nature cell biology

ARTICLES

https://doi.org/10.1038/s41556-019-0333-2

Establishment of porcine and human expanded potential stem cells

Xuefei Gao[1,2,23], Monika Nowak-Imialek[3,4,23], Xi Chen[2,23], Dongsheng Chen[5,6], Doris Herrmann[3,4], Degong Ruan[1,7], Andy Chun Hang Chen[8], Melanie A. Eckersley-Maslin[9], Shakil Ahmad[10], Yin Lau Lee[8], Toshihiro Kobayashi[11], David Ryan[2], Jixing Zhong[5,6], Jiacheng Zhu[5,6], Jian Wu[1], Guocheng Lan[12], Stoyan Petkov[3,4,20], Jian Yang[2,21], Liliana Antunes[2], Lia S. Campos[2], Beiyuan Fu[2], Shengpeng Wang[5,6], Yu Yong[2], Xiaomin Wang[7], Song-Guo Xue[13], Liangpeng Ge[14], Zuohua Liu[14], Yong Huang[14], Tao Nie[7], Peng Li[7], Donghai Wu[7], Duanqing Pei[7,15], Yi Zhang[16], Liming Lu[17], Fengtang Yang[2], Susan J. Kimber[18], Wolf Reik[9], Xiangang Zou[12], Zhouchun Shang[5,6], Liangxue Lai[7], Azim Surani[11], Patrick P. L. Tam[19], Asif Ahmed[10], William Shu Biu Yeung[8], Sarah A. Teichmann[2], Heiner Niemann[3,4,22]* and Pentao Liu[1,2]*

We recently derived mouse expanded potential stem cells (EPSCs) from individual blastomeres by inhibiting the critical molecular pathways that predispose their differentiation. EPSCs had enriched molecular signatures of blastomeres and possessed developmental potency for all embryonic and extra-embryonic cell lineages. Here, we report the derivation of porcine EPSCs, which express key pluripotency genes, are genetically stable, permit genome editing, differentiate to derivatives of the three germ layers in chimeras and produce primordial germ cell-like cells in vitro. Under similar conditions, human embryonic stem cells and induced pluripotent stem cells can be converted, or somatic cells directly reprogrammed, to EPSCs that display the molecular and functional attributes reminiscent of porcine EPSCs. Importantly, trophoblast stem-cell-like cells can be generated from both human and porcine EPSCs. Our pathway-inhibition paradigm thus opens an avenue for generating mammalian pluripotent stem cells, and EPSCs present a unique cellular platform for translational research in biotechnology and regenerative medicine.

Mouse and human embryonic stem cells (ESCs) derived from pre-implantation embryos[1–3] self-renew in long-term cultures and differentiate to all embryonic cell lineages in vitro and in mouse chimeras. The development of well-defined culture conditions, such as 2i/LIF, has substantially facilitated the derivation and maintenance of mouse ESCs[4], and has led to intensive efforts for deriving human ESCs akin to mouse ESCs[5,6]. However, it has been challenging to translate the findings from studies of mouse and human cells to establish ESCs from other mammalian species. The domestic pig shares great genetic, anatomical and physiological similarities with humans, and is considered to be an excellent model for human diseases, cell therapies and even as a donor for porcine xenografts. To this date, bona fide porcine ESCs have not yet been established[7–14]. The published lines usually do not meet with the stringent criteria for pluripotency and are frequently called 'ES-like' cells.

We have recently demonstrated that by targeting key molecular pathways that drive lineage differentiation in the mouse pre-implantation embryo, expanded potential stem cells (mEPSCs) displaying a broad propensity for extra-embryonic and embryonic

[1]School of Biomedical Sciences, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Stem Cell and Regenerative Medicine Consortium, Pokfulam, Hong Kong. [2]The Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge, UK. [3]Institute of Farm Animal Genetics, Friedrich-Loeffler-Institut (FLI), Mariensee, Neustadt, Germany. [4]REBIRTH Centre of Excellence, Hannover Medical School, Hannover, Germany. [5]BGI-Shenzhen, Shenzhen, China. [6]China National GeneBank, BGI-Shenzhen, Shenzhen, China. [7]Key Laboratory of Regenerative Biology of Chinese Academy of Sciences, Guangdong Provincial Key Laboratory of Stem Cell and Regenerative Medicine, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou, China. [8]Department of Obstetrics and Gynaecology, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Pokfulam, Hong Kong. [9]Epigenetics Programme, Babraham Institute, Babraham Research Campus, Cambridge, UK. [10]Aston Medical Research Institute, Aston Medical School, Aston University, Birmingham, UK. [11]Wellcome Trust and Cancer Research UK Gurdon Institute, University of Cambridge, Cambridge, UK. [12]Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, UK. [13]Center for Reproductive Medicine, Shanghai East Hospital, School of Medicine, Tong Ji University, Shanghai, China. [14]Chongqing Academy of Animal Sciences and Key Laboratory of Pig Industry Sciences, Department of Agriculture, Chongqing, China. [15]Guangzhou Regenerative Medicine and Health Guangdong Laboratory, Guangzhou, China. [16]Biotherapy Center, The First Affiliated Hospital of Zhengzhou University, Henan, China. [17]Institute of Immunology, School of Medicine, Shanghai Jiaotong University, Shanghai, China. [18]Faculty of Biology Medicine and Health, University of Manchester, Manchester, UK. [19]Embryology Unit, Children's Medical Research Institute and School of Medical Sciences, Sydney Medical School, Faculty of Medicine and Health, The University of Sydney, Westmead, NSW, Australia. [20]Present address: German Primate Center, Platform Degenerative Diseases, Gottingen, Germany. [21]Present address: Key Laboratory of Arrhythmias, Ministry of Education, Shanghai East Hospital, Tongji University School of Medicine, Shanghai, China. [22]Present address: Hannover Medical School (MHH), TwinCore, Hannover, Germany. [23]These authors contributed equally: Xuefei Gao, Monika Nowak-Imialek, Xi Chen. *e-mail: niemann.heiner@mh-hannover.de; pliu88@hku.hk

lineage differentiation were derived[15,16]. We hypothesized that a similar experimental paradigm of targeting key developmental pathways might be applied to establish porcine stem cells from pre-implantation embryos. However, little is known about the molecular and signalling mechanisms of porcine early pre-implantation embryo development; we thus set out to perform a chemical screen of inhibitors that were used to isolate and maintain mouse mEPSCs, mouse and human ESCs and to delineate the optimal conditions for porcine cells. Our results demonstrate that porcine EPSCs could be established and, importantly, that similar culture conditions permit derivation of human EPSCs.

Results and discussion

Identification of culture conditions for porcine pluripotent stem cells. Although porcine induced pluripotent stem cells (iPSCs) are available, their use for screening is confounded by the leaky expression of the transgenic reprogramming factors after reprogramming or by low levels of expression of the endogenous pluripotency genes[17–20]. To overcome this challenge, we generated porcine iPSCs by expressing eight doxycycline (Dox)-inducible transcription factors, which substantially improved the efficiency of reprogramming both wild-type and transgenic porcine fetal fibroblasts (PFFs), in which a *tdTomato* cassette had been inserted into the 3′ untranslated region of the porcine *OCT4* (*POU5F1*) locus (POT PFFs)[21] of putative iPSC colonies (Fig. 1a–c). The iPSCs from POT PFFs were OCT4–tdTomato$^+$ (Fig. 1c) and expressed high levels of the endogenous pluripotency factors (Fig. 1d). The iPSCs could be passaged as single cells for more than 20 passages in serum-containing medium (M15) plus Dox. Following Dox removal, the iPSCs differentiated within 4–5 d, concomitant with rapid downregulation of the exogenous reprogramming factors and endogenous pluripotency genes, and increased expression of both embryonic and extra-embryonic cell-lineage genes (Fig. 1e–h). These Dox-dependent iPSCs with robust endogenous pluripotency gene expression provided the material for the chemical screen.

Over 400 combinations of 20 small molecule inhibitors and cytokines were tested for their ability to maintain Dox-independent porcine iPSCs in the undifferentiated state (Fig. 1i and Supplementary Table 1). A departure was noted from previous reports that naive mouse ESC medium 2i/LIF (ref. [4]) was able to maintain putative porcine iPSCs[22–24]: porcine iPSCs were rapidly lost with 1.0 μM Mek1 inhibitor PD0325901, irrespective of whether or not Dox was present (Supplementary Fig. 1a–g), indicating that porcine pluripotent stem cells differ from mouse ESCs in the requirements of Mek–ERK signalling[4,25]. Inhibition of p38 and PKC was also non-conducive for porcine iPSCs (Supplementary Fig. 1f,h). Therefore, mouse or human naive ESC conditions[4–6] cannot be directly extrapolated to porcine cells. The Mek1/2, p38 and PKC inhibitors were therefore excluded from the screen. Several conditions were identified that met the screen criteria (Supplementary Fig. 1g), including a minimal requisite condition (no. 517, porcine EPSC (pEPSC) medium (pEPSCM)) comprising inhibitors for GSK3 (CHIR99021), SRC (WH-4-023) and Tankyrases (XAV939; the last two were inhibitors important for mouse EPSCs[15]), and supplements: vitamin C (Vc), Activin A and LIF (Fig. 1i, Supplementary Fig. 1g and Supplementary Table 1). Under these conditions, the Dox-independent iPSCs (pEPSC^{iPS}) remained undifferentiated for 30 passages, expressed endogenous pluripotency factors at levels comparable to the porcine blastocyst and showed no leaky expression of the exogenous reprogramming factors (Fig. 1j and Supplementary Fig. 1i,j).

We next repeated the reprogramming experiment by directly culturing the primary colonies in pEPSCM (Supplementary Fig. 2a) and generated 11 stable pEPSC^{iPS} lines from 16 primary colonies (70% efficiency), 6 of these had no detectable expression of any of the 8 exogenous reprogramming factors but had high levels of endogenous pluripotency genes (Supplementary Fig. 2b).

Establishment of porcine EPSCs from pre-implantation embryos. The pEPSCM condition was subsequently employed to derive stem cell lines from porcine pre-implantation embryos. A total of 26 lines (pEPSCs^{Emb}; 14 male and 12 female) were established from 76 early blastocysts (5.0 days post coitum (dpc)) and 12 cell lines (pEPSCs^{par}) from 252 parthenogenetic blastocysts (Fig. 2a, Supplementary Table 2 and Supplementary Fig. 2c). Like pEPSCs^{iPS}, pEPSCs^{Emb} had high nuclear:cytoplasmic ratios and formed compact colonies with smooth colony edges (Fig. 2a and Supplementary Fig. 2d). The pEPSCs^{Emb} were passaged every 3–4 d at a ratio of 1:8 as single cells, could be maintained for >40 passages on STO feeders without overt differentiation and were genetically stable (Supplementary Fig. 2e). The sub-cloning efficiency was about 10% at a low cell density (2,000 cells well^{-1} in a six-well plate) but routine passaging was performed at a high cell density.

Pluripotency genes were expressed in pEPSCs^{Emb} and pEPSCs^{iPS} at levels comparable to the blastocysts (Fig. 2b and Supplementary Fig. 2b) but were drastically reduced or lost when the pEPSCs were cultured in other previously reported porcine ESC media[8–14] (Supplementary Fig. 2f,g). Porcine EPSCs showed extensive DNA demethylation at the *OCT4* and *NANOG* promoter regions (Fig. 2c) and had *OCT4* distal enhancer activity (Supplementary Fig. 2h). The pEPSCs were amenable to CRISPR/Cas9-mediated insertion of an *H2B–mCherry* expression cassette into the *ROSA26* locus (Supplementary Fig. 2i,j). In vitro, pEPSCs differentiated into tissues expressing genes representative of the three germ layers and, uniquely, trophoblast genes (Fig. 2d and Supplementary Fig. 2k). In immunocompromised mice, pEPSCs^{Emb} formed mature teratomas with derivatives of the three germ layers and contained placental lactogen-1 (PL-1)-, KRT7- and SDC1-positive trophoblast-like cells (Fig. 2e,f). Following incorporation of the pEPSCs into pre-implantation embryos and 48 h of culture, pEPSCs (marked by H2B–mCherry) had colonized both the trophectoderm and the inner cell mass of blastocysts (Supplementary Fig. 3a). A total of 45 conceptuses were harvested from 3 litters at days 26–28 of gestation following the transfer of the chimeric blastocysts to synchronized recipient sows (Supplementary Table 3 and Supplementary Fig. 3b). Flow cytometry analysis of dissociated cells from embryonic and extra-embryonic tissues of the chimeras detected mCherry$^+$ cells in seven conceptuses (Supplementary Fig. 3c and Supplementary Tables 4,5): mCherry$^+$ cells in both the placenta and embryonic tissues in two chimeras (nos 8 and 16); only in embryonic tissues in three chimeras (nos 4, 21 and 34) and exclusively in the placenta of two chimeras (nos 3 and 6). Genomic DNA PCR assays detected mCherry DNA only in those seven mCherry$^+$ chimeras and not in any other conceptuses (Supplementary Fig. 3d and Supplementary Tables 4,5). Despite the low contribution of the donor mCherry$^+$ cells, their descendants were found in multiple embryonic tissues and organs that were identified by tissue lineage markers (Fig. 2g and Supplementary Fig. 3e,f). Therefore, like mEPSCs, pEPSCs^{Emb} and pEPSCs^{iPS} possess an expanded developmental potential for both the embryonic cell lineages and extra-embryonic trophoblast lineages.

Derivation of PGCLCs from pEPSCs^{Emb}. We next investigated whether pEPSCs had the potential to produce PGC-like cells (PGCLCs) in vitro, similarly to mouse and human pluripotent stem cells[26–28]. In porcine embryos (E11.5–E12) in the early primitive streak stage, the first cluster of porcine PGCs can be detected as SOX17$^+$ cells at the posterior end of the nascent primitive streak and these cells later coexpress OCT4, NANOG, BLIMP1 and TFAP2C[28]. *NANOS3* is an evolutionarily conserved PGC-specific factor[29,30] and human *NANOS3* reporter ESCs have been used to study the derivation of PGCLCs[27,28]. We generated and used a *NANOS3–H2B–mCherry* pEPSC^{Emb} reporter line to facilitate the identification of putative PGCLCs (Supplementary Fig. 4a). After

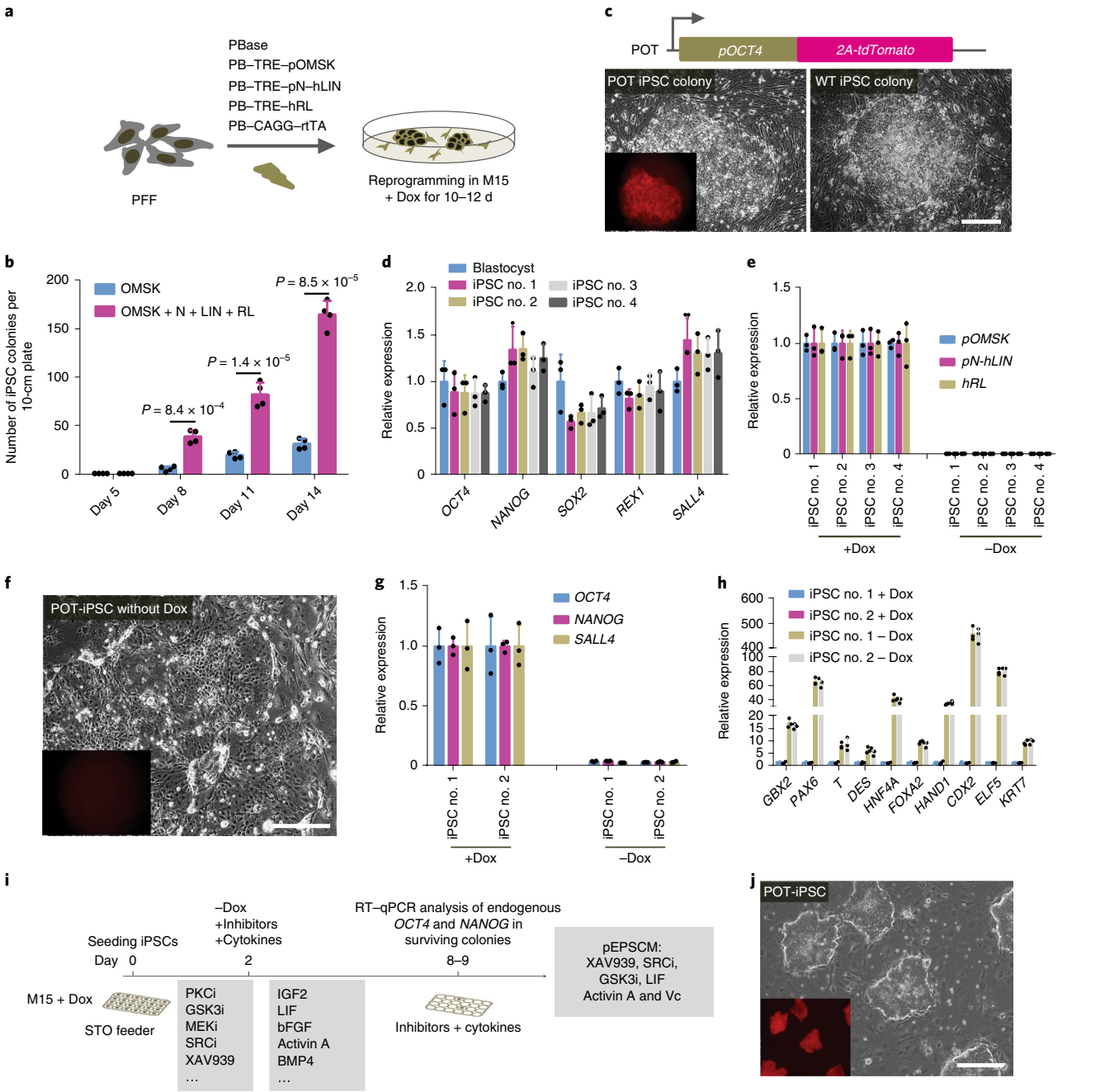

Fig. 1 | Identification of culture conditions for porcine EPSCs. a, Doxycycline-inducible expression of Yamanaka factors *OCT4*, *MYC*, *SOX2* and *KLF4* together with *LIN28*, *NANOG*, *LRH1* and *RARG* in porcine PFFs. Stable genomic integration of complementary DNA in PFFs was achieved by piggyBac (PB) transposition. pOMSK, porcine *OCT4*, *MYC*, *SOX2* and *KLF4*; pN-hLIN, porcine *NANOG* and human *LIN28*; hRL, human *RARG* and *LRH1*. Reverse tetracycline-controlled transactivator (rtTA) was driven by CMV early enhancer/chicken beta actin promoter (CAG). The reprogrammed colonies were single-cell passaged in the presence of Dox in M15 (15% fetal bovine serum). **b**, Coexpression of *LIN28*, *NANOG*, *LRH1* and *RARG* substantially increased the number of reprogrammed colonies from 250,000 PFFs ($n = 4$ independent experiments). **c**, Reprogramming of the porcine *OCT4–tdTomato* knock-in reporter (POT) TAIHU and wide-type (WT) German Landrace PFFs to iPSCs. The inset image shows the loss of tdTomato expression. **d**, The iPSC lines expressed key pluripotency genes, as analysed by quantitative PCR with reverse transcription (RT–qPCR). The iPSC lines nos 1 and 2, and nos 3 and 4 were from WT German Landrace and POT PFFs, respectively. **e**, RT–qPCR analysis of the exogenous reprogramming factors in iPSCs either in the presence of Dox or 5 d after its removal. **f**, POT iPSCs became tdTomato negative 5 d after Dox removal. **g**, RT–qPCR analysis of the expression of endogenous pluripotency genes in iPSCs cultured with or without Dox. The inset image shows the loss of tdTomato expression. **h**, Expression of lineage genes in porcine iPSCs 5–6 d after Dox removal. Gene expression was measured by RT–qPCR. The relative expression levels are shown normalized to *GAPDH*. The experiments were performed three times. **i**, Diagram depicting the screening strategy used to identify the culture conditions for porcine pluripotent stem cells using Dox-dependent iPSCs. Small-molecule inhibitors and cytokines were selected for various combinations. Cell survival, cell morphology and the expression of endogenous *OCT4* and *NANOG* were employed as the read-outs. **j**, Images of POT reporter iPSCs in pEPSCM without Dox. The inset image shows the expression of tdTomato. **d**,**e**,**g**,**h**, $n = 3$ independent experiments. All graphs represent the mean ± s.d. The *P* values were computed using a two-tailed Student's *t*-test. The experiments in **c**,**f**,**j** were repeated independently three times with similar results. The source data are provided in Supplementary Table 1. Scale bars, 100 μm.

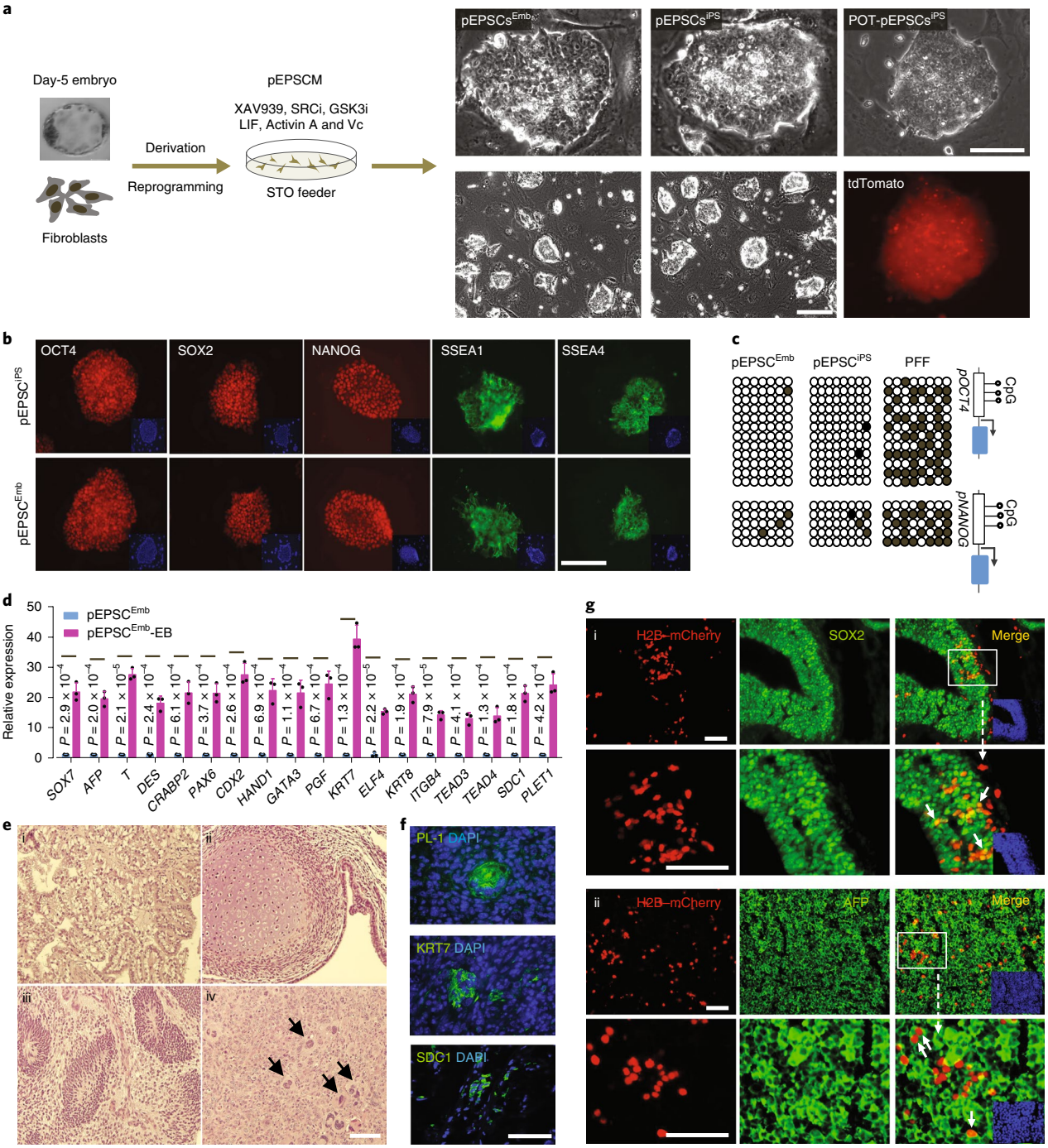

Fig. 2 | Derivation of porcine EPSCs. a, Schematic diagram of the establishment of the pig (*Sus scrofa*) EPSC^{Emb} lines from German Landrace day-5 in vivo-derived blastocysts on STO feeder cells in pEPSCM, and of pEPSC^{iPS} lines by reprogramming German Landrace PFFs and China TAIHU POT knock-in reporter PFFs (left). Images of the established EPSC lines (right) and a fluorescence image of tdTomato expression in POT-pEPSCs^{iPS} (bottom right). Three EPSC^{Emb} lines (male, K3 and K5, and female, K1) and two pEPSC^{iPS} lines (nos 10 and 11) were extensively tested in this study. Gene expression and differentiation were similar for these EPSC lines. **b**, Detection of pluripotency factors and the markers SSEA-1 and SSEA-4 in pEPSCs^{Emb} and pEPSCs^{iPS} through immunostaining. **c**, Bisulfite sequencing analysis of the CpG sites in the *OCT4* and *NANOG* promoter regions in PFFs, pEPSCs^{iPS} and pEPSCs^{Emb}. **d**, Gene expression in EBs (day 7) of pEPSCs^{Emb}. The expression levels of genes of embryonic and extra-embryonic cell lineages were assessed by RT–qPCR. The relative expression levels were normalized to *GAPDH*. Data represent the mean ± s.d.; $n = 3$ independent experiments. The *P* values were calculated using a two-tailed Student's *t*-test. The statistical source data are provided in Supplementary Table 10. **e**, Tissue composition of pEPSC^{Emb} teratoma sections (haematoxylin and eosin staining) with examples of glandular epithelium derived from endoderm (**i**), cartilage derived from mesoderm (**ii**), immature neural tissue derived from ectoderm, which forms neuroepithelial structures (**iii**) and large multinucleated cells reminiscent of trophoblasts (arrows in **iv**). **f**, PL-1-, KRT7- and SDC1-positive cells in pEPSC^{Emb} teratoma sections as revealed by immunostaining. **g**, Detection of pEPSC descendants in the brain (H2B–mCherry^{+}SOX2^{+}; **i**) and liver (H2B–mCherry^{+}AFP^{+}; **ii**) in chimera no. 16. H2B–mCherry and SOX2 are localized to the nucleus, whereas AFP is a cytoplasmic protein. The boxed areas are shown below at higher magnifications. The arrows indicate representative cells that were donor-cell descendants (mCherry^{+}). The nuclei are DAPI stained. Additional chimera analyses are presented in Supplementary Fig. 3e,f. The experiments in **a**,**b**,**e**–**g** were repeated independently three times with similar results. Scale bars, 100 μm.

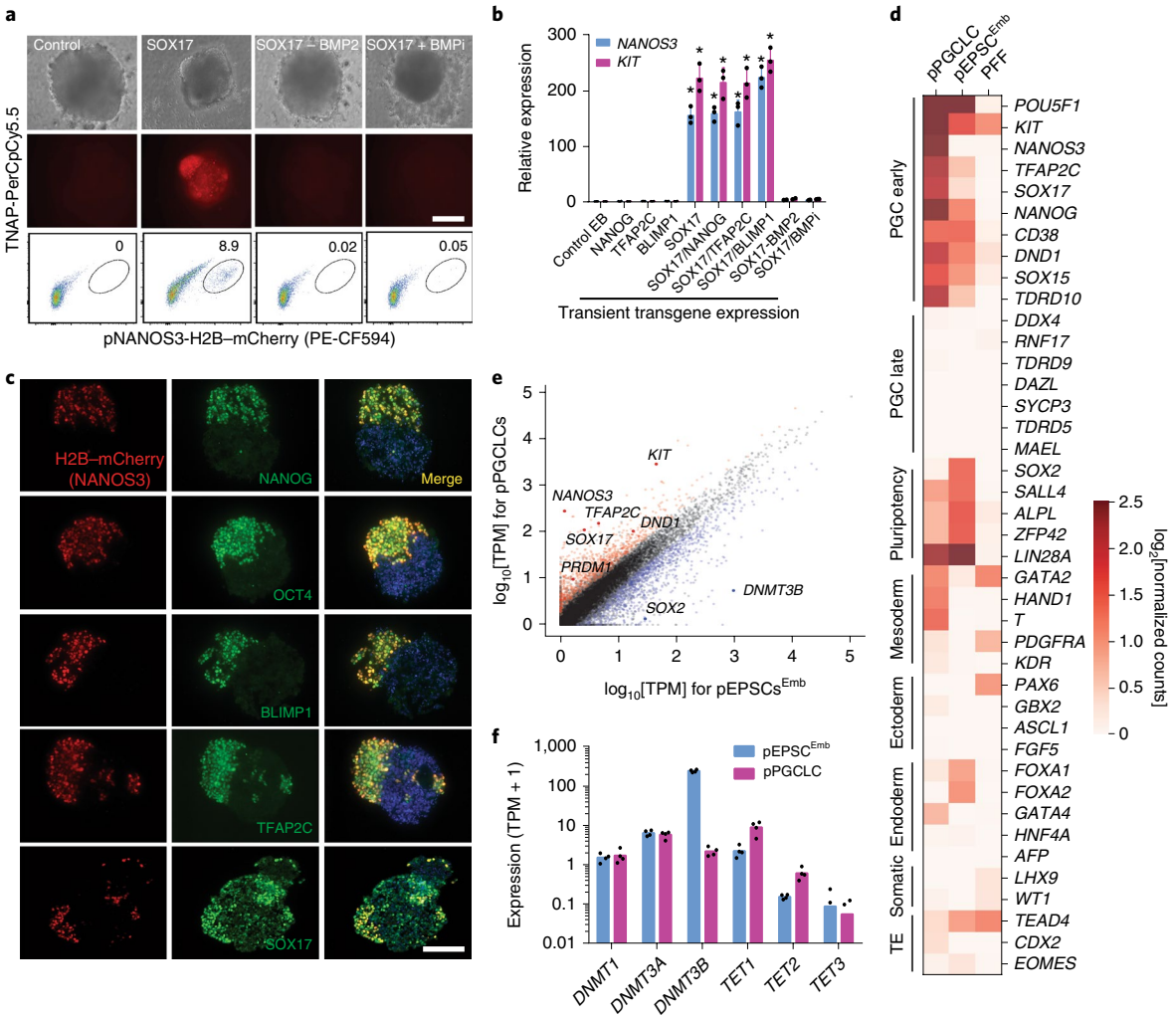

Fig. 3 | In vitro generation of PGCLCs from pEPSCs^{Emb}. **a**, Induction of porcine PGCLCs (pPGCLCs) by transiently expressing *SOX17* in *NANOS3–H2B–mCherry* reporter pEPSCs. The presence of H2B–mCherry$^+$TNAP$^+$ cells in EBs was analysed by fluorescence-activated cell sorting (FACS). The experiments were repeated independently three times with similar results. **b**, RT–qPCR analysis of PGC genes in day-3 EBs following pPGCLC induction. The relative expression levels were normalized to *GAPDH*. Data represent the mean ± s.d.; $n = 3$ independent experiments. The *P* values were calculated using a two-tailed Student's *t*-test. The statistical source data are provided in Supplementary Table 10. **c**, Immunofluorescence analysis of PGC factors in section EBs at days 3–4 following pPGCLC induction. The H2B–mCherry$^+$ cells coexpressed NANOG, OCT4, BLIMP1, TFAP2C and SOX17. The nuclei were DAPI stained. The experiments were performed three times. **d**, RNA-Seq analysis (heat map) of sorted H2B–mCherry$^+$ cells following pPGCLC induction shows the expression of genes associated with PGCs, pluripotency or somatic lineages (mesoderm, endoderm and gonadal somatic cells). **e**, Pairwise gene expression comparison between pEPSCs^{Emb} and pPGCLCs. The key upregulated (red) and downregulated (blue) genes are highlighted. TPM, transcripts per million reads. **f**, Expression of genes related to DNA methylation in pPGCLCs and the parental pEPSCs^{Emb}. The data are from RNA-Seq of sorted H2B–mCherry$^+$ cells following pPGCLC induction. $n = 4$ independent experiments. Scale bars, 100 μm.

transiently expressing the *SOX17* transgene for 12 h, the reporter cells were allowed to form embryoid bodies (EBs; Supplementary Fig. 4b), where cell clusters coexpressing *NANOS3* (mCherry$^+$) and tissue non-specific alkaline phosphatase (TNAP; a PGC marker) were detected within 3–4 d (Fig. 3a).

The derivation of putative porcine PGCLCs was BMP2/4 dependent (Fig. 3a). Interestingly, in contrast to the reported derivation of human PGCLCs[28], the expression of *NANOG*, *BLIMP1* or *TFAP2C* transgenes, either individually or in combination, had no effect on the preponderance of *NANOS3*$^+$ cells (Supplementary Fig. 4c), whereas coexpression of *SOX17* with *BLIMP1* seemed to increase *NANOS3*$^+$ cells (Supplementary Fig. 4c,d).

The putative PGCLCs in the EBs expressed PGC-specific genes (Fig. 3b,c and Supplementary Fig. 4e). Specific RNA-Seq analysis of *NANOS3*$^+$ cells revealed expression of early PGC genes[27] (*OCT4*, *NANOG*, *LIN28A*, *TFAP2C*, *CD38*, *DND1*, *NANOS3*, *ITGB3*, *SOX15* and *KIT*) and reduced *SOX2* expression (Fig. 3d,e). As in PGCLC derivation from human ESCs[27], *DNMT3B* was downregulated in porcine mCherry$^+$/*NANOS3*$^+$ cells, whereas *TET1* and *TET2* were upregulated, relative to the parental pEPSCs^{Emb} (Fig. 3e,f).

Establishment of human EPSCs under conditions similar to those of porcine EPSCs. The findings that inhibition of SRC and Tankyrases is sufficient to convert mouse ESCs to mEPSCs[15], and that the same two inhibitors are required for the derivation of pEPSCs raise the possibility that similar in vitro culture conditions may be developed for additional mammalian species. To explore this possibility, we cultured four established human embryonic stem cell (hESC) lines (H1, H9, Man1/M1 and Man10/M10 cells)[3,31,32] in pEPSCM and passaged them three times. The cells displayed

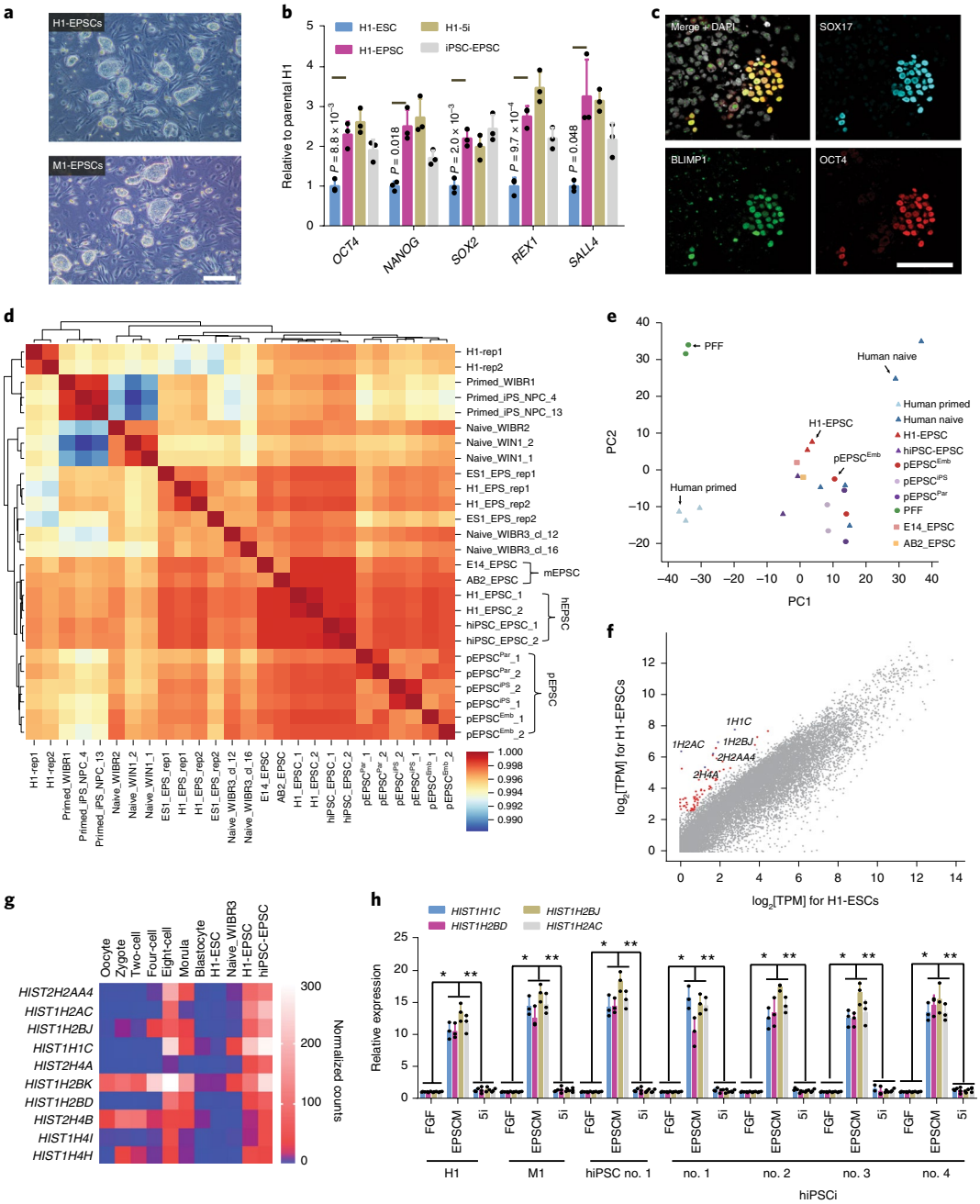

Fig. 4 | Establishment of human EPSCs. a, Images of the established H1-EPSCs (top) and M1-EPSCs (bottom) at passage 25. The experiments were repeated independently three times with similar results. **b**, Expression of pluripotency genes in H1-ESCs, H1-naive ESCs (H1-5i), H1-EPSCs and iPSC-EPSCs. Data represent the mean ± s.d.; $n = 3$ independent experiments. The P values were computed using a two-tailed Student's t-test. **c**, EBs of H1-EPSCs to PGCLCs immunostained for SOX17, BLIMP1 and OCT4. **d**, Hierarchical clustering of gene expression (bulk RNA-Seq) of EPSCs and other human pluripotent stem cells. The correlation matrix was clustered using a Spearman's correlation and complete linkage. The following datasets were used: pEPSC^{Par} (porcine parthenogenetic EPSCs); E14- and AB2-EPSCs (mouse EPSCs)[15]; human primed ESCs (WIBR1, iPS_NPC_4 and iPS_NPC_13) and naive ESCs (WIBR2, WIBR3_cl_12, WIBR3_cl_16, WIN1_1 and WIN1_2)[6,35]; human primed H1 ES cells (H1-rep1 and H1-rep2) and extended pluripotent stem (EPS) cells (H1_EPS_rep1, H1_EPS_rep2, ES1_EPS_rep1 and ES1_EPS_rep2)[36]. **e**, Principal component analysis (PCA) of bulk RNA-Seq data of EPSCs, human primed and naive ESCs, and PFFs. Human naive, $n = 5$; human primed, $n = 3$; H1-EPSC, $n = 2$; hiPSC-EPSC, $n = 2$; pEPSC^{Emb}, $n = 2$; pEPSC^{iPS}, $n = 2$; pEPSC^{Par}, $n = 2$; PFF, $n = 2$; E14_EPSC, $n = 2$ and AB2_EPSC, $n = 1$; n represents the number of biologically independent experiments. **f**, Pairwise comparison of gene expression levels between H1-ESCs and H1-EPSCs showing the highly expressed genes (more than eightfold) in hEPSCs (total 76; red dots) and representative histone genes (blue dots). **g**, Expression levels of selected histone genes in human ESCs, EPSCs and pre-implantation embryos. RNA-Seq data of human ESCs were from ref. [35], whereas embryo cell data were from ref. [37]. **h**, RT–qPCR analysis of four histone 1 cluster genes in seven human ESC or iPSC lines cultured under three different conditions. The hiPSC lines were from the HIPSC project (http://www.hipsci.org): no. 1, HPSI1113i-bima_1; no. 2, HPSI1113i-qolg_3; no. 3, HPSI1113i-oaaz_2 and no. 4, HPSI1113i-uofv_1. The relative expression levels are shown with normalization to *GAPDH*. Data represent the mean ± s.d.; $n = 3$ independent experiments. The P values were computed using a two-tailed Student's t-test. *$P < 0.01$ compared with the FGF condition cultured cells; **$P < 0.01$ compared with 5i-condition cultured cells. The experiments were performed three times. The statistical source data are provided in Supplementary Table 10. Scale bars, 100 μm.

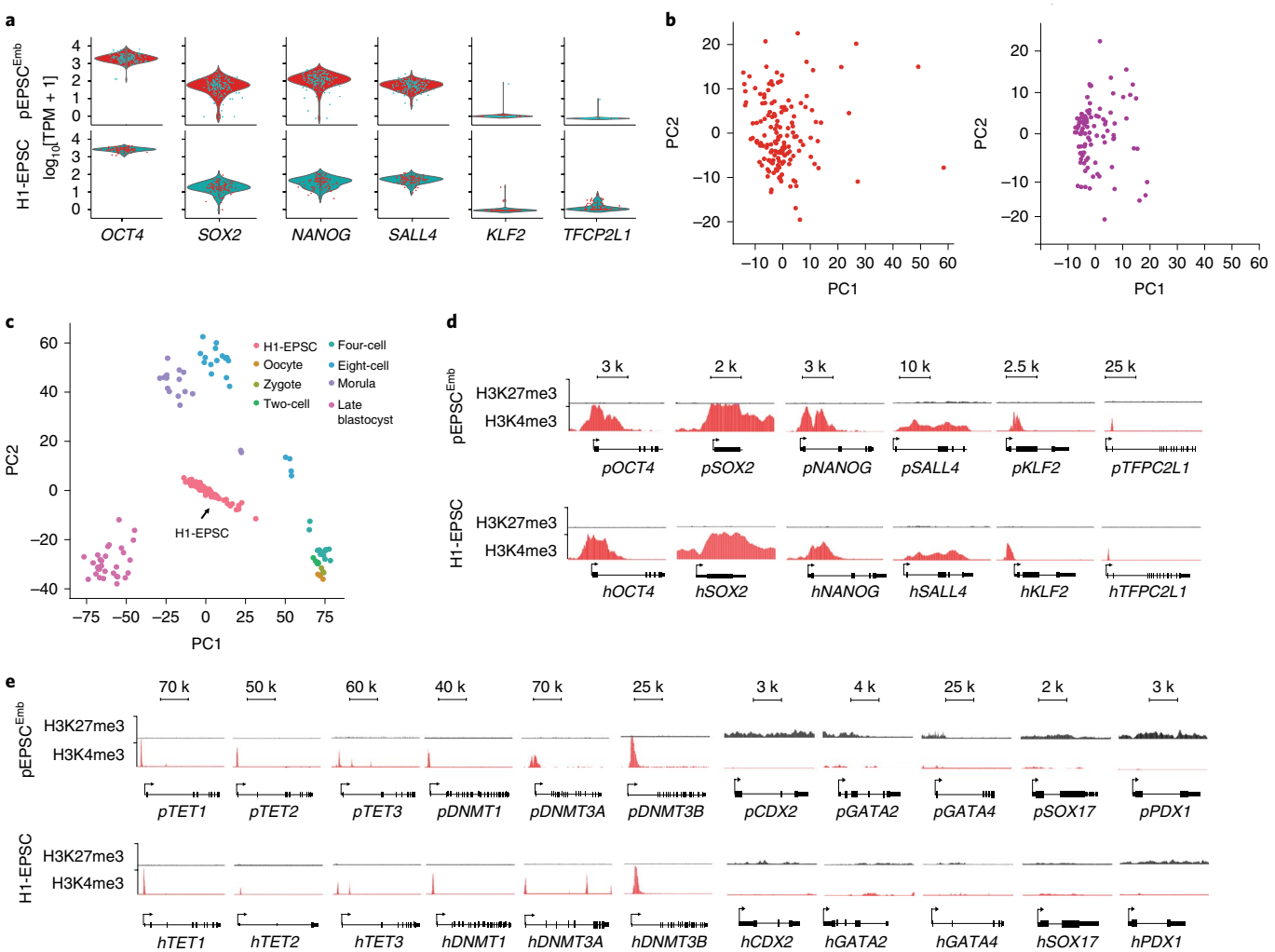

Fig. 5 | Molecular features of porcine and human EPSCs. a, Distribution and probability density of pluripotency gene expression (scRNA-Seq) in pEPSCs^{Emb} (top, $n=150$) and human H1-EPSCs (bottom, $n=96$). n represents the number of cells in each plot. **b**, PCA of the global gene expression pattern (by scRNA-Seq) of pEPSCs^{Emb} (left; $n=150$) and H1-EPSCs (right; $n=96$). n represents the number of cells in each plot. **c**, PCA and comparison of gene expression (by scRNA-Seq) of human H1-EPSCs and human pre-implantation embryos[37]. H1-EPSCs, $n=96$; oocytes, $n=3$; zygotes, $n=3$; two-cell blastomeres, $n=6$; four-cell blastomeres, $n=12$; eight-cell blastomeres, $n=20$; morulae, $n=16$ and late blastocysts, $n=30$; n represents the number of cells. **d**, Chromatin immunoprecipitation-sequencing (ChIP-seq) analysis of H3K27me3 and H3K4me3 marks at pluripotency gene loci in pEPSCs^{Emb} (top) and human H1-EPSCs (bottom). **e**, Histone modifications (H3K4me3 and H3K27me3) at the loci for genes encoding enzymes involved in DNA methylation and demethylation and for cell-lineage genes. The experiments in **d**,**e** were performed three times with similar results.

diverse morphologies and heterogeneous expression of OCT4 (Supplementary Fig. 5a). The removal of Activin A (20.0 ng ml^{-1}) from pEPSCM led to the formation of considerably fewer cell colonies from H1 (<1.0%) and M1 (5.0%) ESCs, whereas none were formed from H9 or M10 (Supplementary Fig. 5a), which reflects the inherent between-line heterogeneity of human ESCs[33,34]. With further refinement of the culture conditions (for example, replacing WH-4-023 with another SRC inhibitor A419259 in human EPSCM (hEPSCM); see Methods), morphologically homogenous and stable cell lines were established from single-cell sub-cloned H1 (H1-EPSCs) and M1 cells (M1-EPSCs; Fig. 4a). Karyotype analysis of H1 and M1 cells grown in hEPSCM on STO feeders revealed genetic stability (at passage 25 post conversion from the parental hESCs; Supplementary Fig. 5b). When human primary iPSC colonies reprogrammed from fibroblasts were directly cultured in hEPSCM, around 70% of the picked colonies could be established as stable iPSC lines (iPSC-EPSCs; Supplementary Fig. 5c), which expressed pluripotency markers with no obvious leakiness of the exogenous reprogramming factors in about half the lines (Fig. 4b and Supplementary Fig. 5d). The H1-EPSCs proliferated more robustly than the H1-ESCs cultured in standard FGF-containing medium (H1-ESC, primed) or under naive 5i/L/A conditions (H1-naive ESC[6]; Supplementary Fig. 5e) and were tolerant of single-cell passaging with about 10% single-cell sub-cloning efficiency in the transient presence of ROCKi. The cell survival at passaging was substantially improved in the presence of 5.0 ng ml^{-1} Activin A or by splitting the cells at higher densities. Human EPSCs expressed pluripotency genes at higher levels than the H1-ESCs (Fig. 4b) and minimal levels of lineage markers (Supplementary Fig. 5f). The expression of core pluripotency factors and surface markers in human EPSCs was confirmed by immunostaining (Supplementary Fig. 5g). H1-EPSCs differentiated to derivatives of the three germ layers in vitro and in vivo (Supplementary Fig. 5h,i). Moreover, H1-EPSCs were successfully differentiated to PGCLCs using in vitro conditions developed for germ-cell-competent hESCs or iPSCs[27,28] (Fig. 4c and Supplementary Fig. 5j).

Our results demonstrate that porcine and human EPSCs could be derived and maintained using a similar set of small-molecule

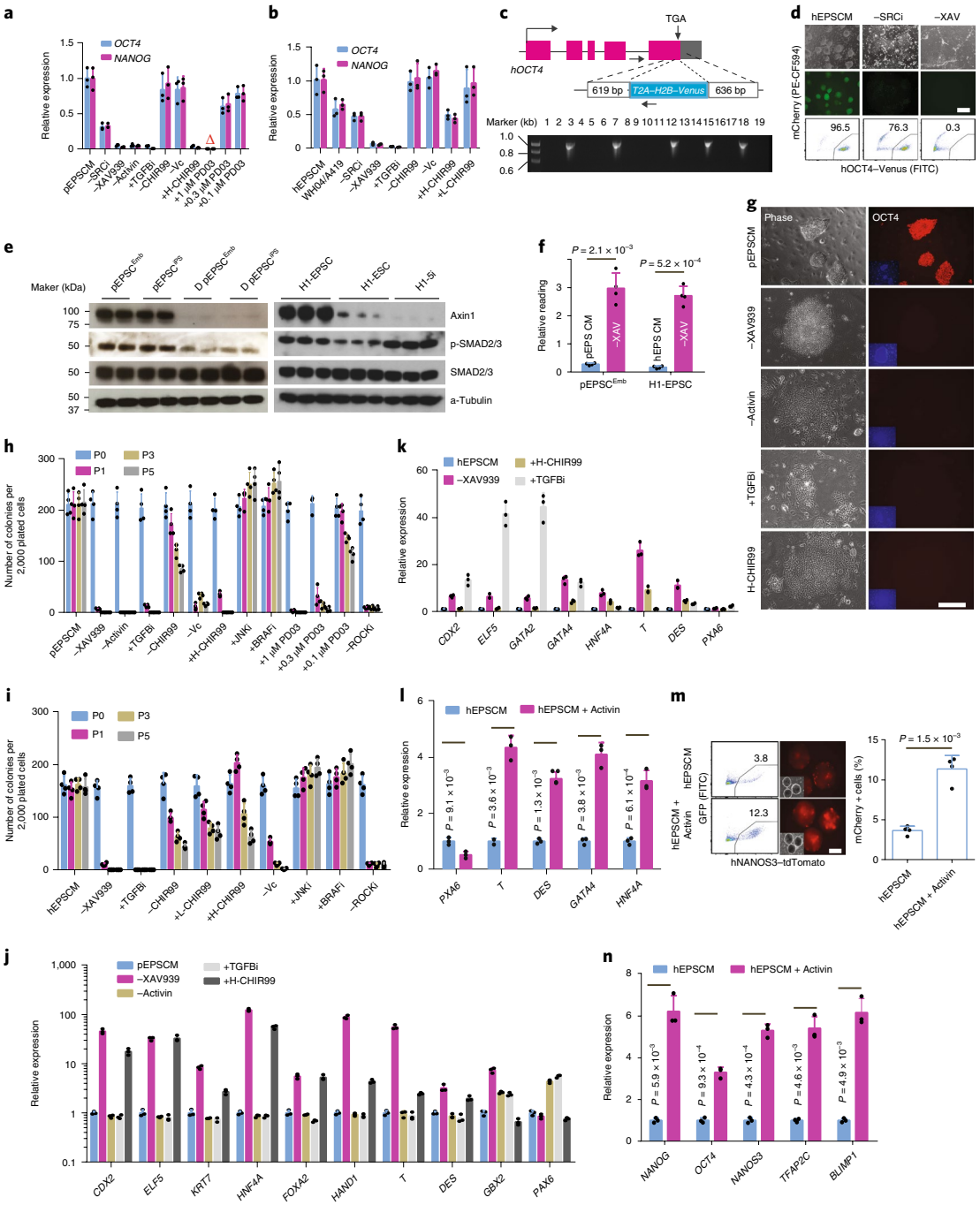

Fig. 6 | The requirement of individual components in EPSCM. a,**b**, Gene expression in pEPSCs^{Emb} (**a**) and H1-EPSCs (**b**) analysed by RT–qPCR. Components were removed individually (−SRCi, −XAV939, −Activin, −Vc and −CHIR99), added (+TGFBi (SB431542), +H-CHIR99 (3.0 μM CHIR99021), +PD03 (PD0325901) and +L-CHIR99 (0.2 μM in hEPSCM)) or replaced (WH04/A419, replacing A419259 with WH-4-23). Red triangle, no cells survived. **c**, Targeting the *H2B–Venus* cassette to the *OCT4* last coding exon in H1-EPSCs. **d**, The effects of removing WH-4-023 (−SRCi) or XAV939 (−XAV) for 7 d were assessed by fluorescence microscopy and flow cytometry of the Venus$^+$ reporter. **e**, Western blot analysis of Axin1 and phosphorylation of SMAD2/3 in EPSCs. EPSCs had higher levels of Axin1 and p-SMAD2/3 (for TGFβ signalling) than the differentiated (D) EPSC^{Emb} or primed H1-ESCs. **f**, TOPflash analysis of canonical Wnt signalling activity in EPSCs. Removing XAV939 (pEPSCM-X and hEPSCM-X) for 5 d increased the TOPflash activity. **g**, Bright-field (left) and immunofluorescence (right) images showing pEPSCs^{Emb} cultured with the indicated changes in medium components. Cells were stained for OCT4 and DAPI. **h**,**i**, Quantitation of the AP$^+$ colonies formed from 2,000 pEPSCs^{Emb} (**h**) or H1-EPSCs (**i**) cultured on STO feeders with different medium components. The colonies were scored for five consecutive passages. −ROCKi, EPSCs passaged without the ROCK inhibitor Y27632. **j**,**k**, RT–qPCR analysis of the expression of lineage genes in pEPSCs^{Emb} (**j**) or hEPSCs (**k**) following the removal of XAV939 or Activin A, inhibition of TGFβ signalling by SB431542, or treatment with 3.0 μM CHIR99021. **l**, Effect on gene expression in EBs generated from H1-EPSCs when supplemented with 5.0 ng ml^{-1} Activin A. **m**,**n**, Effects of 5.0 ng ml^{-1} Activin A on PGCLC (tdTomato$^+$) production from the NANOS3–tdTomato reporter EPSCs assessed by FACS (**m**) and RT–qPCR (**n**). The relative expression levels were normalized to *GAPDH*. All graphs represent the mean ± s.d. **a**,**b**,**j**,**l**,**n**, $n = 3$ independent experiments. **f**–**i**,**m**, $n = 4$ independent experiments. The experiments in **d**,**e**,**g** were repeated independently three times with similar results. The *P* values were computed by a two-tailed Student's *t*-test. The statistical source data are provided in Supplementary Table 10. Scale bars, 100 μm.

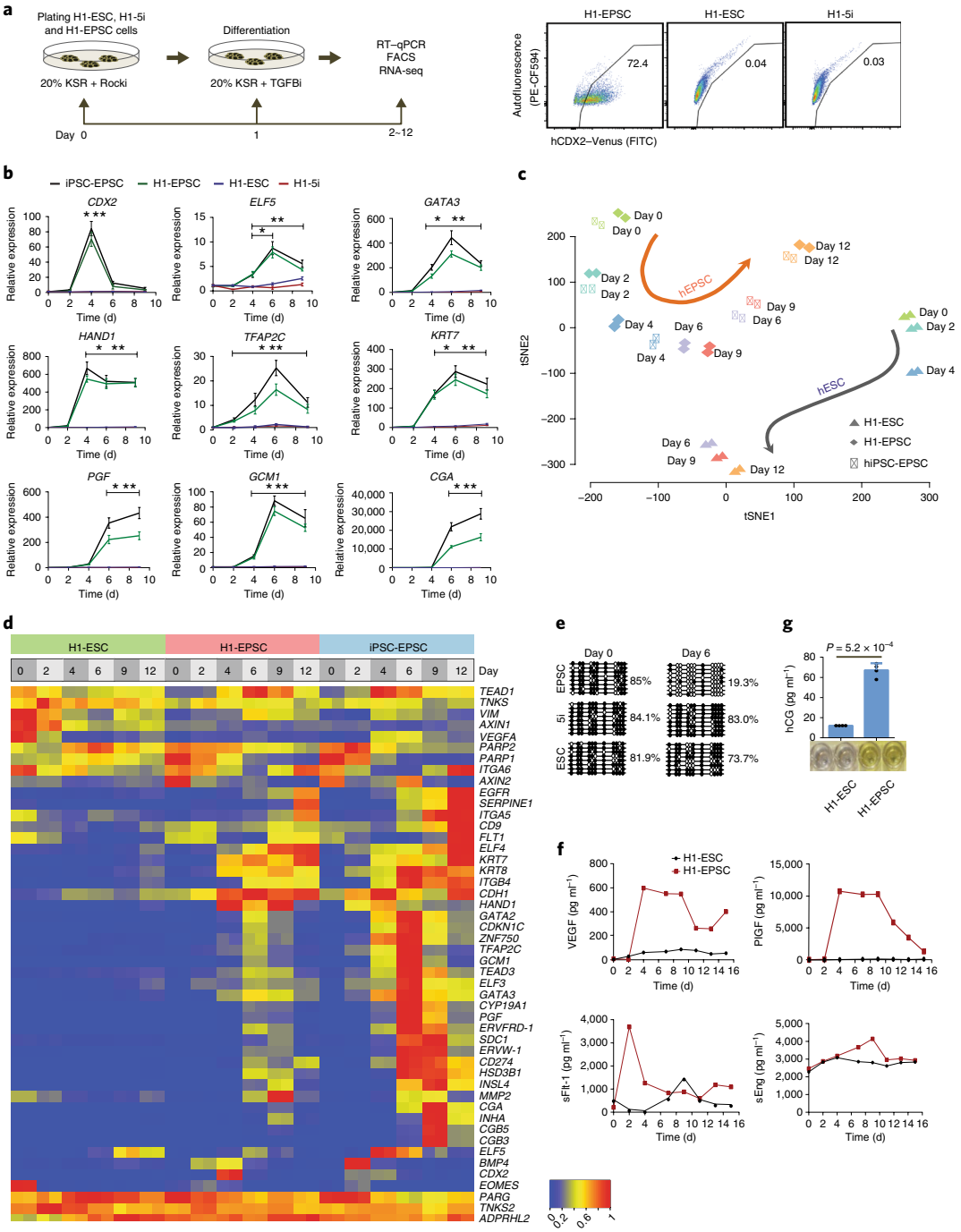

Fig. 7 | Trophoblast differentiation potential of hEPSCs. a, Differentiation of hEPSCs to trophoblasts under TGFβ inhibition (left). Flow cytometry analysis of trophoblast differentiation of *CDX2–H2B–Venus* reporter EPSCs, collected 4 d after TGFβ inhibition (right). *CDX2–H2B–Venus* reporter EPSCs were also cultured in conventional FGF-containing hESCs medium or 5i-naive medium and differentiated under TGFβ inhibition and examined by flow cytometry. The experiments were repeated independently three times with similar results. **b**, The dynamic changes in the expression of trophoblast genes during hEPSC differentiation were assayed by RT–qPCR. The relative expression levels were normalized to *GAPDH*. Data represent the mean ± s.d.; $n = 3$ independent experiments. $^{*}P < 0.01$, compared with H1-ESC cells; $^{**}P < 0.01$, compared with H1-5i cells. The *P* values were calculated using a two-tailed Student's *t*-test. The precise *P* values are presented in Supplementary Table 10. **c**, T-distributed stochastic neighbour embedding (tSNE) analysis of RNA-Seq data of the differentiating human ESCs ($n = 2$) and iPSC-EPSCs ($n = 4$) treated with SB431542. RNA was extracted from cells at days 0–12 of differentiation. The H1-EPSCs and hiPSC-EPSCs showed a different trajectory of differentiation from H1-ESCs. **d**, Heat map showing changes in the expression of trophoblast-specific genes in differentiating H1-ESCs, H1-EPSCs and iPSC-EPSCs collected at several time points of culture for RNA-Seq analysis. **e**, DNA demethylation at the promoter region of the *ELF5* locus in differentiating H1-EPSCs and other cell types following 6 d of SB431542 treatment. Cells from H1-ESCs, H1-naive ESCs (5i) showed no discernible DNA demethylation at the *ELF5* promoter. **f**, Levels of hormones secreted from trophoblasts derived from H1-EPSCs induced by TGFβ inhibition (SB431542). VEGF, PLGF, sFlt-1 and sEng were measured in the conditioned media for culturing the differentiating EPSCs or ESCs for 16 d following treatment with SB431542 for 48 h. **g**, Levels of hCG produced by trophoblasts from SB431542-treated EPSCs or ESCs at day 10 of differentiation, measured by ELISA (bottom). Data represent the mean ± s.d.; $n = 4$ independent experiments. The *P* values were calculated using a two-tailed Student's *t*-test. The statistical source data are presented in Supplementary Table 10.

inhibitors. Global gene expression profiling revealed that pEPSCs and human EPSCs (hEPSCs) were clustered together and were distinct from PFFs or other human pluripotent stem cells[15,35,36] (Fig. 4d,e). Both porcine and human EPSCs expressed high levels of key pluripotency genes and low levels of somatic cell lineage genes (*PAX6*, *T*, *GATA4* and *SOX7*) or placenta-related genes (*PGF*, *TFAP2C*, *EGFR*, *SDC1* and *ITGA5*; Supplementary Fig. 6a–d). Consistent with the high levels of global DNA methylation of pEPSCs and hEPSCs (Supplementary Fig. 6e), the DNA methyltransferase genes *DNMT1*, *DNMT3A* and *DNMT3B* were expressed at high levels, whereas *TET1*, *TET2* and *TET3* were expressed at lower levels (Supplementary Fig. 6f,g). Among the 76 genes (more than eightfold increase) that were expressed at high levels in H1-EPSCs compared with H1-ESCs, 17 encode histone variants with 15 belonging to histone cluster 1 (Fig. 4f and Supplementary Table 6). Interestingly, these histone genes were expressed at low levels in 5i and primed human ESCs but were highly expressed in human eight-cell and morula-stage embryos (Fig. 4g). The significantly higher expression of these histone genes was confirmed in additional hEPSC lines (Fig. 4h). The biological significance of this observation remains to be investigated.

scRNA-Seq reveals substantially homogenous EPSC cultures. EPSCs expressed uniform levels of the core pluripotency factors (Fig. 5a) and were generally homogenous cells in culture in the context of single-cell transcriptome (Fig. 5b). Mouse EPSCs had enriched transcriptomic features of four- to eight-cell blastomeres[15]. Single-cell RNA-Seq (scRNA-Seq) analysis suggested that the hEPSCs transcription profiles, as well as the histone gene expression profiles, were more similar to those of human eight-cell to morula-stage embryos[37,38] than other stages of pre-implantation embryos (Fig. 5c and Supplementary Fig. 6h; Fig. 4g,h and Supplementary Fig. 6i). Interestingly, single-cell transcriptome analysis also revealed low expression levels of naive pluripotency factors, such as *KLF2* in EPSCs (Fig. 5a and Supplementary Fig. 6a,b), which is also expressed at low levels in human early pre-implantation embryos[39]. Although *KLF2*, *TET1*, *TET2* and *TET3* were weakly expressed in both pEPSCs and hEPSCs (Supplementary Fig. 6a,b,f,g), their promoter regions were characterized by active H3K4m3 histone marks (Fig. 5d,e). In contrast to pluripotency genes, the cell-lineage gene loci (for example, *CDX2*, *GATA2*, *GATA4*, *SOX7* and *PDX1*) had high H3K27me3 and low H3K4me3 marks, respectively, in both porcine and human EPSCs (Fig. 5e).

Human and porcine EPSCs have similar signalling requirements. To identify the signalling requirements in EPSCs, we removed individual components from the culture medium. The removal of the SRC inhibitors WH-4-023 or A419259 reduced the expression of pluripotency factors in both EPSCs (Fig. 6a–d). Notably, using WH-4-023 instead of A419259 led to lower pluripotency gene expression in hEPSCs (Fig. 6b). As in mEPSCs, XAV939 enhanced the Axin1 protein content (Fig. 6e) and reduced canonical WNT activities in both EPSCs (Fig. 6f). The withdrawal of XAV939 caused a collapse of pluripotency and differentiation of these EPSCs (Fig. 6a,b,d,g–k). SMAD2/3 were phosphorylated in EPSCs (Fig. 6e). Massive cell loss and downregulation of pluripotency factors in pEPSCs resulted when either Activin A was removed from pEPSCM or the TGFβ inhibitor SB431542 was added (Fig. 6a,g,h,j). Human EPSCs did not require exogenous TGFB in culture but inhibition of TGFβ induced rapid cell differentiation with preferential expression of the trophoblast genes *CDX2*, *ELF5* and *GATA2* (Fig. 6b,i,k). At a relatively low concentration of Activin A (5.0 ng ml^{-1}), hEPSCs showed a stronger propensity for embryonic mesendoderm lineage differentiation (Fig. 6l) and generated more NANOS3-tdTomato^{+} PGCLCs (Fig. 6m,n). The removal of CHIR99021 and vitamin C from EPSCM did not affect pluripotency gene expression but reduced the number of colonies that formed from single cells (Fig. 6a,b,h,i), whereas a high CHIR99021 concentration (3.0 μM) induced differentiation of both EPSCs (Fig. 6a,h,j), as in human or rat naive cells[5,40]. The inhibition of JNK and BRAF might improve culture efficiency but was not essential (Fig. 6h,i). Mouse naive ESCs were cultured 1.0 μM Mek1/2 inhibitor PD0325901 (ref. [4]). We noticed that even 0.1 μM PD0325901 decreased pEPSC survival as measured by colony formation in serial passaging (Fig. 6h).

hEPSCs have potent potential to trophoblasts. We further investigated differentiation of hEPSCs to trophoblasts by generating the *CDX2–Venus* reporter line (Supplementary Fig. 7a). Inhibiting TGFβ with SB431542 resulted in approximately 70% of the reporter cells being *CDX2–Venus*$^{+}$ (Fig. 7a), whereas essentially no *CDX2–Venus*$^{+}$ cells were detected if the reporter cells were previously cultured in FGF or under the 5i-naive ESC conditions. Trophoblast gene levels were rapidly increased in differentiated H1-EPSCs and iPSC-EPSCs but not in H1-ESCs or H1-5i-naive cells (Fig. 7b). Addition of BMP4, which promotes differentiation of human ESCs to putative trophoblasts[41], induced a much higher level of expression of trophoblast genes in EPSCs than in H1-ESCs or H1-5i-naive ESCs (Supplementary Fig. 7b). Inhibition of FGF and TGFβ signalling while simultaneously activating BMP4 was previously reported to effectively induce trophoblast differentiation of human ESCs[42,43]. Under these conditions, the expression of trophoblast genes—especially the late trophoblast genes *GCM1*, *CGA* and *CGB*—was much higher in H1-EPSCs than in H1-ESCs, whereas naive 5i hESCs displayed no sign of trophoblast differentiation (Supplementary Fig. 7c). Global gene expression analysis demonstrated that under TGFβ-signalling inhibition H1-EPSCs and iPSC-EPSCs followed a differentiation trajectory distinct from that of H1-ESCs (Fig. 7c), and that in cells differentiated from EPSCs, but not from H1-ESCs, genes associated with trophoblast development or function were highly expressed, including: (1) *BMP4* (days 2–4); (2) *Syncytin-1* (*ERVW-1*) and *Syncytin-2* (*ERVFRD-1*), which promote cytotrophoblast fusion into syncytiotrophoblast; (3) *p57* (encoded by CDKN1C)[44,45]; (4) *CD274* (encoding PD-L1 or B7-H1) and (5) *EGFR*[46] (Fig. 7d).

We next performed Pearson correlation coefficient analysis of the transcriptome of cells differentiated under TGFβ inhibition with published reference data of primary human trophoblasts (PHTs) and human placenta tissues[43], which again revealed the similarity between cells differentiated from hEPSCs and PHTs and the placenta (Supplementary Fig. 7d). The differentiated cells from H1-EPSCs expressed human trophoblast-specific miRNAs (C19MC miRNAs: hsa-miR-525-3p, hsa-miR-526b-3p, hsa-miR-517-5p and hsa-miR-517b-3p, ref. [47]; Supplementary Fig. 7e,f) displayed DNA demethylation at the *ELF5* locus[48,49] (Fig. 7e) and produced abundant quantities of placental hormones (Fig. 7f,g).

One key mechanism for the derivation and maintenance of mouse, porcine and human EPSCs is blocking poly(ADP-ribosyl)ation activities of the PARP family members TNKS1/2 using small-molecule inhibitors such as XAV939[50,51]. In human cells, poly(ADP-ribose) in proteins is removed by poly(ADP-ribose) glycohydrolase (PARG) and ADP-ribosylhydrolase 3 (ARH3)[52]. Genetic inactivation of *Parp1/2* and *TNKS1/2* in the mouse results in trophoblast phenotypes[53], whereas inactivating *Parg* leads to a loss of functional trophectoderm and trophoblast stem cells (TSCs)[54]. In hEPSCs, *PARG*-deficiency did not appear to cause noticeable changes in EPSCs but adversely affected trophoblast differentiation (Supplementary Fig. 7g–j), which may indicate an evolutionally conserved mechanism for EPSCs and trophoblast development between mice and humans.

Derivation of TSC-like cells from human and porcine EPSCs. When hEPSCs (ESC-converted-EPSCs and iPSC-EPSCs) were cultured in human TSC conditions[46] with low cell density (2,000 cells

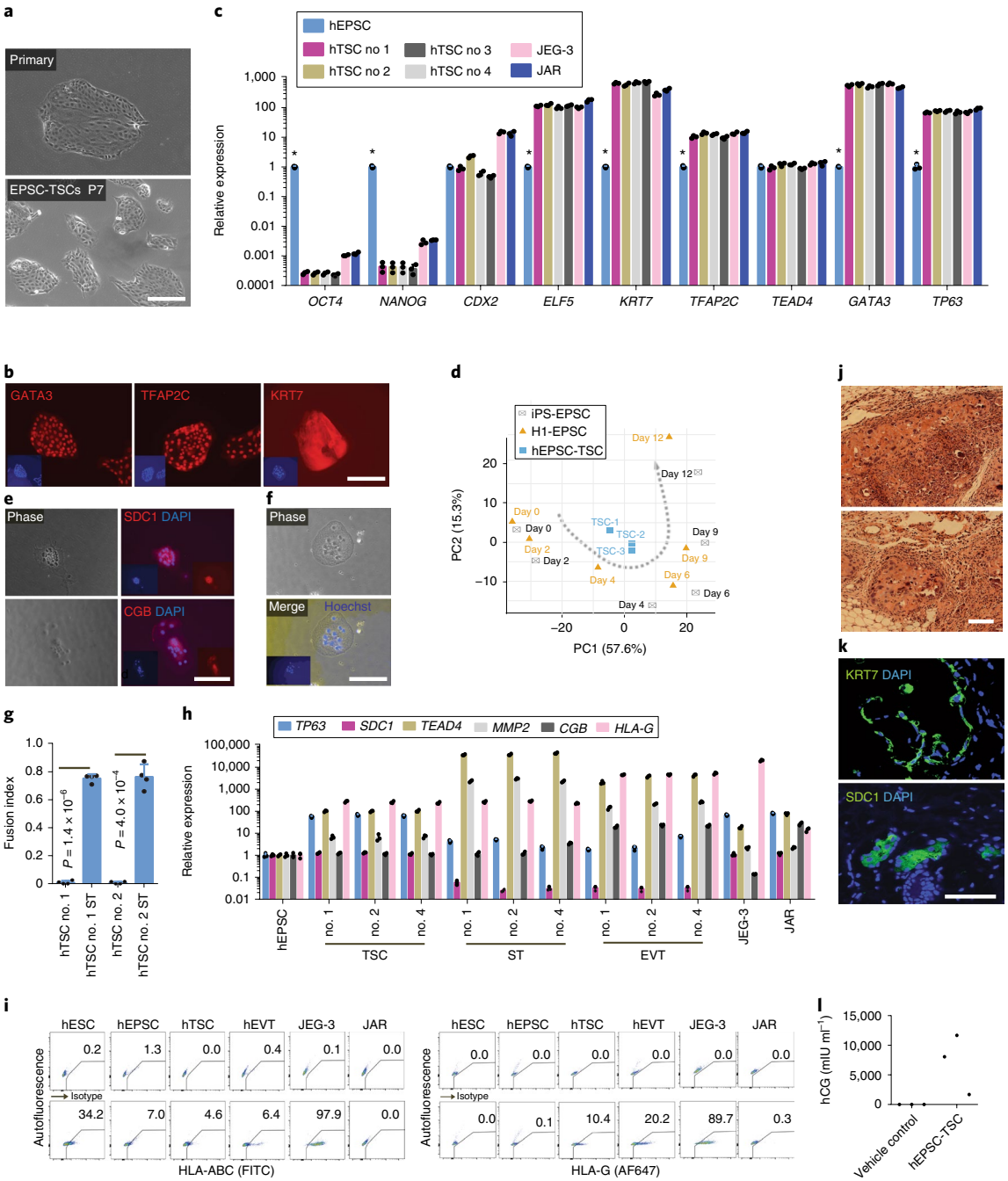

Fig. 8 | Derivation of trophoblast stem cell-like cells from hEPSCs. a, Phase-contrast images of primary TSC colonies formed from individual hEPSCs (top) and of TSCs at passage 7 (bottom). **b**, Expression of the trophoblast markers GATA3, TFAP2C and KRT7 in EPSC-TSCs detected by immunostaining. The nuclei were DAPI stained. Similar results were obtained in four independent EPSC-TSC lines. **c**, RT–qPCR analysis of pluripotency and TSC genes in four EPSC-derived TSC lines and their parental hEPSCs. JEG-3 and JAR are trophoblast cell lines. Data represent the mean ± s.d; $n = 3$ independent experiments. *$P < 0.01$ compared with TSCs. The P values were calculated using a two-tailed Student's t-test. **d**, PCA of gene expression of hTSCs ($n = 3$) and cells differentiated from human EPSCs ($n = 2$) under TGFβ inhibition at several time points. Enriched transcriptomic features of day-4 differentiated EPSCs were observed in hTSCs. **e**, Immunostaining of SDC1 and CGB in hTSC-dervied syncytiotrophoblasts. The nuclei are DAPI stained. **f**, Phase-contrast (top) and Hoechst-staining (bottom) images of multinucleated hTSC-derived syncytiotrophoblasts. **g**, Fusion index of forming syncytiotrophoblasts from hTSCs calculated as the number of nuclei in syncytia/total number of nuclei. Data represent the mean ± s.d.; $n = 4$ independent experiments. The P values were calculated using a two-tailed Student's t-test. **h**, RT–qPCR analysis of trophoblast-specific genes in syncytiotrophoblasts and EVTs derived from three hTSC lines. The expression levels were normalized to *GAPDH*. Data represent the mean ± s.d.; $n = 3$ independent experiments. **i**, Flow cytometry detection of HLA-ABC and HLA-G in hESCs, hEPSCs, hTSCs and hTSC-derived EVT cells (using the protocol from ref. [46]). The choriocarcinoma cells JEG-3 and JAR represent the extravillous and villous trophoblast cells, respectively. Unlike JAR cells, JEG-3 cells expressed HLA-G, HLA-C and HLA-E. **j**, Haematoxylin and eosin staining of lesions formed by hTSCs engrafted subcutaneously in NOD-SCID mice. The two images represent two areas of the same lesion. **k**, Confocal images of immunostaining for SDC1- or KRT7-positive cells in hTSC-derived lesions. The nuclei are DAPI stained. **l**, Serum hCG levels in six NOD-SCID mice 7 d after hTSC engraftment ($n = 3$) or injection of vehicle only ($n = 3$). The experiments in **a**,**b**,**e**,**f**,**h**–**l** were repeated independently three times with similar results. The statistical source data are presented in Supplementary Table 10. Scale bars, 100 μm. ST, syncytiotrophoblast.

per 3.5-cm dish), colonies with TSC morphology formed after 7–9 d (Fig. 8a). These colonies were picked and expanded into stable cell lines under TSC conditions with up to 30% efficiency in line establishment. These hEPSC-derived TSC-like cells were referred to as hTSCs in this study. In contrast, no lines were established from human H1 or M1 ESCs under the hTSC condition, whether or not they were originally cultured under primed or naive ESCs conditions. The hTSCs expressed the trophoblast transcription regulators GATA3 and TFAP2C but had downregulated pluripotency genes (Fig. 8b,c) and showed enriched transcriptomic features of day-4 differentiated human EPSCs under TGFβ inhibition (Fig. 8d). By following published protocols[46], we were able to differentiate hTSCs to multinucleated syncytiotrophoblasts and HLA-G$^+$ extravillous trophoblasts (EVT; Fig. 8e–i). Once injected into immunocompromised mice, hTSCs formed lesions with cells positively stained for SDC1 and KRT7 (Fig. 8j,k). In addition, high levels of hCG were detected in the blood of mice with hTSC-lesions but not in control mice injected with vehicle only (Fig. 8l). Although neither EPSCs expressed high levels of placenta development-related genes (Supplementary Fig. 6c,d), both displayed enriched H3K4me3 at these loci (Supplementary Fig. 8a), clearly underpinning the trophoblast potency of EPSCs. Stable TSC-like lines could also be derived from pEPSCs^{Emb} (pTSCs) under hTSC conditions (Supplementary Fig. 8b). The pTSCs were similar to hTSCs in gene expression profiles and the propensity of lesion formation in immunocompromised mice (Supplementary Fig. 8c–f). When introduced into porcine pre-implantation embryos, descendants of pTSCs were found in the trophectoderm and expressed GATA3 and CDX2 (Supplementary Fig. 8g). Our results therefore provide compelling evidence that human and porcine EPSCs possessed expanded potential that encompasses the trophoblast lineage.

In conclusion, murine, porcine and human EPSCs can now be established under similar in vitro culture conditions. These stem cells share common molecular features and possess expanded potency for both embryonic and extra-embryonic cell lineages that are generally not seen in conventional ESCs or iPSCs. Therefore, EPSCs represent a unique state of cellular potency. The successful generation of EPSCs produces tools for the investigation of embryonic development and opens avenues for translational research in biotechnology, agriculture, and genomic and regenerative medicine.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41556-019-0333-2.

Received: 9 September 2018; Accepted: 24 April 2019;
Published online: 3 June 2019

References

1. Evans, M. J. & Kaufman, M. H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154–156 (1981).
2. Martin, G. R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl Acad. Sci. USA* **78**, 7634–7638 (1981).
3. Thomson, J. A. et al. Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145–1147 (1998).
4. Ying, Q. L. et al. The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519–523 (2008).
5. Takashima, Y. et al. Resetting transcription factor control circuitry toward ground-state pluripotency in human. *Cell* **158**, 1254–1269 (2014).
6. Theunissen, T. W. et al. Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. *Cell Stem Cell* **15**, 471–487 (2014).
7. Ezashi, T., Yuan, Y. & Roberts, R. M. Pluripotent stem cells from domesticated mammals. *Annu. Rev. Anim. Biosci.* **4**, 223–253 (2016).
8. Brevini, T. A. et al. Culture conditions and signalling networks promoting the establishment of cell lines from parthenogenetic and biparental pig embryos. *Stem Cell Rev.* **6**, 484–495 (2010).
9. Vassiliev, I. et al. In vitro and in vivo characterization of putative porcine embryonic stem cells. *Cell Reprogram.* **12**, 223–230 (2010).
10. Haraguchi, S., Kikuchi, K., Nakai, M. & Tokunaga, T. Establishment of self-renewing porcine embryonic stem cell-like cells by signal inhibition. *J. Reprod. Dev.* **58**, 707–716 (2012).
11. Park, J. K. et al. Primed pluripotent cell lines derived from various embryonic origins and somatic cells in pig. *PLoS One* **8**, e52481 (2013).
12. Hou, D. R. et al. Derivation of porcine embryonic stem-like cells from in vitro-produced blastocyst-stage embryos. *Sci. Rep.* **6**, 25838 (2016).
13. Xue, B. et al. Porcine pluripotent stem cells derived from IVF embryos contribute to chimeric development in vivo. *PLoS One* **11**, e0151737 (2016).
14. Ma, Y., Yu, T., Cai, Y. & Wang, H. Preserving self-renewal of porcine pluripotent stem cells in serum-free 3i culture condition and independent of LIF and b-FGF cytokines. *Cell Death Discov.* **4**, 21 (2018).
15. Yang, J. et al. Establishment of mouse expanded potential stem cells. *Nature* **550**, 393–397 (2017).
16. Yang, J., Ryan, D. J., Lan, G., Zou, X. & Liu, P. In vitro establishment of expanded-potential stem cells from mouse pre-implantation embryos or embryonic stem cells. *Nat. Protoc.* **14**, 350–378 (2019).
17. Esteban, M. A. et al. Generation of induced pluripotent stem cell lines from Tibetan miniature pig. *J. Biol. Chem.* **284**, 17634–17640 (2009).
18. Ezashi, T. et al. Derivation of induced pluripotent stem cells from pig somatic cells. *Proc. Natl Acad. Sci. USA* **106**, 10993–10998 (2009).
19. West, F. D. et al. Porcine induced pluripotent stem cells produce chimeric offspring. *Stem Cells Dev.* **19**, 1211–1220 (2010).
20. Petkov, S., Glage, S., Nowak-Imialek, M. & Niemann, H. Long-term culture of porcine induced pluripotent stem-like cells under feeder-free conditions in the presence of histone deacetylase inhibitors. *Stem Cells Dev.* **25**, 386–394 (2016).
21. Lai, S. et al. Generation of knock-in pigs carrying Oct4-tdTomato reporter through CRISPR/Cas9-mediated genome engineering. *PLoS One* **11**, e0146562 (2016).
22. Zhang, W. et al. Pluripotent and metabolic features of two types of porcine iPSCs derived from defined mouse and human ES cell culture conditions. *PLoS One* **10**, e0124562 (2015).
23. Telugu, B. P., Ezashi, T. & Roberts, R. M. Porcine induced pluripotent stem cells analogous to naive and primed embryonic stem cells of the mouse. *Int. J. Dev. Biol.* **54**, 1703–1711 (2010).
24. Du, X. et al. Barriers for deriving transgene-free pig iPS cells with episomal vectors. *Stem Cells* **33**, 3228–3238 (2015).
25. Chen, H. et al. Erk signaling is indispensable for genomic stability and self-renewal of mouse embryonic stem cells. *Proc. Natl Acad. Sci. USA* **112**, E5936–E5943 (2015).
26. Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S. & Saitou, M. Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* **146**, 519–532 (2011).
27. Irie, N. et al. SOX17 is a critical specifier of human primordial germ cell fate. *Cell* **160**, 253–268 (2015).
28. Kobayashi, T. et al. Principles of early human development and germ cell program from conserved model systems. *Nature* **546**, 416–420 (2017).
29. Julaton, V. T. & Reijo Pera, R. A. NANOS3 function in human germ cell development. *Hum. Mol. Genet.* **20**, 2238–2250 (2011).
30. Gkountela, S. et al. The ontogeny of cKIT$^+$ human primordial germ cells proves to be a resource for human germ line reprogramming, imprint erasure and in vitro differentiation. *Nat. Cell Biol.* **15**, 113–122 (2013).
31. Camarasa, M. V. et al. Derivation of Man-1 and Man-2 research grade human embryonic stem cell lines. *In Vitro Cell. Dev. Biol. Anim.* **46**, 386–394 (2010).
32. Ye, J. et al. High quality clinical grade human embryonic stem cell lines derived from fresh discarded embryos. *Stem Cell Res. Ther.* **8**, 128 (2017).
33. International Stem Cell Initiative. Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. *Nat. Biotechnol.* **25**, 803–816 (2007).
34. Koyanagi-Aoi, M. et al. Differentiation-defective phenotypes revealed by large-scale analyses of human pluripotent stem cells. *Proc. Natl Acad. Sci. USA* **110**, 20569–20574 (2013).
35. Theunissen, T. W. et al. Molecular criteria for defining the naive human pluripotent state. *Cell Stem Cell* **19**, 502–515 (2016).
36. Yang, Y. et al. Derivation of pluripotent stem cells with in vivo embryonic and extraembryonic potency. *Cell* **169**, 243–257 (2017).
37. Yan, L. et al. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat. Struct. Mol. Biol.* **20**, 1131–1139 (2013).
38. Dang, Y. et al. Tracing the expression of circular RNAs in human pre-implantation embryos. *Genome Biol.* **17**, 130 (2016).
39. Blakeley, P. et al. Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. *Development* **142**, 3613 (2015).
40. Chen, Y., Blair, K. & Smith, A. Robust self-renewal of rat embryonic stem cells requires fine-tuning of glycogen synthase kinase-3 inhibition. *Stem Cell Rep.* **1**, 209–217 (2013).

41. Xu, R. H. et al. BMP4 initiates human embryonic stem cell differentiation to trophoblast. *Nat. Biotechnol.* **20**, 1261–1264 (2002).
42. Amita, M. et al. Complete and unidirectional conversion of human embryonic stem cells to trophoblast by BMP4. *Proc. Natl Acad. Sci. USA* **110**, E1212–E1221 (2013).
43. Yabe, S. et al. Comparison of syncytiotrophoblast generated from human embryonic stem cells and from term placentas. *Proc. Natl Acad. Sci. USA* **113**, E2598–E2607 (2016).
44. Chilosi, M. et al. Differential expression of p57kip2, a maternally imprinted cdk inhibitor, in normal human placenta and gestational trophoblastic disease. *Lab. Invest.* **78**, 269–276 (1998).
45. Zhang, P., Wong, C., DePinho, R. A., Harper, J. W. & Elledge, S. J. Cooperation between the Cdk inhibitors p27(KIP1) and p57(KIP2) in the control of tissue growth and development. *Genes Dev.* **12**, 3162–3167 (1998).
46. Okae, H. et al. Derivation of human trophoblast stem cells. *Cell Stem Cell* **22**, 50–63 (2018).
47. Lee, C. Q. et al. What is trophoblast? A combination of criteria define human first-trimester trophoblast. *Stem Cell Rep.* **6**, 257–272 (2016).
48. Hemberger, M., Udayashankar, R., Tesar, P., Moore, H. & Burton, G. J. ELF5-enforced transcriptional networks define an epigenetically regulated trophoblast stem cell compartment in the human placenta. *Hum. Mol. Genet.* **19**, 2456–2467 (2010).
49. Ng, R. K. et al. Epigenetic restriction of embryonic cell lineage fate by methylation of *Elf5*. *Nat. Cell Biol.* **10**, 1280–1290 (2008).
50. Huang, S. M. et al. Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. *Nature* **461**, 614–620 (2009).
51. Thorsell, A. G. et al. Structural basis for potency and promiscuity in poly(ADP-ribose) polymerase (PARP) and tankyrase inhibitors. *J. Med. Chem.* **60**, 1262–1271 (2017).
52. Hassa, P. O. & Hottiger, M. O. The diverse biological roles of mammalian PARPS, a small but powerful family of poly-ADP-ribose polymerases. *Front. Biosci.* **13**, 3046–3082 (2008).
53. Hemberger, M. et al. Parp1-deficiency induces differentiation of ES cells into trophoblast derivatives. *Dev. Biol.* **257**, 371–381 (2003).
54. Koh, D. W. et al. Failure to degrade poly(ADP-ribose) causes increased sensitivity to cytotoxicity and early embryonic lethality. *Proc. Natl Acad. Sci. USA* **101**, 17699–17704 (2004).

Acknowledgements

We thank our colleagues of the Wellcome Trust Sanger Institute core facilities for their generous support (J. Bussell, Y. Hooks, N. Smerdon, B. L. Ng and J. Graham) and A. Moffett for critical comments. We thank B. Petersen for performing the surgical embryo transfers and the staff of the experimental pig facility at the Friedrich-Loeffler-Institut Mariensee for their competent and enduring assistance. We acknowledge the following funding and support: the Wellcome Trust (grant nos. 098051 and 206194) to the Sanger Institute and the University of Hong Kong internal funding (P. Liu); a Wellcome Trust Clinical PhD Fellowship for Academic Clinicians (D.J.R.); a PhD fellowship from the Portuguese Foundation for Science and Technology, FCT (grant no. SFRH/BD/84964/2012; L.A.); a Marie Sklodowska-Curie Individual Fellowship (M.A.E.-M.); the BBSRC (grant no. BB/K010867/1), Wellcome Trust (grant no. 095645/Z/11/Z), EU EpiGeneSys and BLUEPRINT (W.R.); a Chongqing Agriculture Development Grant (grant no. 17407 to L.P.G., Z.H.L. and Y.H.); REBIRTH project no. 9.1, Hannover Medical School (H.N.); Shuguang Planning of Shanghai Municipal Education Commission (grant no. 16SG14) and the National Key Research and Development Program (grant no. 2017YFA0104500; L. Lu); the China Postdoctoral Science Foundation (grant no. 2017M622795; D.C.); the Strategic Priority Research Program of CAS (grant nos. XDA16030503 and XDA16030501), the National Key Research and Development Program of China Stem Cell and Translational Research (grant no. 2017YFA0105103) and Key Research and Development Program of Guangzhou Regenerative Medicine and Health Guangdong Laboratory (grant no. 2018GZR110104004; L.Lai); the Shenzhen Municipal Government of China (DRC-SZ [2016] 884; Z.S.); the NHMRC of Australia (Senior Principal Research Fellowship grant no. 1110751; P.P.L.T.); the GRF of Hong Kong (grant nos 17119117 and 17107915) and the National Natural Science Foundation (grant nos. 81671579 (L. Lu), 31471398 (W.S.B.Y.) and U1804281 (Y. Zhang)).

Author contributions

X.G. developed the culture conditions for the pEPSCs and hEPSCs and performed most of the experiments. M.N.-I. performed the pig experiments and wrote the paper. D.H. was critically involved in all porcine experiments. S.P. provided some pig reprogramming factor genes. X.C. performed most of the informatics analyses with support from S.A.T. D.C., S.W., J. Zhong, J. Zhu, Z.S. and J.W. analysed the RNA-Seq data of hEPSCs/ESCs to trophoblasts. A.C.H.C., Y.L.L. and W.S.B.Y. performed the teratoma and TSC lesion experiments at the HKU. M.A.E.-M. and W.R. performed the EPSC DNA methylation analysis. T.K. and A.S. helped X.G. with the PGCLC analysis. S.A. and A.A. measured hormones in cells differentiated from hEPSCs. L.S.C. analysed the teratoma sections. F.Y. and B.F. karyotyped cells. D.Ruan, X.W., L.G., Z.L., Y.H., T.N., D.W., D.P., P. Li, L. Lai, G.L., D. Ryan, J.Y., L.A., Y.Y., S.-G.X., Y.Z., L. Lu and X.Z. were involved in refining the culture conditions or provided intellectual input. S.J.K. provided human M1 and M10 ESCs. P.P.L.T. provided intellectual input and edited the manuscript. H.N. conceptualized the pig experiments, wrote the paper and secured funding for the pig part of the experiments. P. Liu conceived the pEPSC culture screen concept, supervised the research and wrote paper.

Competing interests

Patent applications have been filed relating to the data presented here on behalf of the Sanger Institute and the University of Hong Kong.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41556-019-0333-2.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to H.N. or P.L.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Ethical considerations of working with human cells and animals. The experiments using human ESCs and human cells were approved by the HMDMC of the Wellcome Trust Sanger Institute. The experiments using pig embryos were approved by the Niedersaechsisches Landesamt fuer Verbraucherschutz und Lebensmittelsicherheit (LAVES). The mouse teratoma experiments were performed in accordance with UK Home Office regulations and the Animals (Scientific Procedures) Act 1986 (licence number 80/2552), and were approved by the Animal Welfare and Ethical Review Body of the Wellcome Genome Campus and the Committee on the Use of Live Animals in Teaching and Research, The University of Hong Kong.

Culturing EPSCs. Porcine/human EPSC cells were maintained on STO feeder layers and enzymatically passaged every 3–5 d by a brief PBS wash followed by treatment with 0.25% trypsin/EDTA for 3–5 min. The cells were dissociated and centrifuged (300*g* for 5 min) in 10% fetal bovine serum (FBS)-containing medium. After removing the supernatant, the porcine/human EPSCs were resuspended and seeded in pEPSCM/hEPSCM supplemented with 5 μM Y27632 (Tocris, cat. no. 1254). The addition of 5% FBS (Gibco, cat. no. 10270) and 10% KnockOut Serum Replacement (KSR; Gibco, cat. no. 10828-028) improved the survival of cells during passaging. The medium was switched to pEPSCM/hEPSCM only 12–24 h later. Both pEPSCM and hEPSCM are N2B27-based media. N2B27 basal media (500 ml) was prepared as follows: 482.5 ml DMEM/F-12 (Gibco, cat. no. 21331-020), 2.5 ml N2 supplement (Thermo Fisher Scientific, cat. no. 17502048), 5.0 ml B27 supplement (Thermo Fisher Scientific, cat. no. 17504044), 5.0 ml 100× glutamine penicillin-streptomycin (Thermo Fisher Scientific, cat. no. 11140-050), 5.0 ml 100× NEAA (Thermo Fisher Scientific, cat. no. 10378-016) and 0.1 mM 2-mercaptoethanol (Sigma, cat. no. M6250). To make pEPSCM (500 ml), the following small molecules and cytokines were added into 500 ml N2B27 basal media: 0.2 μM CHIR99021 (GSK3i; Tocris, cat. no. 4423), 0.3 μM WH-4-023 (Tocris, cat. no. 5413), 2.5 μM XAV939 (Sigma, cat. no. X3004) or 2.0 μM IWR-1 (Tocris, cat. no. 3532), 65.0 μg ml^{-1} vitamin C (Sigma, cat. no. 49752-100G), 10.0 ng ml^{-1} LIF (Stem Cell Institute (SCI), University of Cambridge), 20.0 ng ml^{-1} Activin (SCI) and 0.3% FBS (Gibco, cat. no. 10270). To make hEPSCM (500 ml), the following components were added into 500 ml N2B27 basal media: 1.0 μM CHIR99021, 0.1 μM A419259 (Tocris, cat. no. 3914), 2.5 μM XAV939 or 2.0 μM IWR-1, 65 μg ml^{-1} vitamin C and 10 ng ml^{-1} LIF (SCI). In addition, 0.25 μM SB590885 and 2.0 μM SP600125, or 1.0 % ITS-X, could be included to improve EPSC cultures, but they were not essential for the routine maintenance of porcine and human EPSCs. All cell cultures in this paper were maintained at 37 °C with 5% CO_2, unless stated otherwise.

Reprogramming PFFs to iPSCs. Germany Landrace and China TAIHU OCT4–tdTomato PFFs (1.5 × 10^6 per experiment) were transfected using an NHDF Nucleofector kit with 6.0 μg DNA (2.0 μg *PB–TRE-pOSCK* (porcine *OCT4*, *SOX2*, *cMYC* and *KLF4*), 1.0 μg *PB–TRE–pNhL* (porcine *NANOG* and human *LIN28*)[20], 1.0 μg *PB–TRE–hRL* (human *RARG* and *LRH1*)[55], 1.0 μg *PB–EF1a–transposase* and 1.0 μg *PB–EF1a–rTTA*). Dox (1.0 μg ml^{-1}; Sigma, cat. no. D9891) was used to induce expression of the reprogramming factors. For transgene-dependent iPSC generation, the colonies were picked at day 12 into M15 (knockout DMEM (Gibco, cat. no. 10829-018), 15% FBS (Gibco, cat. no. 10270), 1× glutamine penicillin-streptomycin (Thermo Fisher Scientific, cat. no. 11140-050), 1× NEAA (Thermo Fisher Scientific, cat. no. 10378-016) and 0.1 mM 2-mercaptoethanol (Sigma, cat. no. M6250)) supplemented with Dox, 50 μg ml^{-1} vitamin C and 10 ng ml^{-1} bFGF. To directly establish transgene independent iPSCs lines in pEPSCM, Dox was removed at day 9 and the media was switched to pEPSCM. The iPSC colonies were picked and cultured in pEPSCM supplemented with 5 μM Y27632 on day 14 or 15. Y26537 was removed from the culture media 24 h later.

Screening for the culture conditions of pEPSCs. Dox-dependent porcine iPSCs were dissociated using trypsin and seeded in 24-well STO feeder plates at a density of 1 × 10^4 cells well^{-1}. The cells were cultured for 2 d in M15 supplemented with Dox, vitamin C and 10 ng ml^{-1} bFGF before the culture media was switched to the test media (Supplementary Table 1). M15 and N2B27 media were prepared as above. AlbumMax media comprised: DMEM/F-12, 20% AlbumMax II (Gibco, cat. no. 11021-037), 25 mg ml^{-1} human insulin (Sigma, cat. no. 91077 C), 2.0× B27 Supplement, 100.0 μg ml^{-1} IGFII (R&D, cat. no. 292-G2-250), 1× glutamine penicillin-streptomycin, 1 × NEAA and 0.1 mM 2-mercaptoethanol. KSR media (20%) comprised: DMEM/F-12, 20% KSR, 1× glutamine penicillin-streptomycin, 1× NEAA and 0.1 mM 2-mercaptoethanol. Small molecules and cytokines were supplemented as indicated at the following final concentrations: CHIR99021, 0.2 or 3.0 μM; PD0325901, 0.2 or 1.0 μM; WH-4-023, 0.5 μM, PKC inhibitor Go6983, 5.0 μM; SB203580 (p38 inhibitor), 10.0 μM; SP600125 (JNK inhibitor), 4.0 μM; vitamin C, 65.0 μg ml^{-1}; SB590885 (BRAF inhibitor), 0.25 μM; XAV939, 2.5 μM; RO4929097 (Notch-signalling inhibitor), 10.0 μM; LDN193189 (BMP inhibitor), 0.1 μM; Y27632, verteporfin (YAP inhibitor), 10 μM; LIF, 10 ng ml^{-1}; BMP4, 10 ng ml^{-1}; SCF, 50.0 ng ml^{-1}; EGF, 50 ng ml^{-1}; TGFB, 10.0 ng ml^{-1}; bFGF, 10.0 ng ml^{-1} and Activin A, 20.0 ng ml^{-1}. The medium was refreshed daily and the surviving cells were passaged at day 6 with 5 μM Y27632. Endogenous porcine *OCT4* and *NANOG* expression was checked after 4 d.

Sow superovulation. Peripubertal German Landrace gilts were synchronized by feeding 5 ml per day per gilt altrenogest (Regumate; 4 mg ml^{-1}; MSD Animal Health) for 13 d. An injection of 1,500 international units (IU) pregnant mare serum gonadotropin (PMSG) was administered on the last day of altrenogest feeding[56]. Ovulation was induced by intramuscular injection of 500 IU hCG (Ovogest; MSD) 76 h later.

Sows insemination and embryo recovery. The sows were artificially inseminated twice at 40 and 48 h after hCG administration with semen from Germany Landrace boars. Five days later, the sows were killed and the embryos were flushed with Dulbecco's PBS medium supplemented with 1% newborn calf serum. The collected morulae were either directly used for injection experiments or cultured overnight in PZM-3 medium to the blastocyst stage and used for inner cell mass (ICM) isolation.

Generation of parthenogenetic embryos. Oocytes isolated from abattoir ovaries were matured in vitro in 1:1 DMEM high glucose and Ham's F-12 medium supplemented with 60 μg/ml penicillin G potassium salt, 50 μg/ml streptomycin sulfate, 2.5 mM L-glutamine, FBS, murine EGF, 10 IU ml^{-1} PMSG, 10 IU ml^{-1} human chorionic gonadotropin, 100 ng ml^{-1} human recombinant IGF1 and 5 ng ml^{-1} recombinant human bFGF for 40 h in humidified air with 5% CO_2 at 38.5 °C. Matured oocytes were enucleated as described previously[57]. Enucleated oocytes were exposed to a single pulse of 24 V for 45 μs in SOR2 activation medium[57] followed by incubation in 2 mM 6-dimethylaminopurine in PZM-3 medium for 3 h. For the isolation of ICMs, blastocysts from day 6 were cultured for an additional 24 h in D15 medium (DMEM high glucose, 2 mM L-glutamine, 15% FBS, 1% penicillin/streptomycin solution, 1% MEM non-essential amino acids solution and 0.1 mM 2-mercaptoethanol supplemented with LIF).

Outgrowths from porcine blastocysts. Parthenogenetic blastocysts from day 7 and in vivo-derived blastocysts from day 5 were used to obtain ICMs by placing them in Ca^{2+}-TL-HEPES medium by microsurgery using ophthalmic scissors. Isolated ICMs were cultured for 7 d on STO cells in pEPSCM, in the presence of 10 μM Y27632, until initial outgrowths could appear. The outgrowths were mechanically isolated and reseeded onto fresh STO cells in pEPSCM. The cells formed well-defined porcine EPSC^{Emb} colonies 3 d later.

In vitro chimera assay. Small clumps of 6–8 EPSCs^{Emb} or EPSCs^{iPS} expressing mCherry were resuspended in D15 medium containing mouse LIF and 10 μM Y27632 and injected into day 4 or 6 porcine parthenogenetic embryos using a piezo-driven micromanipulator in Opti-MEM I (1×) + GlutamMAX-I reduced serum medium supplemented with 10% FBS. After injection, embryos were cultured in D15 medium at 39 °C in 5% CO_2 and 5% O_2 for 24 h (for day 6 blastocysts) or 48 h (for day 4 embryos). Uninjected embryos (day 4 or 6) were used as controls for embryo development.

In vivo chimera assay. Two days before injection, the pEPSC medium was switched to pEPSCM without WH-4-023 (pEPSCM minus SRCi). One day before injection, the medium was replaced with pEPSCM minus SRCi supplemented with 5 ng ml^{-1} heparin and 10 ng ml^{-1} bFGF. The medium was then replaced with pEPSCM-SRCi supplemented with 5 ng ml^{-1} heparin, 10 ng ml^{-1} bFGF, 10 ng ml^{-1} human LIF, 5 μM Y27632, 20 ng ml^{-1} human recombinant Activin A and 10% FBS 4 h before injection.

During injection, EPSCs^{Emb} were added to a 500 μl drop of M15 medium supplemented with 50 μg ml^{-1} vitamin C, 0.1 μM CHIR99021, 20 ng ml^{-1} human recombinant Activin A, 10 ng ml^{-1} bFGF, 10 ng ml^{-1} human LIF, 5 ng ml^{-1} heparin and 5 μM Y27632 and plated under phase-contrast inverted microscope equipped with a microinjection system. Day 5 porcine morulae were added to a 500 μl drop of Opti-MEM I (1×) + GlutamMAX-I reduced serum medium supplemented with 20 ng ml^{-1} human recombinant Activin A, 10 ng ml^{-1} bFGF, 5 μM Y27632 and 10% FBS. After injection, the morulae were either incubated for 4 h until the embryo transfer in medium used for the injection or cultured overnight and fixed for confocal microscopy analysis.

Evaluation of chimerism in porcine blastocysts cultured in vitro. Porcine chimeric blastocysts were fixed in 3.7% formaldehyde solution for 15 min at room temperature. Thereafter, the embryos were incubated with 0.2 μM SiR-DNA for 30 min at 37 °C to visualize the nuclei. Blastocysts were analysed using a confocal screening microscope.

Flow cytometry of dissected pig chimera tissues and EBs for PGCLCs. The fetuses were dissected into small pieces representing several pieces of the head, trunk and tail. The dissected tissues and placenta were dissociated with 1.0 mg ml^{-1} collagenase IV for 1–3 h at 37 °C on a shaker. The dissociated cells were filtered with a 35-μm nylon mesh and fixed using Fixation medium according to the manufacturer's manual (BD Cytofix, cat. no. 554655). PGC EBs were trypsinized with 0.25% trypsin/EDTA and stained with PerCP-Cy5.5-conjugated anti-TNAP

antibody. NANOS3–H2B–mCherry$^+$/TNAP$^+$ cells were detected using 561 nm (610/20 bandpass filter) and 488 nm (710/50 bandpass filter) channels. FACS data were analysed by FlowJo software. The antibodies used in these experiments are listed in Supplementary Table 7.

Differentiation of porcine EPSCs to PGCLCs. The *piggyBac*-based *PB–TRE–NANOG*, *PB–TRE–BLIMP1*, *PB–TRE–TFAP2C* and *PB–CAG–SOX17–GR* expression constructs were transfected into the pig NANOS3–2A–H2B–mCherry reporter EPSCs^{emb}. Thereafter, the expression of transgenic *NANOG*, *BLIMP1* and *TFAP2C* was induced by 1.0 μg ml^{-1} Dox for the indicated time periods. The SOX17 protein was translocated into the nucleus by the addition of 2.0 μg ml^{-1} dexamethasone (Sigma, cat. no. D2915). Pre-differentiated cells were collected and plated to ultra-low attachment U-bottom 96-well plates (Corning, cat. no. 7007) at a density of 5,000–6,000 cells well^{-1} in 100 μl PGCLC medium. After 3–4 d, the EBs were collected for analysis. PGCLC medium is composed of Advanced RPMI 1640, 1% B27 supplement, 1.0× glutamine penicillin-streptomycin, 1.0× NEAA, 0.1 mM 2-mercaptoethanol and the following cytokines: 500.0 ng ml^{-1} BMP2, 10.0 ng ml^{-1} human LIF, 100.0 ng ml^{-1} SCF, 50.0 ng ml^{-1} EGF and 10.0 μM Y27632. For human PGCLCs, we tested the PGC differentiation potential of two hEPSC lines with the sequential induction method[28].

Teratoma assay of pig and human EPSCs. Porcine and human EPSCs were resuspended in PBS supplemented with 30% matrigel (Corning, cat. no. 354230) and 5.0 μM Y27632. EPSCs (5×10^6) were subcutaneously injected into both dorsal flanks of NSG mice. EPSCs formed visible teratomas between 8 and 10 weeks. When the size of the teratomas reached 1.2 cm^3, they were collected and processed for sectioning.

EB formation assay of EPSCs. Pre-differentiated EPSCs were detached using 0.25% trypsin/EDTA and plated to ultra-low cell attachment U-bottom 96-well plates at a density of 5,000–6,000 cells well^{-1} in 200 μl M10 medium. After 7–8 d of culturing, the EBs were collected for analysis.

CRISPR/Cas9-mediated genome editing in porcine and human EPSC cells. To target an EF1a-H2B–mCherry-iRES-Puro cassette to the porcine ROSA26 locus, Rosa 5′ and 3′ homology arms were synthesized by the IDT Company (650-bp 5′ arm, Chr13: 65756272–65756923; 648-bp 3′ arm, Chr13: 65,755,620–65,756,267). The sequence 5′-CAATGCTAGTGCAGCCCTCATGG-3′ was designed as the target of gRNA/Cas9. For porcine NANOS3, homology arms were also synthesized by the IDT Company (699-bp 5′ arm, Chr2: 65275456–65276148; 699-bp 3′ arm Chr2: 65274749–65275447). A 20-bp sequence (5′-TCCACTTCTGCCTAAGAGGCTGG-3′) preceding the stop codon was targeted by gRNA/Cas9 to introduce the cut and mediate homologous recombination. The same strategy was employed to make the human *OCT4–T2A–H2B–Venus* and *CDX2–T2A–H2B–Venus* reporter EPSC lines. The human OCT4 homology arms are: 619-bp 5′ arm (Chr6: 31164604–31165222) and 636-bp 3′ arm (Chr6: 31163965–31164600). The gRNA/Cas9 targeting sequence is 5′-TCTCCCATGCATTCAAACTGAGG-3′. The CDX2 homology arms are: 478-bp 5′ arm (Chr13: 27963118–27963595) and 557-bp 3′ arm (Chr13: 27962558–27963114). The gRNA/Cas9 targeting sequence is 5′-CCGTCACCCAGTGACCCACCGGG-3′.

Conversion of human ESCs/iPSCs to EPSCs. Trypsinized single cells (5×10^4) were seeded on a STO feeder (9-cm dish) in bFGF-containing media with 5.0 μM Y27632. The standard hESC media used was: DMEM/F-, 20% KSR, 1.0× glutamine penicillin-streptomycin, 1.0× NEAA, 0.1 mM 2-mercaptoethanol and 10.0 ng ml^{-1} bFGF (SCI). A day later, the medium was switched to hEPSCM. After about 5–6 d, hEPSC colonies emerged with most ESCs differentiated and could be expanded in bulk. Single colonies could also be picked and expanded following the method described above.

Reprogramming human fibroblasts to EPSCs. The DNA mixture consisted of 2.0 μg PB–TRE-hOCKS, 1.0 μg PB–TRE-RL, 1.0 μg PB-EF1a-transposase and 1.0 μg PB-EF1a-rtTA. Transfected cells (0.2×10^6; GM00013, Coriell Institute; Amaxa Nucleofector) were seeded on STO (10-cm dish) in M15 supplemented with 50 μg ml^{-1} vitamin C and Dox (1.0 μg ml^{-1}). Dox was removed at days 12–14 and the media was switched to hEPSCM. The surviving colonies were picked to hEPSCM at about day 21 and expanded to stable iPSC-EPSC lines.

Differentiation of hEPSCs to trophoblast lineages. Human EPSCs were dissociated with 0.25% trypsin/EDTA and seeded in gelatinized six-well plates at a density of 0.1×10^6 cells well^{-1}. The cells were cultured in 20% KSR media supplemented with 5 μM Y27632 for 1 d. From the second day, different combinations of SB431542 (10 μM), BMP4 (50 ng ml^{-1}) and the FGF receptor inhibitor PD173074 (0.1 μM) were added into 20% KSR media to start the trophoblast differentiation. The cells were collected at the indicated time points for analysis.

Derivation of stable TSC cell lines from EPSCs. Single hEPSCs or pEPSCs^{Emb} were plated on six-well plates pre-coated with 1.0 mg ml^{-1} Col IV (Corning, cat. no. 354233) at a density of 2,000 cells well^{-1} and cultured in hTSC media as described[46] with a minor modification. After about 7–9 d of culture, the TSC-like colonies were picked, dissociated in TrypLE and re-plated into a 12-well dish pre-coated with 1.0 mg ml^{-1} Col IV. After 4 or 5 passages, the cells were collected for syncytiotrophoblast and extravillous trophoblast (EVT) differentiation tests[46].

TSC lesion assay. TSCs were dissociated with TrypLE and resuspended in PBS supplemented with 30% matrigel and 10 μM Y27632. TSCs (5×10^6 in 100 μl) were subcutaneously injected into both dorsal flanks of eight-week-old male SCID mice. TSCs formed visible lesions within 7–10 d. The lesions were dissected, fixed overnight in 4% phosphate-buffered formalin and embedded in OCT compound and paraffin for sectioning.

RNA-Seq analysis of global gene expression in EPSCs and hTSCs. The cells for RNA preparation were collected from the same batch of culture when the culture had reached 70–80% confluence. Biological replicates were included to allow meaningful conclusions. For human data, protein coding transcripts from GENCODE v27 were used and transcripts from PAR_Y regions were removed from the reference; for mouse data, protein coding transcripts from GENCODE vM16 were used; for pig data, Ensembl build Sscrofa11.1 was used. Human naive and primed ESC RNA-Seq[6] data were downloaded from ENA (Study accession no. PRJNA326944); human embryo single-cell data were downloaded from ENA (Study accession nos PRJNA153427 and PRJNA291062)[37,38]. Mouse EPSC data were from our previous study[15]. The commands used to process the data can be found in the GitHub repository (https://github.com/dbrg77/pig_and_human_EPSC). The expression levels of each selected histone gene in different types of human cells and early embryos were extracted from the expression matrix and visualized as a heat map generated by GraphPad Prism 7.04 (https://www.graphpad.com/scientific-software/prism/). The gene-expression values were linearly transformed into colours (as indicated by the colour legend below each matrix). For scRNA-Seq, we added an extra quality control step where cells with fewer than 10,000 total reads, less than 4,000 detected genes (at least 1 read), more than 80% of reads mapped to ERCC or more than 60% of non-mappable reads were removed before downstream analyses.

Batch correction, PCA and cross-species comparison. The gene count from each sample was collected and log10 transformed. The batch effect (batches here mean different studies) and sequencing depth (total number of reads per sample) were then regressed out using the 'regress_out' function from the NaiveDE package (https://github.com/Teichlab/NaiveDE/tree/master/NaiveDE). PCAs were done on the regressed matrix using Scikit-learn[58]. For cross-species comparisons, only the one-to-one orthologous genes were used.

RNA-Seq analysis of human EPSC differentiation to trophoblasts. A reference index was created based on hg38 from the GENCODE database[59]. Gene expression matrices were generated using Salmon[60] with the following parameters: salmon quant–no-version-check -q -p 6–useVBOpt–numBootstraps 100–posBias–seqBias–gcBias. For tSNE analysis, the R package 'Rtsne' was used for the dimension reduction of gene expression matrices (genes with maximum TPM ≤ 1 were filtered out) and the corresponding result was visualized using a custom R script. For the Pearson's correlation, the RNA-Seq data for reference tissues was downloaded from Chang et al.[61] and the data for reference cells (uESCs, uPHTs, dESCs and dPHTs) was downloaded from the paper by Yabe and colleagues[43]. A list of tissue-specific genes ($n=2,293$) defined by Chang et al.[61] were selected for Pearson correlation coefficients analysis. A pairwise calculation was performed between our data (H1-ESC, H1-EPSC and hiPSC-EPSC) and external references. The expression levels of each trophoblast gene were extracted from the expression matrix and normalized using the following method. The TPM of a given gene was divided by the highest gene-expression level of that gene in a row (12 data points for each cell line, in 36 total values for H1-ESC, H1-EPSC and hiPSC-EPSC). Through this method, each TPM was transformed into a value between zero and one. The overall gene signatures were plotted as a heat map.

PCA analysis of hTSC RNA-Seq. We applied the 'factoextra' R package for PCA analysis and 'limma' R package for batch-effect removal. Genes whose TPM values were lower than one in all samples were removed from the TPM expression matrix.

Construction of scRNA-Seq libraries. The single-cell messenger RNA-Seq library was generated following the SMART-seq2 described protocol[62]. The quality of the library was then assessed by a Bioanalyzer (Agilent) before submission to the DNA sequencing pipeline at the Wellcome Trust Sanger Institute. Pair-ended 75-bp reads were generated by HiSeq2000 sequencers.

ChIP-seq analysis of histone modification profiles in EPSCs. The H3K4me3, H3K27me3, H3K27ac and input ChIP libraries of EPSCs were prepared according to Lee and colleagues[63]. The multiplex sequencing libraries were prepared with the microplex library construction kit (Diagenode, cat. no. C05010014). DNA was amplified for 11 cycles and the library was checked on a Bioanalyzer (Ailgent) using a high-sensitivity DNA kit. The library concentration was checked by

qPCR using KK4824 and equal molarities of different libraries were pooled and sequenced by HiSeq2500.

Bowtie2 (version 2.3.4)[64] with default settings was used to map 50-bp single-end reads to the UCSC reference genomes (build susScr11 for pig and hg38 for human). For the ChIP-seq data from human naive and primed ESCs[6], raw reads were downloaded from ENA (study accession no. PRJNA255308) and processed in the same manner.

Peak calling was performed using MACS2 (2.1.1.20160309)[65]. For the identification of enriched regions of punctate marks (H3K4me3 and H3K27ac) from pig samples, peak calling was performed with flags '-g 2.7e9 -q 0.01 -f BAM–nomodel–extsize 200'. For the identification of enriched regions of broad marks (H3K27me3), peak calling was performed with flags '-g 2.7e9 -q 0.01 -f BAM–nomodel–extsize 200–broad'. For human data, peak calling was done in the same way with a change of genome size '-g hs' during the peak calling. The resulting bedGraph files were converted to bigWig files using the script bdg2bw (https://gist.github.com/jl32587/34370c995460f9d5ad65). The bigWig files were visualized using the UCSC genome browser[66].

Statistical analysis and reproducibility. No statistical methods were used to predetermine the sample sizes. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment. The statistical analysis was conducted with Microsoft Excel or Prism 7.04 (GraphPad). The *P* values were calculated using a two-tailed Student's *t*-test. The figure legends indicate the exact number of measurements, the number of independent experiments and the statistical test used for each analysis performed. Experiments were repeated independently with similar results obtained.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Sequencing data are deposited into ArrayExpress and the accession numbers are E-MTAB-7252 (ChIP-Seq), E-MTAB-7253 (bulk RNA-Seq) and E-MTAB-7254 (scRNA-Seq). Re-analysed previously published data are available under the accession codes ENA PRJNA326944, ENA PRJNA153427, ENA PRJNA291062 and GSE73017. The source data for the figures and supplementary figures are in Supplementary Table 10. All other relevant data are available from the corresponding author on request.

Code availability

The software and algorithms for data analyses used in this study are all well-established from previous work and are referenced throughout the manuscript. No custom code was used in this study.

References

55. Wang, W. et al. Rapid and efficient reprogramming of somatic cells to induced pluripotent stem cells by retinoic acid receptor gamma and liver receptor homolog 1. *Proc. Natl Acad. Sci. USA* **108**, 18283–18288 (2011).
56. Petersen, B. et al. Development and validation of a highly efficient protocol of porcine somatic cloning using preovulatory embryo transfer in peripubertal gilts. *Cloning Stem Cells* **10**, 355–362 (2008).
57. Holker, M. et al. Duration of in vitro maturation of recipient oocytes affects blastocyst development of cloned porcine embryos. *Cloning Stem Cells* **7**, 35–44 (2005).
58. Pedregosa, F. & Varoquaux, G. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
59. Harrow, J. et al. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res.* **22**, 1760–1774 (2012).
60. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
61. Chang, C. W. et al. Identification of human housekeeping genes and tissue-selective genes by microarray meta-analysis. *PloS One* **6**, e22859 (2011).
62. Picelli, S. et al. Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181 (2014).
63. Lee, T. I., Johnstone, S. E. & Young, R. A. Chromatin immunoprecipitation and microarray-based analysis of protein location. *Nat. Protoc.* **1**, 729–748 (2006).
64. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
65. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
66. Kent, W. J. et al. The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).

nature research

Corresponding author(s): Pentao Liu

Last updated by author(s): Mar 29, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	iRODS Version 4.2.2
Data analysis	Microsoft Excel 2016 and GraphPad Prism 6 were used to analyze statistical data and draw graphs in the study. FlowJo V10 was used to analyze the flow cytometry data. Photoshops CS5 was used to crop images from unprocessed gel images. All codes used to process and analyse the data are available at the GitHub repository: https://github.com/dbrg77/pig_and_human_EPSC.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Sequencing data are deposited into ArrayExpress, and the accession numbers are E-MTAB-7252 (ChIP-seq), E-MTAB-7253 (bulk RNA-seq) and E-MTAB-7254 (single cell RNA-seq). Previously published data re-analysed are available under accession code ENA PRJNA326944, ENA PRJNA153427, ENA PRJNA291062, and GSE73017. Source data for Figures and Supplementary Figures are in Supplementary table 10. All other relevant data are available from the corresponding author on request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes were determined based on preliminary experiments and our experience with the specific type of experiment and commonly used sample sizes in comparable publications within this field of research. We independently performed all experiments (including porcine EPSC derivation from preimplantation embryos and porcine/human EPSCs from somatic cells; EPSC differentiation in vitro/vivo, PGCLC derivation; TSC derivation from porcine/human EPSC; RNAseq and ChIPseq) at least 3 times, allowing to judge the degree of variability between replicates. The sample sizes and number of repeats are defined in each figure legends.
Data exclusions	No data exclusion from the analysis.
Replication	Each experiment was repeated independently at least three times and sample sizes and number of repeats are defined in each figure legends. All the experimental findings were reliably reproduced.
Randomization	For pig embryo microinjections, embryos and recipient sows were randomly allocated into experimental groups.
Blinding	The investigators were blinded to group allocation during data collection.

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).*
Research sample	*State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.*
Sampling strategy	*Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.*
Data collection	*Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.*
Timing	*Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Non-participation	*State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.*
Randomization	*If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.*

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.*
Research sample	*Describe the research sample (e.g. a group of tagged Passer domesticus, all Stenocereus thurberi within Organ Pipe Cactus National*

Research sample	*Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.*
Sampling strategy	*Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.*
Data collection	*Describe the data collection procedure, including who recorded the data and how.*
Timing and spatial scale	*Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Reproducibility	*Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.*
Randomization	*Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.*
Blinding	*Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.*

Did the study involve field work? ☐ Yes ☐ No

Field work, collection and transport

Field conditions	*Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).*
Location	*State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).*
Access and import/export	*Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).*
Disturbance	*Describe any disturbance caused by the study and how it was minimized.*

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☐	☐ Palaeontology
☐	☒ Animals and other organisms
☐	☐ Human research participants
☐	☐ Clinical data

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☐	☒ Flow cytometry
☐	☐ MRI-based neuroimaging

Antibodies

Antibodies used	mCherry (Abcam, ab167453, 1:200 for IF), mCherry (LSBio, LS-C204207, 1:200 for IF), OCT4 (SantaCruz,SC-5279, 1:50 for IF), NANOG (Abcam, ab80892,1:50 for IF), SOX2 (R&D, AF2018,1:50 for IF), Tuj (R&D, MAB1195,1:50 for IF), SOX17 (R&D, AF1924,1:200 for IF), GATA3 (R&D, AF2605,1:200 for IF), CDX2(BioGenex, AM392-5M:clone CDX2-88, 1:400 for IF), GATA4 (R&D, AF2606, 1:50 for IF), AFP (R&D, MAB1368,1:100 for IF), A-SMA (R&D, MAB1420,1:100 for IF), TFAP2C (SantaCruz, SC-8977,1:50 for IF), BLIMP1 (eBioscience,14-5963-82,1:100 for IF), KRT7 (SantaCruz, 5F28,1:50 for IF), PL-1 (SantaCruz, SC-34713, 1:100 for IF), TRA-1-60 (STEMCELL, 60064, 1:100 for IF), TRA-1-81 (STEMCELL, 60065, 1:100 for IF), SSEA1 (STEMCELL, 60060, 1:100 for IF), SSEA4 (STEMCELL, 60062,1:100 for IF), SMAD2/3 (Cell signaling, 8685, 1:1000 for western blot), pSMAD2/3 (Cell signaling, 8828,1:1000 for western blot), AXIN2 (Cell signaling, 2151s, 1:1000 for western blot), a-TUBLIN (Abcam, ab7291, 1:1000 for western blot), CGB (Abcam, ab9582,1:200 for IF), SDC1 (Abcam, ab39969,1:200 for IF), H3K4me3 (Diagenode, C15410003-50,1µg/Chip), H3K27me3 (Abcam, ab6002,1µg/Chip), Alexa488-AP (BD Pharmingen, 561495,1:50 for Flow), AF647-CD38 (Biolegend, 303514,1:50 for Flow), PerCpCy5.5-TNAP (BD Pharmingen, 561508,1:50 for Flow), HLA-G (MEM-G/9) (NOVUS,

NB35003-314A647,1:50 for Flow), HLA-ABC (W6/32) (eBioscience, 11-9983-41,1:50 for Flow).

Validation

Antibodies were validated according to manufacturer's instruction.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

Human ES cell lines H1 and H9 were published previously in reference 4. Man1 and Man10 were previously published in reference 31 and 32. GM00013 cells were purchased from Coriell Institute. Germany Landrace and China TAIHU porcine fetal fibroblasts were previously published in reference 20 and 21, respectively. Pig TSC-#1 and TSC-#3 were derived from porcine EPSCs that established from Germany Landrace preimplantation embryos. Mouse feeder cell line STO was previously published in reference 55.

Authentication

Human ES cell lines H1, H9, Man1 and Man10 were obtained directly from the author institutes.
GM00013 cells were directly purchased from Coriell Institute.

Mycoplasma contamination

Mycoplasma tested, all cell lines are negative.

Commonly misidentified lines
(See ICLAC register)

No commonly misidentified cell lines were used.

Palaeontology

Specimen provenance

Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).

Specimen deposition

Indicate where the specimens have been deposited to permit free access by other researchers.

Dating methods

If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

Mice and pigs were used in this study. The mice used were 8-week-old male NSG and SCID mice. The male and female pigs were German Landrace of 7-9 months of age (90-120 kg bodyweight).

Wild animals

No wild animals were used in this study.

Field-collected samples

No field-collected samples were used in this study.

Ethics oversight

The Wellcome Trust Sanger Institute in UK, Niedersaechsisches Landesamt fuer Verbraucherschutz und Lebensmittelsicherheit in Germany and The University of Hong Kong approved the protocols used in this study.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics

Describe the covariate-relevant population characteristics of the human research participants (e.g. age, gender, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."

Recruitment

Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.

Ethics oversight

Identify the organization(s) that approved the study protocol.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	*Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.*
Study protocol	*Note where the full trial protocol can be accessed OR if not available, explain why.*
Data collection	*Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.*
Outcomes	*Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.*

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	For ChIP-seq (E-MTAB-7252): https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-7252/.
Files in database submission	The H3K4me3 and H3K27me3 ChiP-seq data of porcine and human EPSCs
Genome browser session (e.g. UCSC)	Porcine cells ChiP-Seq data: http://genome-euro.ucsc.edu/cgi-bin/hgTracks?hgS_doOtherUser=submit&hgS_otherUserName=dbrg77&hgS_otherUserSessionName=XG2_pEPSC_treat_pileup. Human cells ChiP-Seq data: http://genome-euro.ucsc.edu/cgi-bin/hgTracks?hgS_doOtherUser=submit&hgS_otherUserName=dbrg77&hgS_otherUserSessionName=XG2_hEPSC_treat_pileup.

Methodology

Replicates	For ChIP experiments, 3 biological replicates were performed.
Sequencing depth	50 base pair single end reads were mapped to the UCSC reference genomes (build susScr11 for pig and hg38 for human) using bowtie2 (version 2.3.4) with default setting. For the human reference hg38, all the alternative loci were removed (chr*_alt) before mapping. Reads mapped to the mitochondrial genome were removed, and reads mapped to the nuclear genome were filtered by samtools with flags '-q 30' to filter reads with relatively low mapping quality (MAPQ less than 30). For the ChIP-seq data from human naïve and primed ESCs, raw reads were downloaded from ENA (Study accession PRJNA255308) and processed in the same manner.
Antibodies	Antibodies against H3K4me3 and H3K27me3 were used for ChIP-Seq analysis. Please see Supplementary Table 7 for details.
Peak calling parameters	Peak calling was performed using MACS2 (2.1.1.20160309). For identification of enriched regions of punctate marks (H3K4me3 and H3K27ac) from pig samples, peak calling was performed with flags '-t chip.bam -c input.bam -g 2.7e9 -q 0.01 -f BAM --nomodel --extsize 200 -B --SPMR'. For identification of enriched regions of broad marks (H3K27me3), peak calling was performed with flags '-t chip.bam -c input.bam -g 2.7e9 -q 0.01 -f BAM --nomodel --extsize 200 -B --SPMR --broad'. For human data, peak calling was done in the same way, with a change of genome size '-g hs' during the peak calling. The resulting bedGraph files were converted to bigWig files using the script bdg2bw (https://gist.github.com/jl32587/34370c995460f9d5ad65). The bigWig files were visualised using UCSC genome browser.
Data quality	For H3K4me3 peaks, all samples have more than 20,000 peaks that have FDR<1% and fold enrichment more than 5. For H3K27me3 peaks, they are broad. We visually inspect the signal along the genome browser link provided above. This is the most efficient way of examining the data.
Software	bowtie2-2.2.9, macs2 (v1.1.1.20160309), samtools v1.3, bedtools v2.27.1

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	The half fetuses of day 26-28 chimeras were dissected into small tissue pieces representing several body sections (head, trunk and tail). The dissected tissues and placenta were digested with 1.0 mg/ml collagenase IV (Thermo Fisher Scientific, Cat. No. 17104019) for 1-3 hours at 37 °C on a shaker. A pipette was used to blow the tissue blocks and dissociate them into single cells. The dissociated cells were filtered with a 35 µm nylon mesh (Corning, Cat. No. 352235) to remove tissues clumps. After centrifugation, the cells were fixed using Fixation Medium according to the manufacturers' manual (BD Cytofix, Cat. No. 554655) and the washed cells were stored at 4 °C in PBS supplemented with 0.1% NaN3 (Sigma, Cat. No. 199931) and 5% FBS (Gibco, Cat. No. 10270) before analysed with flow cytometry. All the samples were analysed using BD LSR Fortessa cytometer. 561nm (610/20 bandpass filter) and 488nm (525/50 bandpass filter) channels were used to detect mCherry and excluded autofluorescence.
Instrument	LSR II Fortessa cytometry (Bector Dickinson)
Software	FlowJo V10
Cell population abundance	Purity was computer determined by Summit software. A value of % total between 80-90% was obtained. The % total was the % of cells sorted on the stream out of all sorted cells.
Gating strategy	Because of the various sizes of the dissociated cells analyzed , the majority of starting cells were included in FSC/SSC gates. The boundary between positive and negative is defined according the negative control. Please see an example in Supplementary Figure 3c.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type	*Indicate task or resting state; event-related or block design.*
Design specifications	*Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.*
Behavioral performance measures	*State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).*

Acquisition

Imaging type(s)	*Specify: functional, structural, diffusion, perfusion.*
Field strength	*Specify in Tesla*
Sequence & imaging parameters	*Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.*
Area of acquisition	*State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.*

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software	*Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*
Normalization	*If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.*
Normalization template	*Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.*
Noise and artifact removal	*Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).*
Volume censoring	*Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.*

Statistical modeling & inference

Model type and settings	*Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).*

Effect(s) tested

Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference
(See Eklund et al. 2016)

Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.

Correction

Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).

Graph analysis

Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).

Multivariate modeling and predictive analysis

Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.