-derived stem cells. *Dev. Cell* **22**, 887–901 (2012).
22. M. A. Li, S. J. Pettitt, K. Yusa, A. Bradley, Genome-wide forward genetic screens in mouse ES cells. *Methods Enzymol.* **477**, 217–242 (2010).
23. G. Di Minin, A. Postlmayr, A. Wutz, HaSAPPy: A tool for candidate identification in pooled forward genetic screens of haploid mammalian cells. *PLOS Comput. Biol.* **14**, e1005950 (2018).
24. J. Wu, B. Huang, H. Chen, Q. Yin, Y. Liu, Y. Xiang, B. Zhang, B. Liu, Q. Wang, W. Xia, W. Li, Y. Li, J. Ma, X. Peng, H. Zheng, J. Ming, W. Zhang, J. Zhang, G. Tian, F. Xu, Z. Chang, J. Na, X. Yang, W. Xie, The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature* **534**, 652–657 (2016).
25. E. Martí, X. Altafaj, M. Dierssen, S. de la Luna, V. Fotaki, M. Alvarez, M. Pérez-Riba, I. Ferrer, X. Estivill, Dyrk1A expression pattern supports specific roles of this kinase in the adult central nervous system. *Brain Res.* **964**, 250–263 (2003).
26. M. Ohira, N. Seki, T. Nagase, E. Suzuki, N. Nomura, O. Ohara, M. Hattori, Y. Sakaki, T. Eki, Y. Murakami, T. Saito, H. Ichikawa, M. Ohki, Gene identification in 1.6-Mb region of the Down syndrome region on chromosome 21. *Genome Res.* **7**, 47–58 (1997).
27. D. Yu, C. Cattoglio, Y. Xue, Q. Zhou, A complex between DYRK1A and DCAF7 phosphorylates the C-terminal domain of RNA polymerase II to promote myogenesis. *Nucleic Acids Res.* **47**, 4462–4475 (2019).
28. C. D. Vona, D. Bezdan, A. B. M. M. K. Islam, E. Salichs, N. López-Bigas, S. Ossowski, S. de la Luna, Chromatin-wide profiling of DYRK1A reveals a role as a gene-specific RNA polymerase II CTD kinase. *Mol. Cell* **57**, 506–520 (2015).
29. T. Rayon, S. Menchero, A. Nieto, P. Xenopoulos, M. Crespo, K. Cockburn, S. Cañon, H. Sasaki, A. K. Hadjantonakis, J. L. de la Pompa, J. Rossant, M. Manzanares, Notch and hippo converge on Cdx2 to specify the trophectoderm lineage in the mouse blastocyst. *Dev. Cell* **30**, 410–422 (2014).
30. T. Frum, A. Ralston, Cell signaling and transcription factors regulating cell fate during formation of the mouse blastocyst. *Trends Genet.* **31**, 402–410 (2015).
31. Y. Xu, J. Zhao, Y. Ren, X. Wang, Y. Lyu, B. Xie, Y. Sun, X. Yuan, H. Liu, W. Yang, Y. Fu, Y. Yu, Y. Liu, R. Mu, C. Li, J. Xu, H. Deng, Derivation of totipotent-like stem cells with blastocyst-like structure forming potential. *Cell Res.* **32**, 513–529 (2022).
32. H. Shen, M. Yang, S. Li, J. Zhang, B. Peng, C. Wang, Z. Chang, J. Ong, P. Du, Mouse totipotent stem cells captured and maintained through spliceosomal repression. *Cell* **184**, 2843–2859.e20 (2021).
33. M. Yang, H. Yu, X. Yu, S. Liang, Y. Hu, Y. Luo, Z. Izsvák, C. Sun, J. Wang, Chemical-induced chromatin remodeling reprograms mouse ESCs to totipotent-like stem cells. *Cell Stem Cell* **29**, 400–418.e13 (2022).
34. Y. Hu, Y. Yang, P. Tan, Y. Zhang, M. Han, J. Yu, X. Zhang, Z. Jia, D. Wang, K. Yao, H. Pang, Z. Hu, Y. Li, T. Ma, K. Liu, S. Ding, Induction of mouse totipotent stem cells by a defined chemical cocktail. *Nature* **617**, 792–797 (2022).
35. G. Yang, L. Zhang, W. Liu, Z. Qiao, S. Shen, Q. Zhu, R. Gao, M. Wang, M. Wang, C. Li, M. Liu, J. Sun, L. Wang, W. Liu, X. Cui, K. Zhao, R. Zang, M. Chen, Z. Liang, L. Wang, X. Kou, Y. Zhao, H. Wang, Y. Wang, S. Gao, J. Chen, C. Jiang, Dux-mediated corrections of aberrant h3k9ac

during 2-cell genome activation optimize efficiency of somatic cell nuclear transfer. *Cell Stem Cell* **28**, 150–163.e5 (2021).
36. F. Yang, X. Huang, R. Zang, J. Chen, M. Fidalgo, C. Sanchez-Priego, J. Yang, A. Caichen, F. Ma, T. Macfarlan, H. Wang, S. Gao, H. Zhou, J. Wang, DUX-miR-344-ZMYM2-mediated activation of MERVL LTRs induces a totipotent 2C-like state. *Cell Stem Cell* **26**, 234–250.e7 (2020).
37. P. G. Hendrickson, J. A. Doráis, E. J. Grow, J. L. Whiddon, J. W. Lim, C. L. Wike, B. D. Weaver, C. Pflueger, B. R. Emery, A. L. Wilcox, D. A. Nix, C. M. Peterson, S. J. Tapscott, D. T. Carrell, B. R. Cairns, Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat. Genet.* **49**, 925–934 (2017).
38. D. Rodriguez-Terrones, X. Gaume, T. Ishiuchi, A. Weiss, A. Kopp, K. Kruse, A. Penning, J. M. Vaquerizas, L. Brino, M. E. Torres-Padilla, A molecular roadmap for the emergence of early-embryonic-like cells in culture. *Nat. Genet.* **50**, 106–119 (2018).
39. C. Kime, H. Kiyonari, S. Ohtsuka, E. Kohbayashi, M. Asahi, S. Yamanaka, M. Takahashi, K. Tomoda, Induced 2C expression and implantation-competent blastocyst-like cysts from primed pluripotent stem cells. *Stem Cell Reports* **13**, 485–498 (2019).
40. R. Li, C. Zhong, Y. Yu, H. Liu, M. Sakurai, L. Yu, Z. Min, L. Shi, Y. Wei, Y. Takahashi, H. K. Liao, J. Qiao, H. Deng, E. Nuñez-Delicado, C. Rodriguez Esteban, J. Wu, J. C. Izpisua Belmonte, Generation of blastocyst-like structures from mouse embryonic and adult cell cultures. *Cell* **179**, 687–702.e18 (2019).
41. A. Ralston, B. J. Cox, N. Nishioka, H. Sasaki, E. Chea, P. Rugg-Gunn, G. Guo, P. Robson, J. S. Draper, J. Rossant, Gata3 regulates trophoblast development downstream of Tead4 and in parallel to Cdx2. *Development* **137**, 395–403 (2010).
42. R. K. Ng, W. Dean, C. Dawson, D. Lucifero, Z. Madeja, W. Reik, M. Hemberger, Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5. *Nat. Cell Biol.* **10**, 1280–1290 (2008).
43. K. P. Koh, A. Yabuuchi, S. Rao, Y. Huang, K. Cunniff, J. Nardone, A. Laiho, M. Tahiliani, C. A. Sommer, G. Mostoslavsky, R. Lahesmaa, S. H. Orkin, S. J. Rodig, G. Q. Daley, A. Rao, Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* **8**, 200–213 (2011).
44. K. Wu, H. Liu, Y. Wang, J. He, S. Xu, Y. Chen, J. Kuang, J. Liu, L. Guo, D. Li, R. Shi, L. Shen, Y. Wang, X. Zhang, J. Wang, D. Pei, J. Chen, SETDB1-mediated cell fate transition between 2c-like and pluripotent states. *Cell Rep.* **30**, 25–36.e6 (2020).
45. M. Leeb, S. Dietmann, M. Paramor, H. Niwa, A. Smith, Genetic exploration of the exit from self-renewal using haploid embryonic stem cells. *Cell Stem Cell* **14**, 385–393 (2014).
46. A. Lackner, R. Sehlke, M. Garmhausen, G. Giuseppe Stirparo, M. Huth, F. Titz-Teixeira, P. van der Lelij, J. Ramesmayer, H. F. Thomas, M. Ralser, L. Santini, E. Galimberti, M. Sarov, A. F. Stewart, A. Smith, A. Beyer, M. Leeb, Cooperative genetic networks drive embryonic stem cell transition from naïve to formative pluripotency. *EMBO J.* **40**, e105776 (2021).
47. T. Ishiuchi, H. Ohishi, T. Sato, S. Kamimura, M. Yorino, S. Abe, A. Suzuki, T. Wakayama, M. Suyama, H. Sasaki, Zfp281 shapes the transcriptome of trophoblast stem cells and is essential for placental development. *Cell Rep.* **27**, 1742–1754.e6 (2019).
48. W. Zhang, C. Yao, Y. Luo, Q. Li, Q. Zhao, Y. Zhao, Y. Wang, M. Geng, Q. Wang, M. Xu, S. Sun, D. Wu, Q. Gao, X. Wu, L. Shuai, Rif1 and Hmgn3 regulate the conversion of murine trophoblast stem cells. *Cell Rep.* **38**, 110570 (2022).
49. F. Cambuli, A. Murray, W. Dean, D. Dudzinska, F. Krueger, S. Andrews, C. E. Senner, S. J. Cook, M. Hemberger, Epigenetic memory of the first cell fate decision prevents complete ES cell reprogramming into trophoblast. *Nat. Commun.* **5**, 5538 (2014).
50. Y. Yang, B. Liu, J. Xu, J. Wang, J. Wu, C. Shi, Y. Xu, J. Dong, C. Wang, W. Lai, J. Zhu, L. Xiong, D. Zhu, X. Li, W. Yang, T. Yamauchi, A. Sugawara, Z. Li, F. Sun, X. Li, C. Li, A. He, Y. du, T. Wang, C. Zhao, H. Li, X. Chi, H. Zhang, Y. Liu, C. Li, S. Duo, M. Yin, H. Shen, J. C. I. Belmonte, H. Deng, Derivation of pluripotent stem cells with in vivo embryonic and extraembryonic potency. *Cell* **169**, 243–257.e25 (2017).
51. J. Yang, D. J. Ryan, W. Wang, J. C. H. Tsang, G. Lan, H. Masaki, X. Gao, L. Antunes, Y. Yu, Z. Zhu, J. Wang, A. A. Kolodziejczyk, L. S. Campos, C. Wang, F. Yang, Z. Zhong, B. Fu, M. A. Eckersley-Maslin, M. Woods, Y. Tanaka, X. Chen, A. C. Wilkinson, J. Bussell, J. White, R. Ramirez-Solis, W. Reik, B. Göttgens, S. A. Teichmann, P. P. L. Tam, H. Nakauchi, X. Zou, L. Lu, P. Liu, Establishment of mouse expanded potential stem cells. *Nature* **550**, 393–397 (2017).
52. X. Gao, M. Nowak-Imialek, X. Chen, D. Chen, D. Herrmann, D. Ruan, A. C. H. Chen, M. A. Eckersley-Maslin, S. Ahmad, Y. L. Lee, T. Kobayashi, D. Ryan, J. Zhong, J. Zhu, J. Wu, G. Lan, S. Petkov, J. Yang, L. Antunes, L. S. Campos, B. Fu, S. Wang, Y. Yong, X. Wang, S. G. Xue, L. Ge, Z. Liu, Y. Huang, T. Nie, P. Li, D. Wu, D. Pei, Y. Zhang, L. Lu, F. Yang, S. J. Kimber, W. Reik, X. Zou, Z. Shang, L. Lai, A. Surani, P. P. L. Tam, A. Ahmed, W. S. B. Yeung, S. A. Teichmann, H. Niemann, P. Liu, Establishment of porcine and human expanded potential stem cells. *Nat. Cell Biol.* **21**, 687–699 (2019).
53. T. S. Macfarlan, W. D. Gifford, S. Driscoll, K. Lettieri, H. M. Rowe, D. Bonanomi, A. Firth, O. Singer, D. Trono, S. L. Pfaff, Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**, 57–63 (2012).
54. Masaki Yagi, Satoshi Kishigami, Akito Tanaka, Katsunori Semi, Eiji Mizutani, Sayaka Wakayama, Teruhiko Wakayama, Takuya Yamamoto, Yasuhiro Yamada, Derivation of ground-state female ES cells maintaining gamete-derived DNA methylation. *Nature* **548**, 224–227 (2017). 10.1038/nature23286.
55. J. Cadinanos, A. Bradley, Generation of an inducible and optimized piggyBac transposon system. *Nucleic Acids Res.* **35**, e87 (2007).
56. A. G. Uren, H. Mikkers, J. Kool, L. van der Weyden, A. H. Lund, C. H. Wilson, R. Rance, J. Jonkers, M. van Lohuizen, A. Berns, D. J. Adams, A high-throughput splinkerette-PCR method for the isolation and sequencing of retroviral insertion sites. *Nat. Protoc.* **4**, 789–798 (2009).
57. W. Zhang, Y. Tian, Q. Gao, X. Li, Y. Li, J. Zhang, C. Yao, Y. Wang, H. Wang, Y. Zhao, Q. Zhang, L. Li, Y. Yu, Y. Fan, L. Shuai, Inhibition of apoptosis reduces diploidization of haploid mouse embryonic stem cells during differentiation. *Stem Cell Reports* **15**, 185–197 (2020).
58. Q. L. Ying, J. Wray, J. Nichols, L. Batlle-Morera, B. Doble, J. Woodgett, P. Cohen, A. Smith, The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519–523 (2008).
59. A. Yanagida, D. Spindlow, J. Nichols, A. Dattani, A. Smith, G. Guo, Naive stem cell blastocyst model captures human embryo lineage segregation. *Cell Stem Cell* **28**, 1016–1022. e4 (2021).
60. K. Peng, X. Li, C. Y. Wu, Y. Wang, J. Yu, J. Zhang, Q. Gao, W. Zhang, Q. Zhang, Y. Fan, Y. Yu, L. Shuai, Derivation of haploid trophoblast stem cells via conversion in vitro. *iScience* **11**, 508–518 (2019).

Acknowledgments: We thank all members of the Group of Stem Cells and Genetics for discussion and help. **Funding:** This work was supported by National Key Research and Development Program of China, grant no. 2019YFA0109901 to L.S.; National Natural Science Foundation of China, grant nos. 32022025 and 82371671 to L.S., 82201840 to Q.G., 82301879 to W.Z., and 31970579 to X.W.; The Fundamental Research Funds for the Central Universities, grant no. 63223030 to L.S.; The Natural Science Foundation of Tianjin City, grant nos. 22JCZDJC00480 to Y.L. and 20JCYBJC01400 to Y.C; and Strategic Collaborative Research Program of the Ferring Institute of Reproductive Medicine, grant no. FIRMD181102 to L.S. This work was also funded by National Clinical Research Center for Obstetrics and Gynecology (Peking University Third Hospital), no. BYSYSZKF2021007 to L.S. **Author contributions:** Conceptualization: W.Z. and L.S. Methodology: W.Z., S.S., Q.W., X.L., M.X., Q.L., Y.Z., K.P., C.Y., Y.W., Y.C., and Y.L. Investigation: L.S. Visualization: W.Z. and L.S. Supervision: X.W., Q.G., and L.S. Writing—original draft: W.Z., Q.G., and L.S. Writing—review and editing: L.S. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The RNA-seq, WGBS, and screening raw dataset in this study have been deposited in the Genome Sequence Archive of the Beijing Institute of Genomics (BIG) Data Center with accession numbers CRA011851, https://ngdc.cncb.ac.cn/search/?dbId=&q=CRA011851.

Submitted 4 May 2023
Accepted 17 November 2023
Published 20 December 2023
10.1126/sciadv.adi5683