Nature. Author manuscript; available in PMC 2017 December 07.

Published in final edited form as:
Nature. 2017 June 15; 546(7658): 416–420. doi:10.1038/nature22812.

Principles of early human development and germ cell program from conserved model systems

Toshihiro Kobayashi[1,2,*], **Haixin Zhang**[4,*], **Walfred W.C. Tang**[1,2,**], **Naoko Irie**[1,2,**], **Sarah Withey**[4,**], **Doris Klisch**[4], **Anastasiya Sybirna**[1,2,3], **Sabine Dietmann**[1,3], **David A. Contreras**[4], **Robert Webb**[4], **Cinzia Allegrucci**[5], **Ramiro Alberio**[4,#], and **M. Azim Surani**[1,2,#]

[1]Wellcome Trust/Cancer Research UK Gurdon Institute, University of Cambridge, Tennis Court Road, Cambridge CB2 1QN, UK

[2]Department of Physiology, Development and Neuroscience, University of Cambridge, Downing Street, Cambridge CB2 3DY, UK

[3]Wellcome Trust Medical Research Council Stem Cell Institute, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK

[4]School of Biosciences, University of Nottingham, LE12 5RD, Loughborough, UK

[5]School of Veterinary Medicine and Sciences, University of Nottingham, LE12 5RD, Loughborough, UK

Abstract

Human primordial germ cells (hPGCs), the precursors of sperm and eggs, originate during week 2-3 of early postimplantation development1. Using *in vitro* models of hPGC induction2–4, recent studies suggest striking mechanistic differences in specification of human and mouse PGCs5. This may partly be due to the divergence in their pluripotency networks, and early postimplantation development6–8. Since early human embryos are inaccessible for direct studies, we considered alternatives, including porcine embryos that, as in humans, develop as bilaminar embryonic discs. Here we show that porcine PGCs (pPGCs) originate from the posterior pre-primitive streak competent epiblast by sequential upregulation of SOX17 and BLIMP1 in response to WNT and BMP signalling. Together with human and monkey *in vitro* models simulating peri-gastrulation

*First authors contributed equally
**Second authors contributed equally
#co-corresponding authors
D.A.C current address CEIEPAA-FMVZ-UNAM, Tequisquiapan, Queretaro, 76790, Mexico

Author contributions

T.K. designed experiments and performed cell culture, plasmid construction, IF, qPCR, RNA-seq, Westerns, data analysis and wrote the paper; W.W.C.T. designed experiments and analysed RNA-seq. N.I. performed preliminary work and designed experiments, and S.D. performed bioinformatics. A.S. helped with a hPSC reporter. H.Z., S.W., D.K. and CA designed and performed IF and culture of pig embryos and epiblasts. D.A.C. and R.W designed and performed in situ hybridizations and IF. R.A. supervised the project, designed experiments, performed dissections and wrote the paper. M.A.S. supervised the project, designed experiments, and wrote the paper. All authors contributed to the manuscript.

Data availability

RNA-seq data have been deposited in Gene Expression Omnibus (GEO) under accession number GSE85378. All other data are included within the paper, source data and Supplementary Information.

Europe PMC Funders Author Manuscripts

development, we show conserved principles for epiblast development for competency for PGC fate, followed by initiation of the epigenetic programme9–11, regulated by a balanced SOX17–BLIMP1 gene dosage. Our combinatorial approach using human, porcine and monkey *in vivo* and *in* vitro models, provides synthetic insights on early human development.

First, we sought the origin of porcine PGCs (pPGCs) in ~E9.5-E16 peri-gastrulating embryos. At ~E9.5–E10, key pluripotency genes NANOG, OCT4 and SOX2 are detected in the epiblast of bilaminar embryos (Fig. 1a). In ~E11 pre-primitive streak (PS) stage embryos with an incipient anterior-posterior axis (Extended Data Fig. 1a), BRACHYURY (T) expression is evident in the posterior pseudo-stratified epiblast cells, together with NANOG and OCT4, but SOX2 is downregulated (Fig. 1b).

In the midline of early-PS stage embryos (~E11.5-E12), we see the first cluster of SOX17 positive (+ve) cells in the posterior end of the nascent PS (arrows in Fig. 1c,d; Extended Data Fig. 1b); most of these express BLIMP1, except for those at the anterior end. Expression of SOX17 precedes BLIMP1; NANOG is retained and upregulated in SOX17/BLIMP1 +ve pPGCs (Fig. 1d; Extended Data Fig. 1b). In ~E12.5-E13.5 embryos, pPGCs exhibit co-expression of SOX17, BLIMP1, NANOG, TFAP2C, OCT4, and pPGC cell surface marker Sda/GM212, but have low levels of T (Fig. 1e, Extended Data Fig. 1c,d). This pPGC cluster of ~60 SOX17/BLIMP1 +ve cells located at the border between embryonic and extraembryonic tissues in early-PS stage embryos (~E12), increases to >300 PGCs by ~E15.5 (Extended Data Fig. 2a-c). A 6-hour (h) pulse of EdU labelling shows that DNA synthesis ceases soon after the detection of Sda/GM2 epitope (Fig. 1f, Extended Data Fig.2d), indicating that the sharp increase in pPGCs is likely due to the additional recruitment from T+ve competent progenitors. Thereafter, pPGCs enter quiescence and pause prior to migration, as in mice13 (Fig. 1f, Extended Data Fig. 2c). Notably, PRDM14 expression in pPGCs is weak and apparently cytoplasmic (Extended Data Fig. 1f), while SOX2 is undetectable (Extended Data Fig. 2e,f).

Initiation of the germline-specific epigenetic program9,14 is evident in nascent pPGCs with a global reduction in 5-methylcytosine (5mC) (Extended Data Fig. 3b,c) and concomitant enrichment of 5-hydroxymethylcytosine (5hmC) (Fig. 1g,h). Consistently, UHRF1 is downregulated (Fig. 1i) and TET1 is upregulated (Extended Data Fig. 3a). Progressive reduction in H3K9me2 and G9a expression is also evident (Extended Data Fig. 3b,c) and global DNA demethylation follows as pPGCs migrate towards the gonads (Extended Data Fig. 3d). Expression of SOX17, BLIMP1, TFAP2C, OCT4 and NANOG continues in pPGCs (arrowheads in Fig. 1e, Extended Data Fig. 2e,f), as seen in equivalent hPGCs *in vivo*4,9. Thus, pPGCs originating from peri-gastrulation porcine epiblast exhibit close similarities with hPGCs following sequential SOX17 and BLIMP1 expression, and the onset of the epigenetic program4,9.

To determine when porcine epiblasts gain competency for pPGC fate, we isolated epiblast discs away from the hypoblast and trophectoderm at different stages of development (Fig. 2a,b,d,e). We exposed them to cytokines, including BMP2 or 43,4(henceforth called Cytokines or Cy) for 64h to induce pPGCs *ex vivo* (Fig. 2f). While no response was seen in ~E10.5-E11 early bilaminar disc epiblasts, an efficient induction of pPGCs occurs in

Europe PMC Funders Author Manuscripts

epiblasts from ~E11.5 pre-PS embryos, with co-expression of SOX17, BLIMP1, NANOG, OCT4 and TFAP2C (Fig. 2g,h, Extended Data Fig. 3e). Note a strong BMP2/4 signal from the posterior pre-streak and early PS pig embryos *in vivo*15,16 (Fig. 2a), with nuclear pSMAD1/5/8 localisation in the posterior epiblast (Fig. 2c). BMP4 inhibitor (LDN193189) abrogates (0/5) pPGC induction (Fig. 2g, Extended Data Fig. 3e). The competency for pPGC fate wanes during early-PS stage concomitantly with the onset of gastrulation and mesendoderm differentiation (Fig. 1c). WNT signalling from the posterior pre-PS epiblast is also apparently important (Fig. 2a), as evidenced by the high proportion of T +ve cells, a downstream target on WNT17, during pPGC induction (Fig. 1b). A WNT inhibitor (WNTi) diminishes pPGC induction, but not in early-PS stage epiblast when they are already competent for pPGC fate (Fig. 2g, Extended Data Fig.3e). Expression of T in response to WNT is a common feature of competency for germ cell fate3,4,17.

Next, since we cannot examine early human embryos directly, we decided to simulate human peri-gastrulation development and induce hPGC fate *in vitro*. We established an *in vitro* model with human pluripotent stem cells (hPSC) for mesendoderm (ME) and peri-gastrulation development18–20 (Fig. 3a). These hPSCs in a defined conventional medium (henceforth; Conv-hPSCs, see Methods), are probably equivalent to the pre-gastrulating porcine epiblast21. To follow gain and loss of competency for hPGCs specification in the course of ME differentiation, we used cells with a sensitive *NANOS3-tdTomato* (NT) reporter (see Extended Data Fig. 4). By checking for hPGCs induction by Cytokines at 6h intervals (Fig. 3a), we observed a peak of competency at ~12h during ME differentiation, declining thereafter (Fig. 3b,d), illustrating a step-wise transition in cell states. PGC-competent cells show moderate upregulation of PS markers; T, MIXL, GSC, EVX1 and EOMES, and a slight reduction of SOX2 (Fig. 3e, Extended Data Fig. 5a, b). At 18h onwards, there is upregulation of late PS markers; TBX6, MESP1/2, and *SNAIL* (SNAI1/2), which triggers epithelial-mesenchymal transition, occurring during gastrulation22 (Extended Data Fig. 5b). The response of cells to signalling also changes, since BMP2/4 now induces mesodermal cells (Extended Data Fig. 5b, e-h), while Activin A and a BMP inhibitor efficiently induces definitive endoderm (DE)18 (Fig. 3c,d. Extended Data Fig. 5e-h). Thus, from ~12h, onwards, the ME precursors (or Pre-ME) are competent for PGC fate, but from ~18h (ME), competency for PGC declines, with concomitant gain of competency for DE/mesoderm fates (Fig. 3a-c).

Both WNT and ACTIVIN/NODAL signalling are necessary for competency for hPGCs in 12h Pre-ME (Extended Data Fig. 5c). Inhibition of BMP during ME differentiation dramatically reduces the efficiency of hPGC induction, suggesting a role for endogenous BMP (Extended Data Fig. 5d). Addition of BMP during Pre-ME(12h) however does not affect hPGC competency, but it favours differentiation into lateral mesoderm (LM) at 24h, which originates from mid/posterior PS (Extended Data Fig. 5e-h). By contrast, DE differentiation occurs in the anterior PS as seen in PS-stage porcine embryos (arrowheads in Fig 1d, Extended Data Fig.1c), suggesting that BMP confers posterior characteristics to PS during gastrulation (Extended Data Fig. 5i). The combined evidence from porcine embryos, and simulations in *in vitro* with Conv-hPSCs, shows progressive and transient acquisition of competency for hPGC, followed by competency for DE/mesoderm.

To test if this model fapplies to non-human primates, we used *Cynomolgus* monkey PSCs (cmPSCs) to induce ME differentiation (Extended Data Fig. 6a,b). Remarkably, cmPSCs also show a temporal gain of competency for cmPGC fate at ~12h during ME differentiation, and thereafter DE and mesoderm at 24h (Extended Data Fig. 6c-f). A recent study suggests that cmPGCs originate from the epiblast derived nascent amnion, but also potentially from the posterior epiblast23; a dual origin is however also possible. A detailed molecular analysis of amnion is important, since SOX17 and BLIMP1 also have roles elsewhere, including extraembryonic tissues24,25, which exhibit great developmental diversity and less developmental constrain than epiblast26. Further traceable analysis of *in vivo* non-human primate embryo, and potentially of human embryos *in vitro* is merited.

Next, we examined the combinatorial roles of the key transcription factors considering that SOX17 is the key regulator of both hPGC and DE (Fig. 3d)4,26. Ectopic expression of SOX17 in Pre-ME at 12h induces hPGCs resulting in NT/OCT4/BLIMP1+ve cells, but in ME at 24h it induces FOXA2+ve (DE) efficiently (Fig. 3f,g). During the induction of hPGCs in Pre-ME by BMP (Extended Data Fig. 7a,b), we first detect SOX17 at 12h, of which ~30-40% show co-expression with BLIMP1, but not TFAP2C (Extended Data Fig. 7c,d), as in the pig embryo (Extended Data Fig. 1e). TFAP2C is detected at ~18h, as the number of SOX17/BLIMP1/TFAP2c+ve cells increases progressively. These putative hPGCs eventually form a cluster in the middle of embryoids by 48h (Extended Data Fig. 7b,c,e). At 12h onwards NANOG expression increases progressively in SOX17/BLIMP1+ve cells (Extended Data Fig. 7e,f). However, unlike in mouse27, NANOG alone cannot induce hPGCs (Extended Data Fig. 8a-c), and PRDM14 expression appears weak2,4, and might be cytoplasmic in E14 pPGCs (Extended Data Fig. 1f) in hPGCs *in vivo*5. PRDM14 is critical for mouse PGC fate and epigenetic reprogramming28, but its role, if any, in hPGCs and pPGCs requires further studies.

Next, we tested SOX17, BLIMP1 and TFAP2C individually and in combination for the induction of hPGCs in the reversibly PGC-competent hPSCs (Comp-hPSC)4, with the NT reporter and inducible *SOX17*, *BLIMP1* and *TFAP2C* transgenes (Extended Data Fig. 4c-j, see Methods, and Extended Data Fig. 8d-g). BLIMP1 and TFAP2C individually and together elicit a negligible response after 2 days, while SOX17 alone produces a modest response, with or without TFAP2C (Fig. 4a). Remarkably, SOX17 with BLIMP1 produce a strong response, and a large proportion of NT-positive cells (Fig. 4a), suggesting that they act synergistically and rapidly (within ~24h), compared to 96h for Cytokines to induce a similar response (Fig. 4b,c). This response is preceded by downregulation of SOX229 and upregulation of ‘naïve’ pluripotency genes including KLF4 and TFCP2L1, as in hPGCs *in vivo*9 (Fig. 4e, Extended Data Fig. 8h). The response to SOX17-BLIMP1 is otherwise similar to that induced by Cytokines as judged by global RNA-seq (Fig. 4d, Extended Data Fig. 8i,j).

For further mechanistic insights, we induced ectopic *SOX17* with or without *BLIMP1* expression in *SOX17* null Comp-hPSC4 (Extended Data Fig. 9a, Extended Data Fig. 10a,b). Although *SOX17* initiates hPGC fate, there is also a significant increase in FOXA2 expression, an endodermal gene (Fig. 4f). Accordingly, FACS-purified cells show FOXA1, FOXA2 and HNF1β expression, alongside PGC markers, OCT4, NANOG, BLIMP1 and

TFAP2C (Extended Data Fig. 9b), and the cell surface marker CXCR4 shared by both endoderm19 and PGC in humans (Extended Data Fig. 9c,d). This likely occurs due to high ectopic SOX17 gene dosage exceeding the levels induced by BMP in wildtype Comp-hPSC (Extended Data Fig. 9b, e-h), without a proportionate increase in BLIMP1 (Extended Data Fig. 9i). Notably, BLIMP1 represses endodermal genes during hPGC specification9. Indeed, concomitantly high BLIMP1 and SOX17 induce hPGCs robustly and rapidly while suppressing endoderm genes (Fig. 4g, Extended Data Fig. 10c), and also initiate the germline-specific epigenetic program9, with downregulation of DNMT3A, DNMT3B and UHRF1, and upregulation of TET2 (Extended Data Fig. 8k), as in pPGCs (Extended Data Fig. 3b,d).

The *ex vivo* peri-gastrulation porcine embryos, and *in vitro* mimics for human and monkey development exemplify conserved principles for PGC specification in the context of epiblast development in flat disc embryos (Fig. 4h). We show how SOX17 likely regulates both PGCs and DE. We propose that signalling followed by epigenetic priming of regulatory elements might confer competency for PGC fate, and thereafter DE/mesoderm, as an integral part of epiblast development program towards gastrulation. Remarkably, SOX17-BLIMP1 are necessary and sufficient for inducing PGCs, and for initiating the germline-specific epigenetic program. Once specified, PGCs are irreversibly committed to the germline fate (Extended Data Fig. 2e,f, Extended Data Fig. 10d,e). Notably, our integrated approach provides insights on early human development and cell fate decisions.

Methods

Porcine embryo collection

All the procedures involving animals have been approved by the School of Biosciences Ethics Review Committee, The University of Nottingham. Embryos were collected from crossbred Large White and Landrace sows (2-3 years old) between days 10 and 16 after artificial insemination (AI). Embryos from days 10-16 (embryonic day (E) 10-16; (bilaminar (n=4); Pre-PS (n=7), Early-PS (n=6); PS (n=10), Late-PS (n=13), ~E14 (n=8), ~E16 (n=10)) were flushed from the uterine horns with 30-40 ml warm PBS (supplemented with 1% FCS). Embryos older than E15 were retrieved by cutting the uterus longitudinally from the antimesometrial side and picked with forceps. Embryos at each stage were retrieved from multiple sows. All embryos were grouped by stages and washed with DMEM-F12 (Thermo Fisher Scientific) + 20% knockout serum replacement (KSR: Thermo Fisher Scientific) supplemented with 25 mM HEPES (wash medium, WM) and transported to the laboratory in a portable incubator at 38.5°C.

Immunohistochemistry of porcine embryos and gonads

Embryos were processed as previously described30. Briefly, embryos and gonads were fixed in 2.5% paraformaldehyde (PFA) in PBS overnight (ON) at 4ºC and transferred to methanol at -20°C for long term storage. For less advanced embryos (E9.5-10.5) 2% agarose blocks were made before embedding in wax. Agarose blocks were then trimmed and dehydrated with increasing concentrations of ethanol (70%, 90% and 100%) for 2 h each step. After dehydration, blocks were processed in xylene followed by hot paraffin wax for 2 h and then

cooled down in a metal mold. E14-15 embryos and gonads were dehydrated, embedded and processed as described above. All paraffin blocks were cut into 4 μm thick sections and mounted onto SuperFrost plus slides. Sections were air dried ON and immunohistochemistry was conducted subsequently. For cryo-sections, fixed embryos were incubated in 30% sucrose/PBS ON at 4ºC prior to mounting in OCT compound. Cryo-sections were cut at 6 μm onto Superfrost plus glass slides. Sections were left to air dry for 1-2 h before IF. Unused slides were kept at -80°C in air tight containers.

Immunofluorescence staining of porcine embryos

Sections were dewaxed with xylene for 30 min and rehydrated with decreasing concentrations of ethanol (100%, 90% and 70%) and transferred to PBS for 15 min. Antigen retrieval was then performed by boiling the slides in 0.01M Citrate Buffer (PH 6.0) for 10 min. Sections were permeabilized with 1% Triton-X100 in PBS for 15 min. Triton-X100 was washed three times for 5 min each, and blocking solution (PBS supplemented with 5% BSA or 10% Donkey serum/5% BSA/PBS) was added for 1-1.5 h. After blocking, sections were incubated with the desired primary antibody (see Supplementary Table 2) ON at 4°C in a humidified chamber. Slides were then washed three times with 0.1% Tween-20/PBS. Finally slides were incubated with fluorescent (TRITC, Cy3 Alexa Fluorophore 488, 555, 568 and/or 647; eBioscience, Abcam, Thermo Fisher Scientific)-conjugated secondary antibodies for 40 min at room temperature (RT). Slides were mounted with Fluoroshield with DAPI (Sigma) and sealed with nail varnish. Slides were kept at -20 °C until observed.

Porcine epiblast cultures and quantification of PGC induction

Embryos collected 11.5 days post AI were dissected under a dissecting microscope as described previously31. Trophectoderm and primitive endoderm were carefully removed with forceps and pure epiblasts were placed in pre-incubated in GK15 medium supplemented with individual inhibitors (i), including 10 μM WNTi (IWR-1-endo, Selleckchem), and 100 nM BMPi (LDN193189, Sigma,). GK15 medium is composed of GMEM (Thermo Fisher Scientific) supplemented with 15% KSR, 0.1 mM 2-mercaptoethanol (Thermo Fisher Scientific), 0.1 mM NEAA (Thermo Fisher Scientific), 1 mM sodium pyruvate (Sigma), 100 U/ml Penicillin- 0.1 mg/ml Streptomycin (Sigma), and 2 mM L-Glutamine (Sigma). After 30 min incubation with the inhibitors, cytokines were supplemented to the final concentration of 500 ng/ml BMP4 (R&D systems), 1mg/ml hLIF (Peprotech) and 200 ng/ml SCF (R&D systems). Epiblasts collected from two independent batches of animals and from 4 animals each time were grouped by stage of development (bilaminar, Pre-PS, Early-PS) and randomly distributed into three treatment groups. After 64 h incubation in a low oxygen incubator at 38.5ºC they were fixed in 2.5% PFA at 4ºC ON. After washing in PBS they were incubated in 30% sucrose/0.0025% sodium azide at 4ºC ON and then mounted in OCT compound as described above. IF was carried out with the antibodies listed in Supplementary Table 2 with fluorescent-conjugated secondary antibodies.

All immunostained sections of single epiblasts were imaged at 100x magnification at antibody-specific wavelengths. NANOG-positive cells could be split into two categories; high expression and lower expression, based on immunofluorescence intensity. Image

acquisition was performed using SimplePCI capture software and the Cell Counter application of Fiji was used to mark protein of interest and NANOG-high cells. Separate fluorophore images were overlaid and cells marked as positive for all three proteins of interest were counted. Only robustly stained cells were considered positive. No statistical methods were used to predetermine sample size. Mann-Whitney test was used to test for significance.

Wholemount *In situ* hybridization

RNA *in situ* hybridization (ISH) was carried out as previously described15. Briefly, embryos (NP (n=2); 3-5S (n=3); 6-8S (n=3)) were rehydrated in decreasing concentrations of methanol and then washed in PBS-T (PBS + 0.1% Tween-20). Next, embryos were equilibrated in (1:1) PBST:Hybridisation buffer (HB: Formamide: 50%, 2xSSC (pH5), EDTA (5mM, pH8), 0.05 mg/ml Yeast RNA, 0.2% Tween20, 0.5% CHAPS, 0.1 mg/ml Heparin) for at least 10 min, and then transferred to HB before incubating at hybridisation temperature (HT) for a minimum of 2 h. After incubation, HB containing pig OCT4 probe was added and incubated ON. The following day the probe was removed and embryos were washed 3-5 times with wash buffer (WB: 50% Formamide, 1xSSC (pH5), 0.1% Tween20) at HT, and then equilibrated with (1:1) WB:MABT solution (MABT solution: 100 mM Maleic acid, 150 mM NaCl, 0.005% Tween 20) before washing with MABT solution at RT. The embryos were then blocked with MABT and of 2% blocking reagent (Roche) for 1 h followed by blocking solution with MABT, 2% blocking reagent and 10% normal goat serum for at least 2 h. The blocking solution was then replaced with a 1:2000 dilution of anti-Dig-AP Fab fragments (Roche) and incubated overnight with gentle agitation at 4ºC. Embryos were rinsed in MABT several times before the development step with NBT/BCIP. After color reaction embryos were washed with 5xTBST (TBST: 0.7 M NaCl, 0.01 M KCl, 0.125 M Tris (pH7.5), 0.5% Tween20) solution. The colour reaction was repeated until signal was detected. Once the colour reaction was satisfactory, the embryos were re-fixed with 4% PFA for 1 h, rinsed in PBST and observed under a microscope.

EdU labelling

Embryos collected (n=7) were rinsed in WM before 6 h incubation in 10 uM EdU at 38.5ºC. After rinsing the embryos were fixed in 4% PFA for 2 h and then transferred to 1% PFA for long term storage. Embryos were processed for immunohistochemistry as described above, except that the EdU staining protocol was performed following manufacturer instruction (Click-iT EdU, Invitrogen) prior to addition of primary antibodies used for co-labelling.

Cell culture

NANOS3-mCherry hPSC, SOX17KO hPSC, SOX17KO+iS hPSC lines were established previously4. NANOS3-tdTomato hPSC line was newly generated by gene targeting (See Vector construction and gene Introduction). A cynomolgus monkey PSC (cmPSC) line (MF12 embryonic stem cells (ESCs)) is kindly gifted from E. Curnow, Washington National Primate Research Centre.

Undifferentiated Conv-hPSC were maintained on vitronectin (Thermo Fisher Scientific)-coated plate in Essential 8 medium32 (Thermo Fisher Scientific) according to

manufacturer's protocol. Cells were passed every 3-5 days using 0.5 mM EDTA/PBS without breaking cell clumps.

Undifferentiated Comp-hPSC were maintained on Mitomycin-C treated or irradiated mouse embryonic fibroblast (MEF) (purchased from MTI-GlobalStem or prepared in house) in knockout DMEM (Thermo Fisher Scientific) supplemented with 20% KSR, 0.1 mM NEAA, 0.1 mM 2-mercaptoethanol, 100 U/ml Penicillin- 0.1 mg/ml Streptomycin, 2 mM L-Glutamine, 20 ng/ml human LIF (Stem Cell Institute, university of Cambridge (SCI)), 8 ng/ml bFGF (SCI), 1 ng/ml TGFβ1 (Peprotech), 3 μM GSK3i (CHIR99021,Miltenyi Biotec), 1 μM ERKi (PD0325901,Miltenyi Biotec), 5 μM p38i (SB203580,TOCRIS bioscience), and 5 μM JNKi (SP600125, TOCRIS bioscience), as reported4,33. Cells were passed every 2-4 days using TrypLE express or 0.25% trypsin/EDTA (Thermo Fisher Scientific). Before harvesting Comp-PSC on MEF, 10 μM of ROCKi (Y-27632, TOCRIS bioscience) was added into the medium.

Undifferentiated cmPSC were maintained on Mitomycin-C treated or irradiated MEF in DMEM/F-12 (Thermo Fisher Scientific) supplemented with 20% KSR, 0.1 mM NEAA, 0.1 mM 2-mercaptoethanol, 100 U/ml Penicillin- 0.1 mg/ml Streptomycin, 2 mM L-Glutamine, 20 ng/ml bFGF (SCI). Before differentiation, cmPSC were maintained on MEF in Essential 8 medium supplemented with 5% KSR, 2.5 μM WNTi (IWR-1; Tocris) at least for 2 passage. Cells were passed every 3-4 days using 0.25% trypsin/EDTA (Thermo Fisher Scientific). Before harvesting cmPSC on MEF, 10 μM of ROCKi (Y-27632, TOCRIS bioscience) was added into the medium.

For ME induction and subsequent DE/LM induction, we optimized published protocol18,34,35 with slight modification. For the basal medium, we used aRB27 medium which is composed of Advanced RPMI 1640 Medium (Thermo Fisher Scientific) supplemented with 1% B27 supplement (Thermo Fisher Scientific), 0.1 mM NEAA, 100 U/ml Penicillin- 0.1 mg/ml Streptomycin, 2 mM L-Glutamine. For ME induction, trypsinized Conv-hPSC were seeded on vitronectin-coated dish at 200,000 cells/well in 12 well plate and cultured in ME induction medium for 6-36 h as indicated in the figures. ME induction medium was composed of aRB27 medium supplemental with 100 ng/ml Activin-A (SCI), 3 μM GSK3i and 10 μM of ROCKi. For DE induction, ME induction medium was replaced to DE induction medium after washing with PBS once and cells were cultured for further 2-3 days. DE induction medium was composed of aRB27 medium supplemented with 100 ng/ml Activin-A and 0.5 μM BMPi (sigma). For LM induction, ME induction medium was replaced to LM induction medium after washing with PBS once and cells were cultured for further 1 day. LM induction medium was composed of aRB27 medium supplemented with 100 ng/ml BMP2 (SCI), 2.5 μM WNTi (IWR-1) and 10 μM TGFβi (SB 431542; Tocris).

Embryoids were formed according to previous report4. Briefly, Conv- or Comp-hPSCs, Pre-ME or ME were trypsinized into single cells and harvested into Corning® Costar® Ultra-Low attachment multiwell 96 well plate (Sigma) or PrimeSurface 96V plate (S-BIO) at 4000 cells/well in GK15 medium or aRB27 medium. To improve the cell aggregation, in some experiments, we added 0.25% (v/v) poly-vinyl alcohol (sigma) in the basal medium,

according to published protocol36. For hPGC induction with Cytokines, 500 ng/ml BMP4 or BMP2 (SCI), 10 ng/ml human LIF (SCI), 100 ng/ml SCF, 50 ng/ml EGF (R&D Systems), and 10 μM ROCKi were added to GK15 medium or aRB27 medium. Notably, an use of aRB27 medium for the basal medium improved the efficiency of hPGC induction, especially, when we used the Pre-ME cells (see Extended Data Fig. 4i,j). A detailed protocol of the *in vitro* differentiation system described above is also available at *Protocol Exchange*37.

For induction of exogenous transgenes, 100 uM Dexamethasone (Sigma), 0.5 ug/ml Doxycycline (Sigma) and/or 0.5 uM Shiled-1 (Clontech) were added to the medium.

Vector construction and gene introduction

For construction of *NANOS3-tdTomato* knock-in targeting vector, 5' and 3' arms amplified from *NANOS3-mCherry* knock-in targeting vector4, *rox-PGK-PuroΔtk-rox* amplified from *pL1L2-PGKpuroΔtk* (a kind gift from SCI), *MC1*-promoter driven *DTA* cassette amplified from *pMC1-DTA*, were cloned into *pBluescript KS(+)* (Stratagene) with an in-fusion HD cloning kit (Takara Bio). For efficient gene targeting, we used the CRISPR/Cas9 system. Guide RNAs targeting around the stop codon sequence of NANOS3 genes (5'-caccgAGCCTCCTAGGTGGACATGG-3', 5'-aaacCCATGTCCACCTAGGAGGCTc-3') were cloned into pX330 (Addgene) or eSpCas9(1.1) (Addgene). For construction of *pCAG-Dre-IRES-Hygromycin*, *puromycin resistance gene* in *pCAGGS-Dre-IRES-Puro* was replaced to *hygromycin resistance gene*. For construction of Dox/Shiled-1 inducible system, *SOX17* cDNA without stop codon amplified from cDNA of hPGC and *Destabilized-domain* amplified from pBMN YFP-FKBP (Addgene) were cloned into PiggyBAC vector used previously38. For construction of Dox-inducible system, *Tet3G* amplified from *pCMV-Tet3G* (Clonetech) was cloned into PiggyBAC vector which contains CAG-promoter and IRES- neomycin-resistance gene. *TRE3G* cassette amplified from *pTRE3G* (Clonetech), human *BLIMP1* cDNA amplified from TRE-Tight BLIMP1-EGFP38, human *NANOG* cDNA amplified from *pEP4-E02S-CK2M-EN2L* (Addgene) and human *TFAP2C* cDNA amplified from hPGC were cloned into PiggyBAC vector. All fragments were amplified by PCR using PrimeSTAR MAX or PrimeSTAR GXL DNA polymerase (Takara Bio) according to the manufacturer's protocol.

For gene targeting of NANOS3-tdTomato reporter, WIS2 male hESC and H9 female hESC lines were used. For the gene targeting method, electroporation or lipofection was carried out as described before4. In brief, for electroporation, 1~5 x 10e6 Comp-hPSCs suspended in PBS were mixed with targeting vector and CRISPR/Cas9 plasmid was transferred to a Gene Pulser cuvette (Bio-Rad). Electroporation was carried out using Gene Pulser equipment (Bio-Rad). For lipofection, reverse transfection was carried out. 2 x 10e5 Comp-hPSCs were suspended in 100-200 μl of Opti-MEM containing targeting vector, CRISPR/Cas9 plasmid and lipofectamine 2000 (Thermo scientific) complex, and left them for 5 min at room temperature. After electroporation or lipofection, PSCs were seeded onto 4 drug resistant (DR4) MEF (GlobalStem or SCI) and 48 h later, 0.5ug/ml puromycine (Sigma) was added to the culture medium for selection. After the selection, puromycin resistant hPSC colonies were picked up and judged correct targeting by PCR using following primers

(Primer F1; 5'-GGGGCCAGTCTAACTAGGTGTG-3', Primer R1; 5'-GGTCTTCCTGGAAATCCAGCCG-3', Primer R2; 5'-TGCCGGTGCCATGCCCCAGGAACA-3' see Extended Data Fig. 4a, b). The targeted clones were expanded and then used for excision of *Rox*-flanked *PGK-PuroΔtk* by transient transfection of *pCAG-Dre-IH.* After selection with 25 ug/ml hygromycin B and subsequently with 0.2 uM FIAU, colonies were picked up and judged the excision by adding puromycin or PCR using following primers (Primer F2; 5'-CCGTCCCATGCACGTCTTTATC-3', Primer F3; 5'-GAGGGCAGAGGAAGTCTGCTAACA-3', Primer R3; 5'-GGTCTTCCTGGAAATCCAGCCG-3' see Extended Data Fig. 4a, b). In the representative clone derived from Wis2 hESC used in this study, G-banding karyotype analysis was performed at Medical Genetics Laboratories, Cambridge University Hospitals NHS Foundation Trust following their standard protocols, and the normal karyotypes were confirmed (20 out of 20 cells).. All clones showed correlation of NANOS3-tdTomato and TNAP expression by FACS analysis (representatives are shown in Exteneded Data Fig. 4d-f). We also confirmed NANOS3-tdTomato +ve cells are specifically expressing the other PGC markers by immunostaining (Exteneded Data Fig. 4g). Notably, fluorescence of NANOS-tdTomato was brighter than NANOS3-mCherry which we made in previous report4 (Exteneded Data Fig. 4h).

For gene introduction using PiggyBAC system, reverse transfection was carried out. 2 x 10e5 SOX17 KO Comp-hPSCs, NANOS3-tdTomato Comp-hPSCs were suspended in 100-200 μl of Opti-MEM containing plasmids and lipofectamine 2000 (Thermo scientific) complex, and left them for 5 min at room temperature. The cells were seeded on DR4 MEF at several different concentrations, and then 48 h later, drugs for selection were added to the culture medium.

Quantitative reverse transcription PCR (RT-qPCR)

Total RNA was extracted using PicoPure® RNA Isolation Kit (Thermo Fisher Scientific) and cDNA was synthesized using QuantiTect Reverse Transcription Kit (QIAGEN) according to manufacturer's protocols. RT-qPCR were performed and analyzed as described previously39 and the primers sequences used in the paper are listed in Supplementary Table 3. Values normalized to GAPDH and relative to control samples (Comp-hPSC or hPGC) are shown in the figures. Error bars are mean ± SD from technical triplicate with two experiments.

Immunofluorescence staining of cells and cryo-section of embryoids

Cells were cultured on μ-Slide 8 well (ibidi) and fixed in 4% PFA for 10 min at room temperature. Embroids were fixed in 4% PFA for 2-4 h or ON at 4°C and embedded in OCT compound for frozen sections. Each sample was incubated with primary antibodies for 1-2 h at RT or ON at cold room and with fluorescent-conjugated secondary antibodies for 1 h at RT. Antibodies used here are listed on Supplementary Table 2. After antibody treatment, samples were stained with DAPI (Sigma) to mark nuclei and were observed under confocal laser scanning microscopy.

Europe PMC Funders Author Manuscripts

Flow cytometry analysis

PSCs, Embryoids, DE and LM were trypsinized with 0.25% trypsin/EDTA at 37°C for 5-15 min and were stained with Alexa-488 or 647 conjugated anti-TNAP antibody (BD Bioscience), PerCP-Cy5.5 conjugated anti-CXCR4 antibody (Biolegend), PE-Cy7 conjugated anti-PDGFRa antibody (Biolegend) and/or Alexa-647 conjugated anti-CD38 antibody and subjected to FACS LSR Frotessa cytometry (BD Bioscience). FACS data were analyzed by Flowjo software.

Western blot analysis

Western blot analysis was carried out as described before4. Briefly, whole-cell extracts were prepared by day 4 embryoids in lysis buffer composed of 50mM Tris-HCl (pH7.5), 0.15M NaCl, 0.1% SDS, 1% Triton X-100, 1% Sodium deoxycholate and cOmplete mini EDTA free (Roche Applied Science, Penzberg, Germany). After electrophoresis, proteins were transferred to nitrocellulose membranes. Membranes were incubated in Western Blocking Reagent (Roche Applied Science) and treated with antibodies. Primary antibodies against SOX17 (goat IgG; R&D systems, see Supplementary Table 2), and H3 (rabbit IgG; Abcam ab1791) were used. Horseradish peroxidase-conjugated secondary antibodies against goat or mouse IgG were added (Dako, Life technologies). After antibody treatment, blots were developed using ECL Western Blotting Detection System (GE Healthcare).

Preparation of RNA-sequencing (RNA-seq) libraries

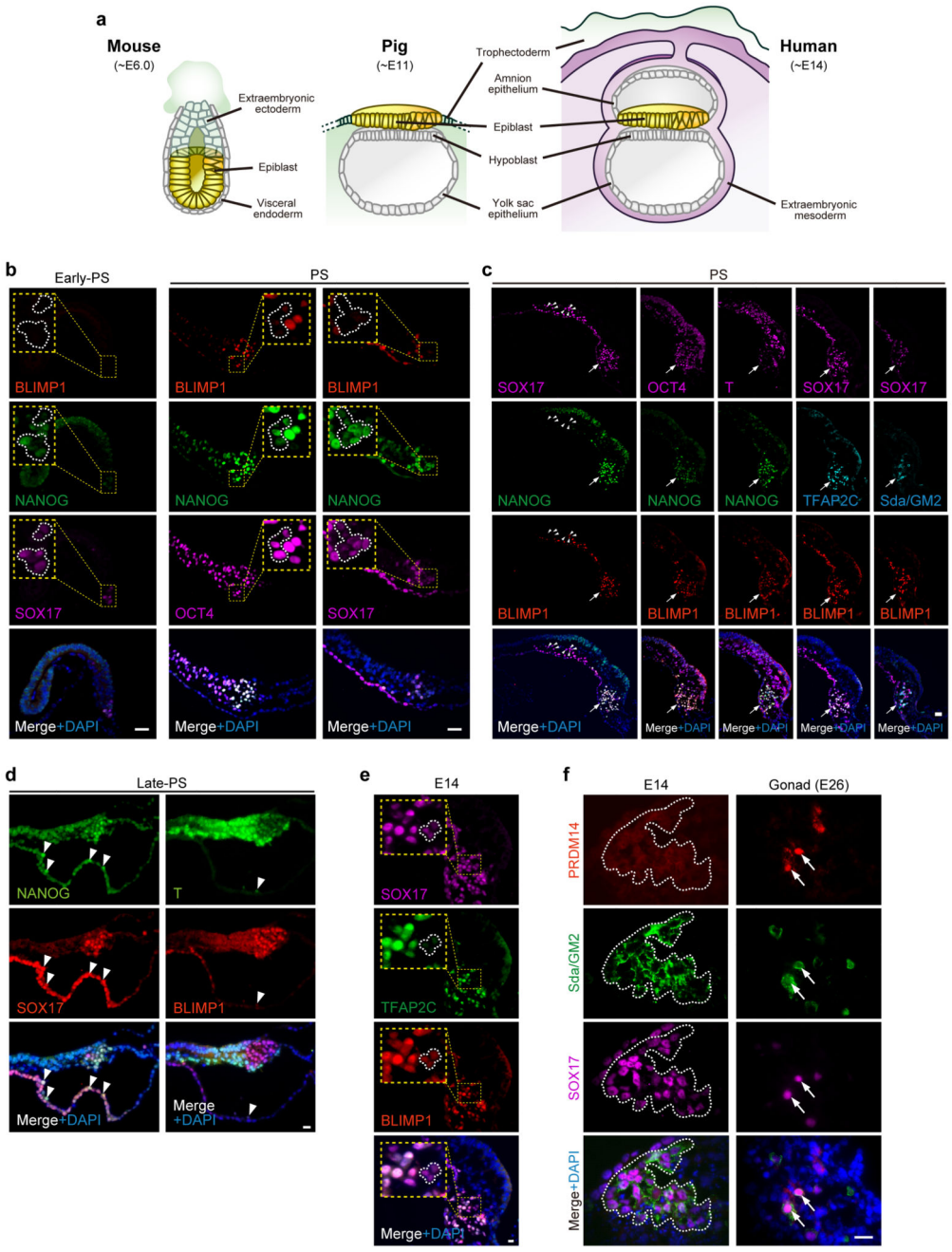
Total RNA (5 ng) was reverse transcribed and amplified into cDNA using Ovation RNA-Seq System V2 (Nugen). Amplified cDNA was sonicated into 250 bp by Covaris S220 Focused-ultrasonicators. Subsequently, RNASeq library was generated with 500 ng of fragmented cDNA using Ovation Rapid DR Multiplex System (Nugen). Library was quantified by qPCR using KAPA Library Quantification Kit (Kapa Biosystems). Libraries were subjected to single-end 50 bp sequencing on HiSeq 2500 sequencing system (Illumina). Every 4 indexed libraries were multiplexed to one lane of a flowcell, resulting in >40 millions single end reads per sample

Bioinfomatics analysis

Adapter-and quality-trimmed RNA-seq reads were mapped to the human reference genomes (UCSC GRCh37/hg19) using *TopHat2* (http://ccb.jhu.edu/software/tophat, version: 2.0.13) guided by ENSEMBL 83 gene models. Raw counts per transcripts were obtained using featureCounts, only the longest transcript per gene was kept. Replicates were evaluated, raw counts were normalized, and the differential expression of transcripts was statistically evaluated by the R Bioconductor *DESeq2* package (www.bioconductor.org). Expression-normalized transcript counts were further normalized by transcript length (per kB). Transcript annotations in all bioinformatics analyses were based on ENSEMBL (Release 83) considering protein coding, long-noncoding RNA and processed transcripts. Hierarchical clustering was performed with the *R hclust* function using the Ward's method. Principal components were computed by singular value decomposition with the R *princomp* functioned on scaled DESeq-normalized expression levels. Only the 80% most highly expressed transcripts were used for clustering and principal component analysis. t-statistic

Stochastic Neighbor Embedding (t-SNE) analysis was performed using R Rtsne package with default parameters and “perplexity = 3”. Gene set enrichment analysis (GSEA) was performed with the R phenoTest package on RNAseq data (log2(normalized counts)). hPGC-specific genes were defined as follows: (1) upregulated in week 7 male hPGC over gonadal somatic cells (log2(fold change) >3 and adjusted p value <0.05); and (2) upregulated in d4 hPGC+Cy over comp-hPSC in our previous study4; and (3) upregulated in week 7 male hPGC over comp-hPSC.

Extended Data



Extended Data Fig.1. Expression of key germ cell genes in early pPGCs.

a. Representation of a mouse, pig and human embryos before gastrulation

b. Sections of Early-PS and PS stage embryo showing SOX17, BLIMP1, NANOG and OCT4. Yellow dashed insets show cells at high magnification and white dashed lines mark SOX17 +ve and/or BLIMP1 -ve cells. Scale bar: 20 μm.

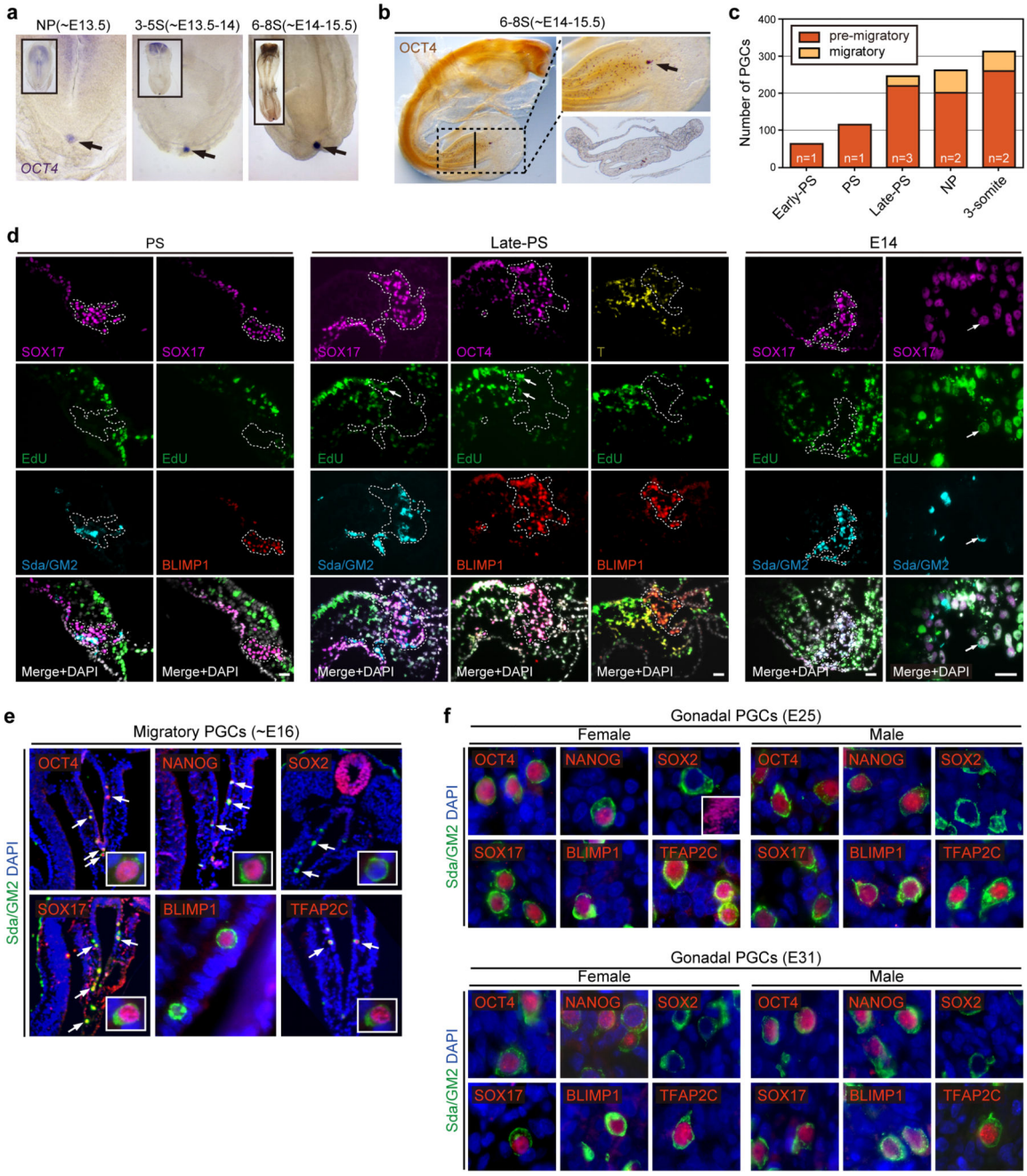
Europe PMC Funders Author Manuscripts

c. PS (~E12) stage embryo with a cluster of pPGCs (arrow) with multiple combinations of PGC gene expression (SOX17, BLIMP1, NANOG, TFAP2C, OCT-4, Sda/GM2 and mesoderm gene, T). Arrowheads at the anterior streak point to primitive endoderm (SOX17/BLIMP1+ve and NANOG negative cells. Scale bar: 20 μm.

d. Late-PS; (~E12.5-E13.5) embryo with a pPGC cluster (arrow) showing NANOG, SOX17 (split color image of Fig.1e), BLIMP1, and T expression. Arrowheads mark early migratory PGCs in the primitive endoderm. Scale bar: 25 μm.

e. E14 embryo stained for SOX17, BLIMP1, and TFAP2C. Yellow dashed insets show cells at high magnification and white dashed lines mark SOX17/BLIMP1 +ve and TFAP2C -ve cells. Scale bar: 20 μm.

f. Immunostaining for PRDM14 co stained with Sda/GM2 and SOX17 in E14 (pPGC cluster) embryos and E26 gonads. Arrows point to pPGCs in the gonad. Scale bar: Scale bar: 20 μm.



Extended Data Fig.2. Proliferation and development of early pPGCs.

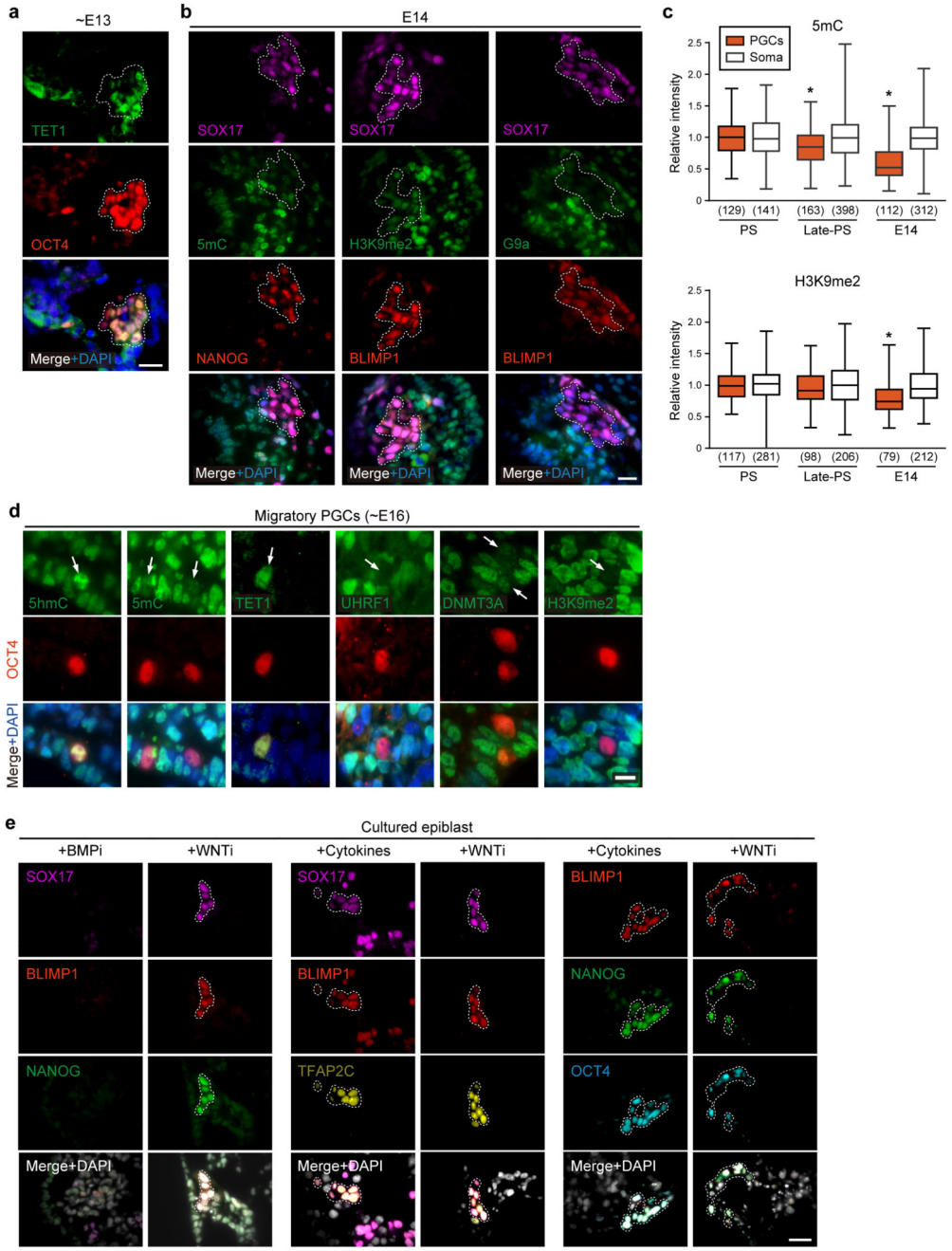
a. *OCT4 RNA in situ* hybridization identifies the pPGC cluster (arrow) in the posterior end of ~E13.5-E15.5 embryos. Insets show whole embryos.

b. Wholemount OCT4 IHC of a porcine embryo. Dashed square marks the area shown at higher magnification on the top right. Arrow points to the pPGC cluster. Bottom right: Cross section of the embryo (line in the wholemount image) shows migratory pPGCs (red cells).

c. Number of pPGCs at different stages as indicated.

Europe PMC Funders Author Manuscripts

d. Immunostaining of EdU labelled embryos at the indicated stages with different antibody combinations identifying the PGCs. The pPGC cluster is highlighted with dashed white line. Arrows show SOX17/EdU +ve and SOX17/BLIMP1/EdU +ve cells. Scale Bar: 20 μm.
e. Immunostained migratory pPGCs (arrows); inset show cells at higher magnification.
f. Immunostained gonadal pPGCs.



Extended Data Fig.3. Epigenetic reprogramming in pre- and early migratory pPGCs, and key germ cell markers in migratory pPGC and cultured porcine epiblast.
a. A cluster of pPGCs (dashed line) at ~E13 stained for TET1 and OCT4. Scale bar: 20 μm.

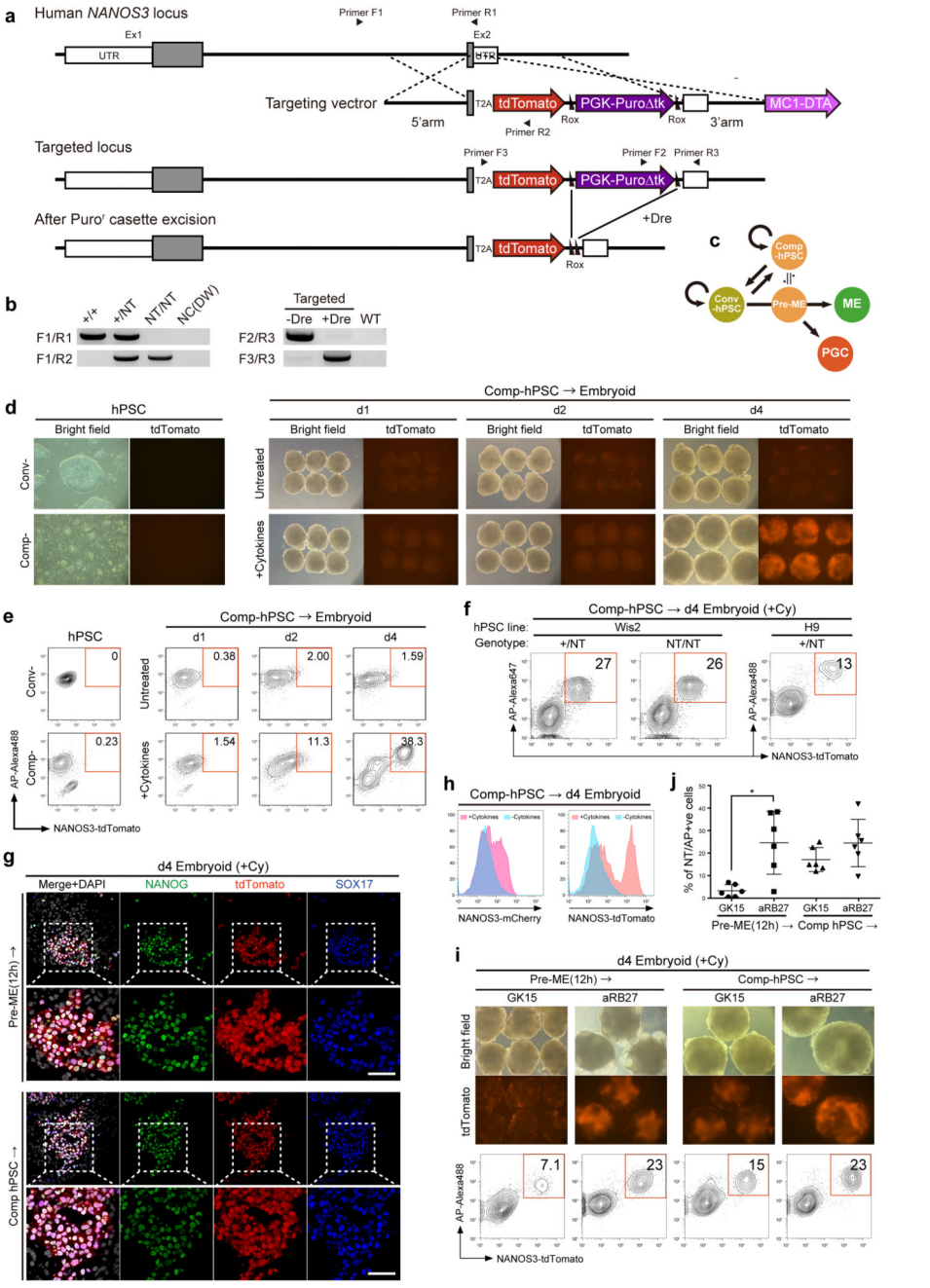
b. Serial sections of E14 embryos immunostained for different epigenetic markers combined with BLIMP1, NANOG and SOX17. Dashed lines highlight pPGC clusters. Scale bar: 20 μm.

c. Quantification of 5mC and H3M9me2 in embryos of different stages. Numbers of cells analyzed are indicated. (* $p<0.01$; Mann-Whitney test)).

d. Serial sections of ~E16 embryos showing migratory pPGCs immunostained for the indicated epigenetic marks. Scale bar: 20 μm.

e. Triple immunostaining of epiblasts cultured under different conditions. Scale bar: 10 μm.

Europe PMC Funders Author Manuscripts



Extended Data Fig.4. Characterization of NANOS3-tdTomato (NT) reporter hPSC

a. Targeting strategy for making NANOS3-tdTomato reporter.

b. Representative genotyping of targeted clones using genomic DNA.

c. Conventional (Conv) and PGC-competent (Comp) hPSCs states are reversible; the latter is equivalent to Pre-ME (12h at ME induction). Conv-hPSCs are cultured in Essential 8 medium on vitronectin coated dishes (see Methods). Comp-hPSCs are cultured in the hPSC medium containing inhibitor(i)s (GSK3i, ERKi, p38i, JNKi) on MEF (see Methods).

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

d. NT reporter Conv- and Comp-hPSC, and day 1-4 embryoids induced with or without Cytokines.

e. FACS pattern and % NT/AP +ve cells shown in Extended Data Fig. 4d.

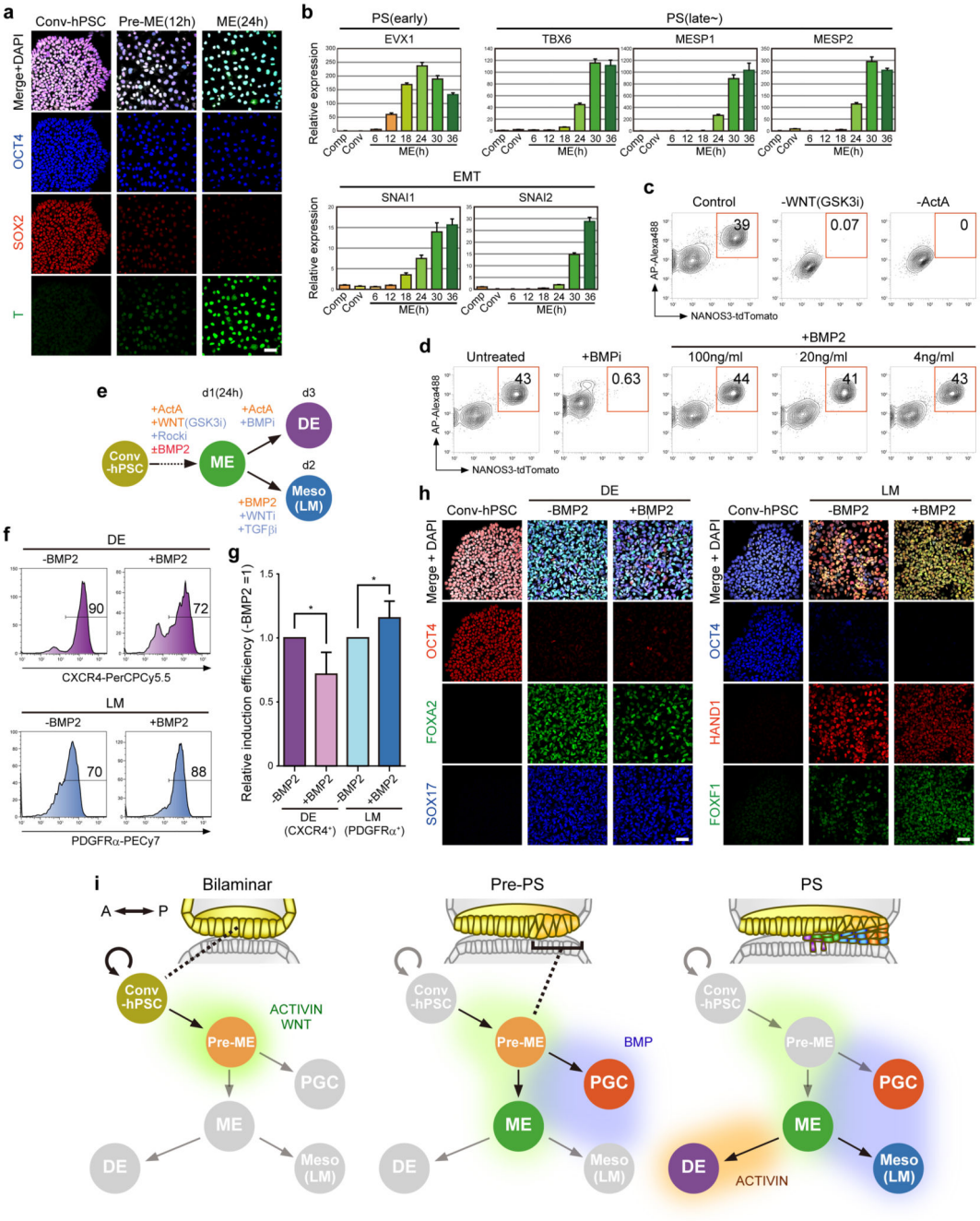
f. FACS pattern and % NT/AP +ve cells in multiple clones derived from Wis2 or H9 hESC lines.

g. Immunostaining of day 4 embryoids induced from Pre-ME(12h) or Comp-hPSC by BMP containing cytokines. Scale bar: 50 μm.

h. Comparison of sensitivity of 2 NANOS3 reporter cell lines. FACS patterns of day 4 embryoids induced from Comp-hPSC (harboring NANOS3-mCherry reporter or NANOS3-tdTomato reporter) with or without Cytokines.

i. Comparison of hPGC induction efficiency derived from Pre-ME(12h) or Comp-hPSC. Representative images and FACS patterns are shown.

j. Scatter plot shows % NT/AP +ve cells in indicated condition (n=6). Paired T-test was used to test for statistical significance (* $p<0.05$)



Extended Data Fig.5. Characterization of Pre-ME and ME induced from Conv-hPSC

a. Immunostaining of Conv-hPSCs during 12–24h ME induction. Scale bar: 50 μm.

b. Gene expression (RT-qPCR) change during ME induction.

c. FACS patterns of day 4 embryoids induced from Pre-ME(12h) with cytokines. Pre-ME was induced with or without GSK3i or Activin A.

d. FACS patterns of day 4 embryoids induced from Pre-ME(12h) with cytokines. Pre-ME was induced with or without BMP2 or the inhibitor.

e. Schematics of DE or lateral mesoderm (LM) differentiation from ME

Europe PMC Funders Author Manuscripts

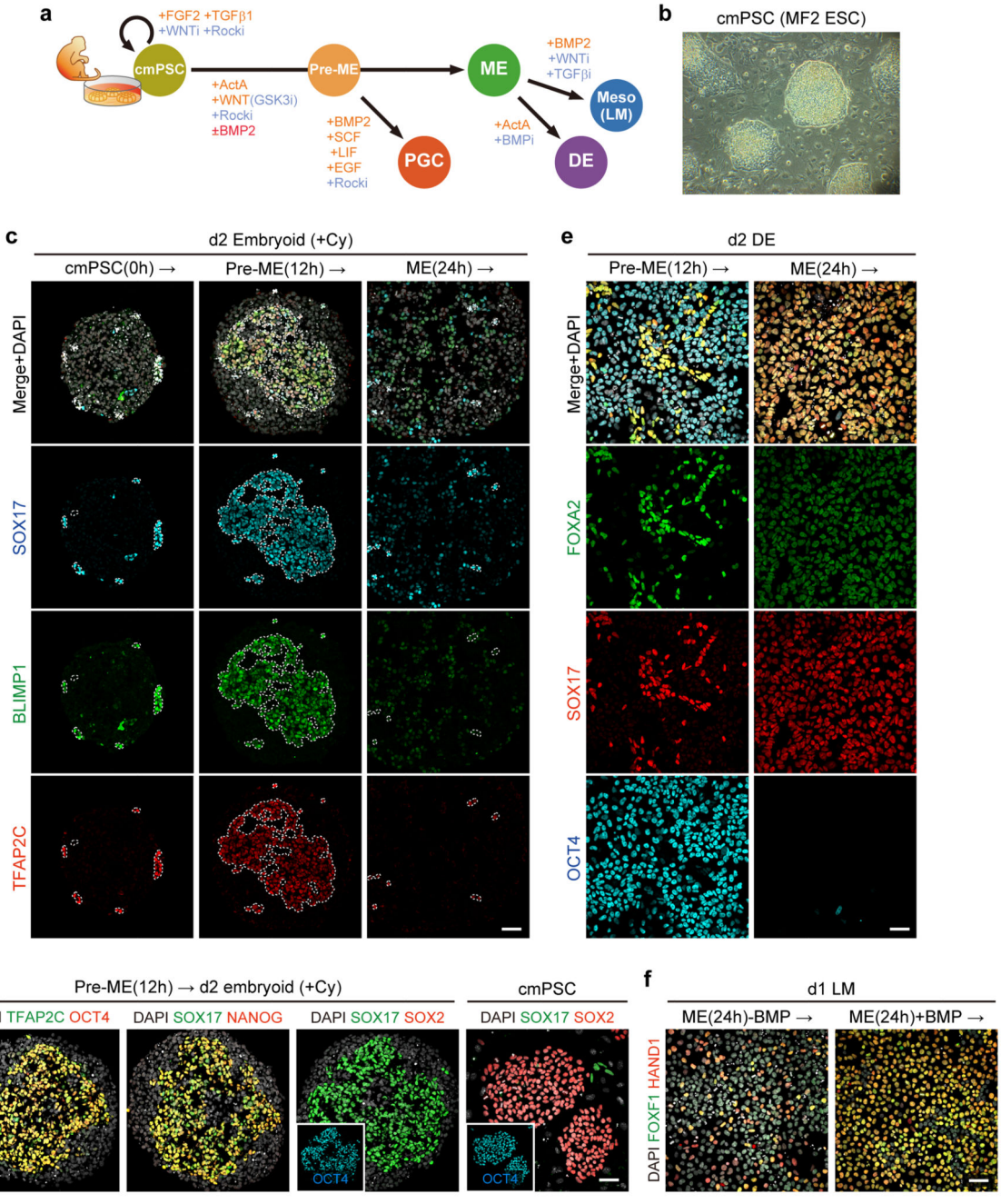
Europe PMC Funders Author Manuscripts

f. FACS patterns of day 2 DE (% CXCR4 +ve) and day 1 LM (% PDGFRa +ve) induced from 24h of ME induction.

g. Relative induction efficiency of DE or LM from ME induced with or without BMP2 (DE: n=5, LM: n=6). Paired T-test was used to test for statistical significance (* $p<0.05$)

h. Immunostaining of DE and LM in Extended Data Fig.5f. Scale bar: 50 μm.

i. Schematic of spatial-temporal progression from Conv-hPSC to ME and the signaling.



Extended Data Fig.6. Robust induction of cynomolgus monkey PGC (cmPGC) from cells during ME differentiation

a. Schematics of *in vitro* differentiation of cmPSC. The same system was adopted as shown for Conv-hPSC differentiation in Fig.3.

b. Bright field Image of undifferentiated cmPSC.

c. Immunostaining of day 2 embryoids induced with Cytokines from cells at 0h (=cmPSC), 12h and 24h during ME differentiation. Dashed lines highlight SOX17/BLIMP1/TFAP2C +ve cmPGCs. Scale bar: 50 μm.

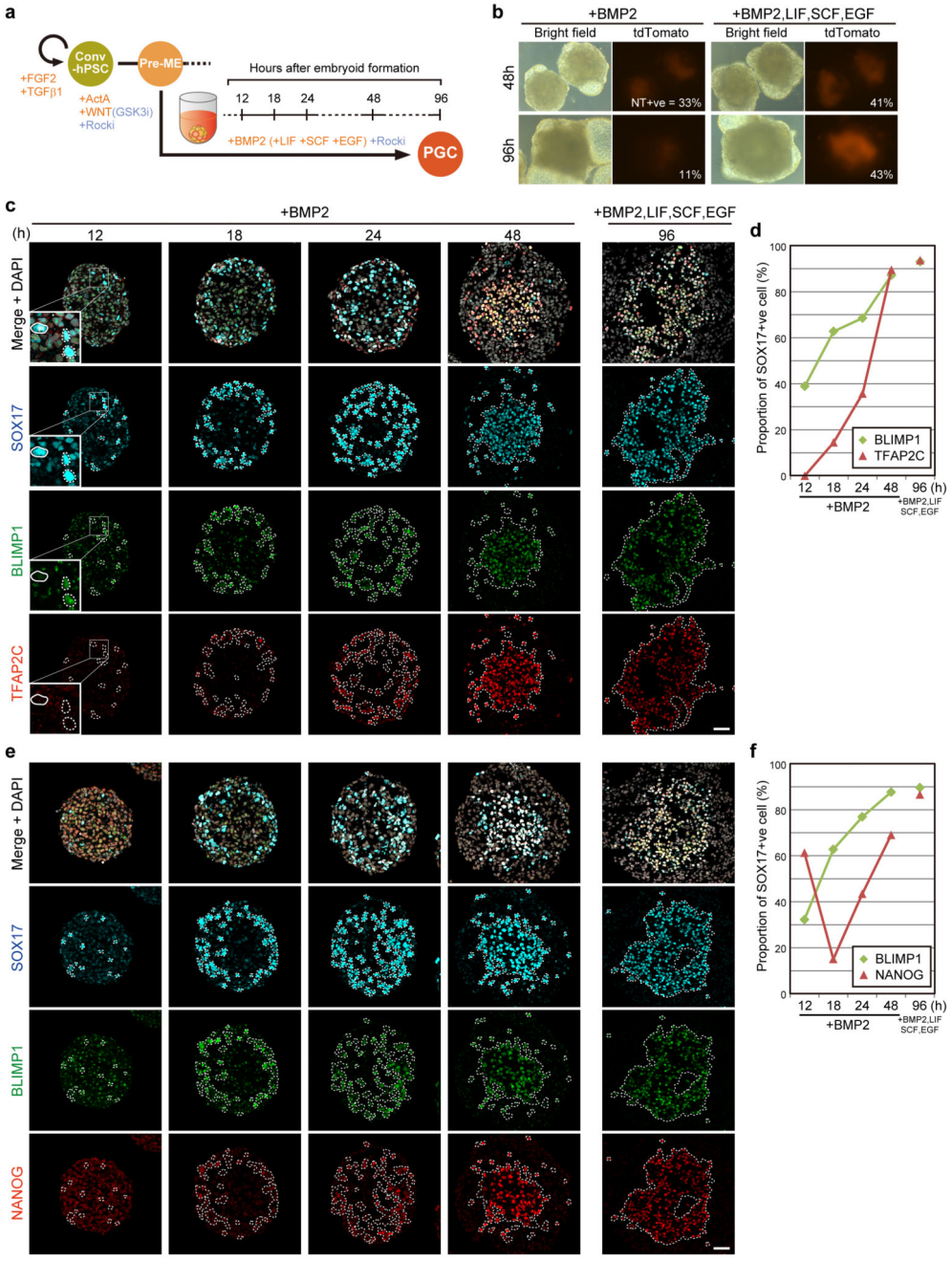
d. Immunostaining of day 2 Pre-ME(12h)-derived embryoids in Extended Data Fig. 6c for pluripotency markers. Notably, cmPGCs express SOX17 but not SOX2. In contrast, cmPSC colonies express SOX2 but not SOX17.

e. Immunostaining of day 2 cmDE induced from Pre-ME(12h) and ME(24h). Scale bar: 50 μm.

f. Immunostaining of day1 cmLM induced from ME(24h). ME were induced with or without BMP. Notably, adding BMP during ME differentiation increased the efficiency for FOXF1/HAND1 +ve LM cells, as shown in Conv-hPSC (Extended Data Fig. 5e-i)
Scale bar: 50 μm.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts



Extended Data Fig.7. Chronology of transcription factors expression during hPGC induction

a. Schematic of hPGC induction from Pre-ME(12h).

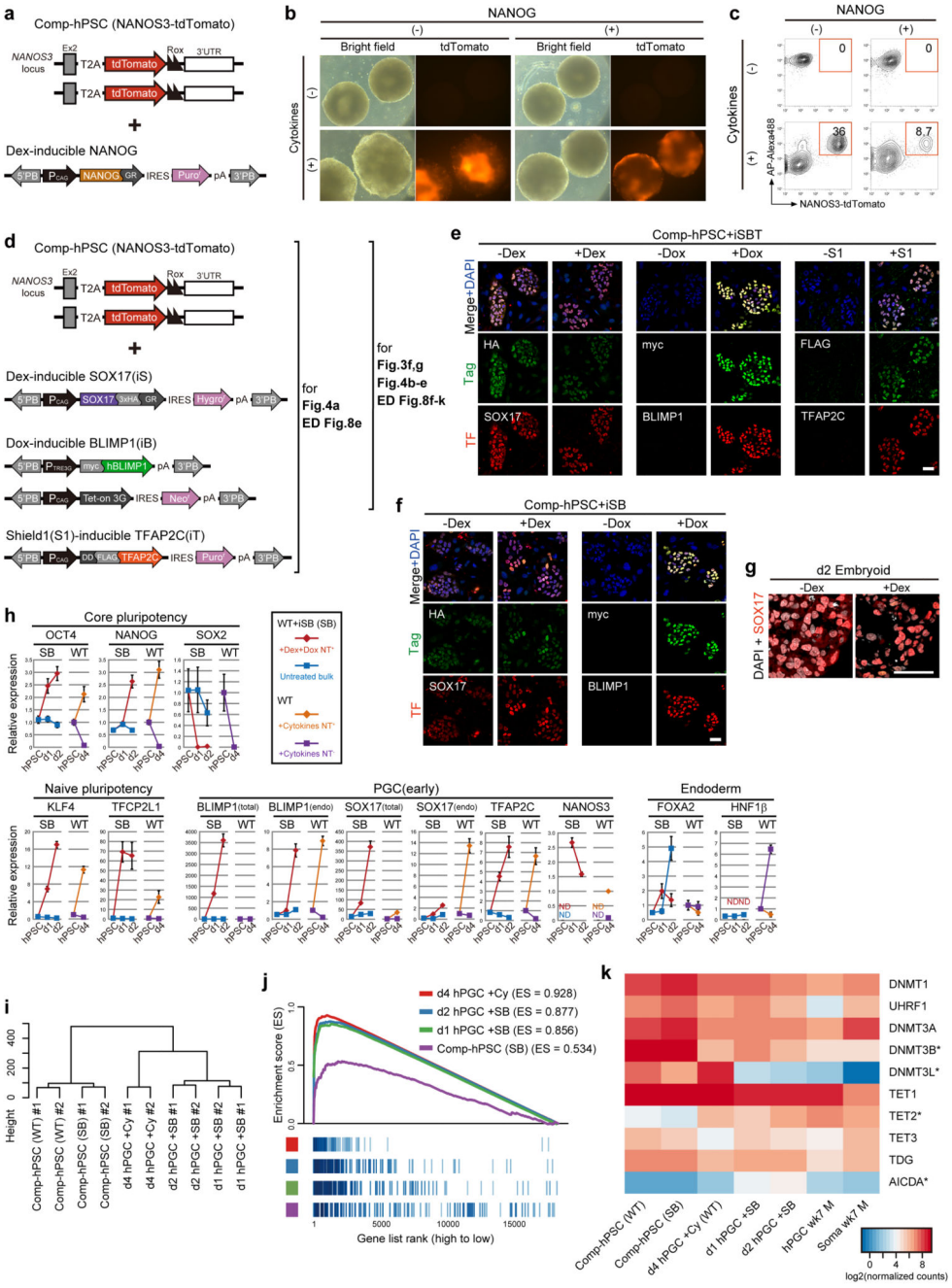
b. Images of day 2 and 4 embryoids in response to BMP2 alone or BMP2 with LIF, SCF and EGF(=Cytokines). Notably, BMP2 alone can induce hPGC at almost the same efficiency as the full cytokines, but do not survive during extended culture, as shown previously.

c. Immunostaining of embryoids induced with BMP2 alone or BMP2 with LIF, SCF and EGF showing expression of SOX17, BLIMP1 and TFAP2C. Scale bar: 50 μm.

d. Proportion of SOX17 +ve cells indicated in Extended Data Fig.7c.

e. Immunostaining of embryoids induced with BMP2 alone or BMP2 with LIF, SCF and EGF showing expression of SOX17, BLIMP1 and NANOG. Scale bar: 50 μm.

f. Proportion of SOX17+ve cells in Extended Data Fig.7e.



Extended Data Fig.8. Effect of NANOG on hPGC induction, characterization of NANOS3-tdTomato reporter hPSC containing inducible SOX17, BLIMP1 with or without TFAP2C, and similarity between Cytokine- and SOX17/BLIMP1-induced hPGC

a. Represents overexpression of Dex-inducible NANOG transgenes in NT reporter Comp-hPSC.

Europe PMC Funders Author Manuscripts

b. Day4 embryoids following induction of NANOG (by Dex), with or without cytokines as indicated.

c. FACS patterns after induction of hPGCs; NT/AP +ve cells (%) shown in Extended Data Fig. 8b.

d. Represents overexpression of Dex-inducible SOX17, Dox-inducible BLIMP1, Shield1(S1)-inducible TFAP2C transgenes in NT reporter Comp-hPSC.

e. Immunostaining of NT reporter Comp-hPSC +iSBT 1 day after induction of SOX17, BLIMP1 and TFAP2C by addition of Dex, Dox or S1. Scale bar: 50 μm.

f. Immunostaining of NT reporter Comp-hPSC +iSB 1 day after induction of SOX17 or BLIMP1 by addition of Dex or Dox. Scale bar: 50 μm.

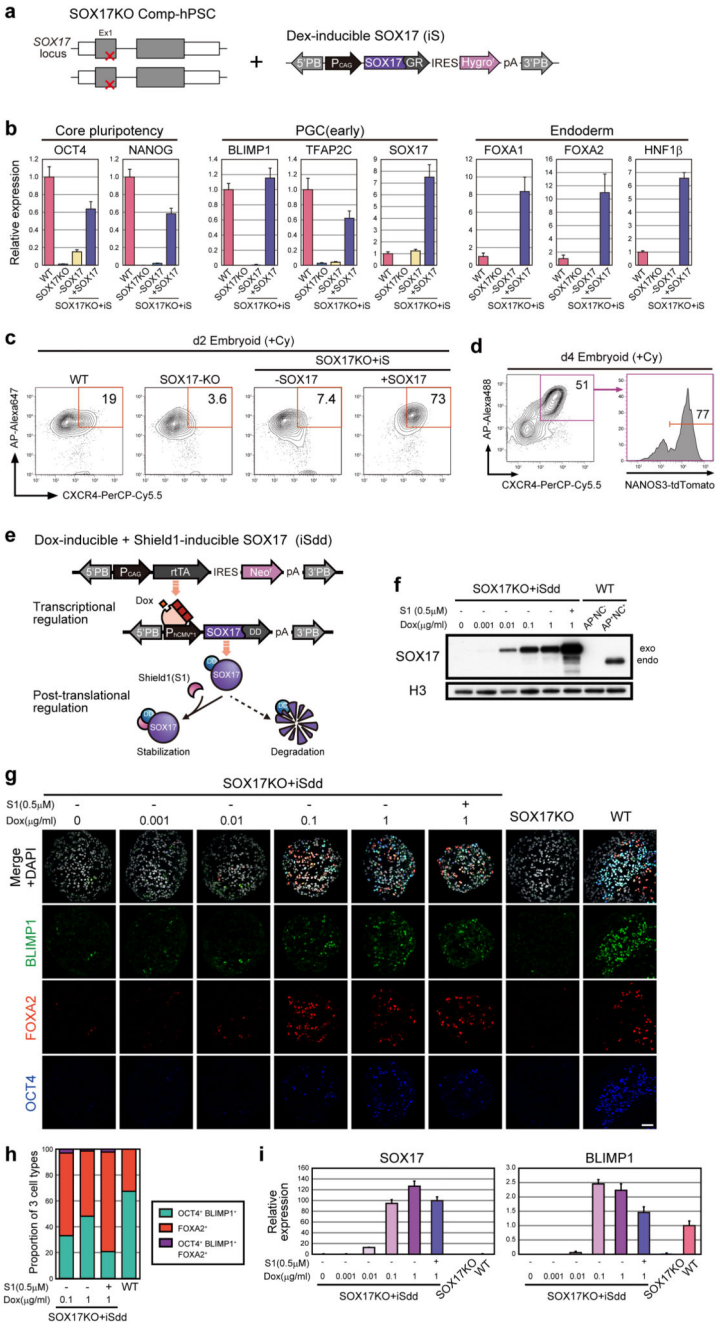
g. Immunostaining of day 2 embyroid induced with or without Dex to induce nuclear localization of SOX17. Notably, accumulation of SOX17 signal is observed in +Dex condition.

h. Changes in gene expression (RT-qPCR) during hPGC induction: Comp-hPSCs control (AP +ve cells); NT +ve hPGCs induced by SOX17/BLIMP1 or Cytokines; NT -ve cells in cells exposed to Cytokines.

i. Unsupervised hierarchical clustering (UHC) of gene expression.

j. Gene set enrichment analysis (GSEA) of 123 hPGC-specific genes (Supplementary Table 1) on the transcriptome of cytokine- and SOX17/BLIMP1-induced hPGCs.

k. Heat map showing expression of epigenetic modifiers related to global DNA demethylation. Same datasets as shown in Fig. 4e were used for analysis.



Extended Data Fig.9. Response of *SOX17* KO Comp-hPSC to SOX17

a. Overexpression of Dex-inducible SOX17 (iS) in SOX17 KO Comp-hPSC.

b. Gene expression (RT-qPCR) on day 4 of FACS-sorted NANOS3-mCherry (NC)/Alkaline phosphatase (AP) +ve hPGCs by RT-qPCR.

c. FACS analysis of d2 embryoids induced from WT Comp-hPSC, SOX17 KO Comp-hPSC and SOX17KO Comp-hPSC rescued with SOX17GR transgene (iS) (% CXCR4/AP+ cells).

d. FACS pattern of day 2 embyoid induced from NT reporter Comp-hPSC showing AP/CXCR4 +ve cells expressing NT.

Europe PMC Funders Author Manuscripts

e. Represents SOX17 inducible system (iSdd). Expression of SOX17 fused with destabilized domain (DD) can be induced by Doxycycline (Dox); addition of Shield1 (S1) can stabilize SOX17-DD protein.

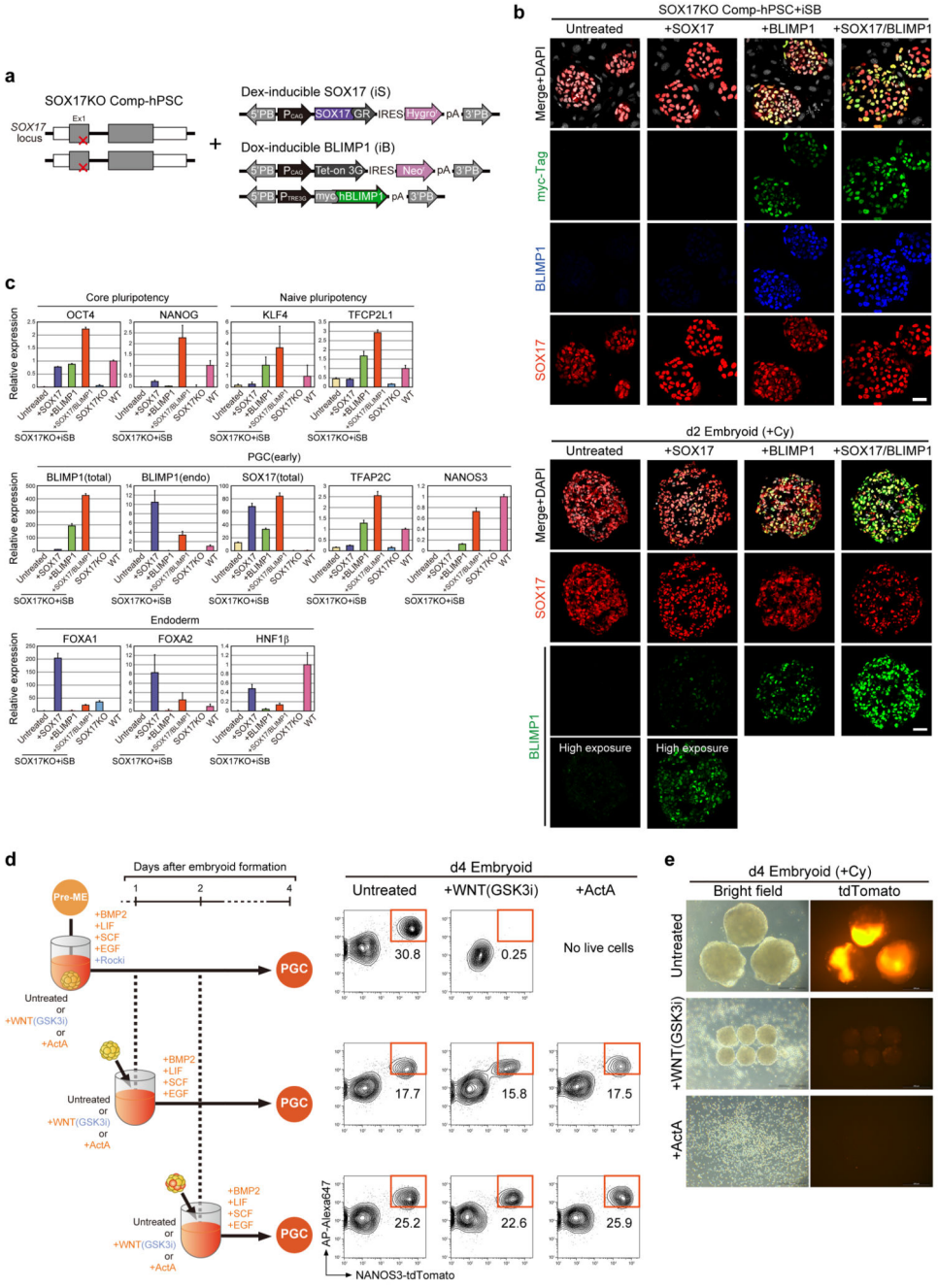
f. Western blots showing SOX17 expression level in day 5 embryoids from SOX17KO+iSdd Comp-hPSC. Embryoids were induced with Cytokines. To induce SOX17, different concentration of Dox and S1 were added. As controls, NC/AP +ve hPGCs and NC/AP -ve cells from WT Comp-hPSC-derived embryoids induced with Cytokines were used. Histone H3 (H3) was used for internal control.

g. Immunostaining of day 4 embryoids from SOX17KO+iSdd Comp-hPSC. Embryoids from SOX17 KO and WT Comp-hPSC induced with Cytokines were used as controls. Scale bar: 50 μm.

h. Quantification of immunostaining data in Extended Data Fig. 9g. The numbers of OCT4/BLIMP1 +ve hPGCs, FOXA2 +ve endodermal cells and OCT4/BLIMP1 +ve hPGCs expressing FOXA2 were counted from 3 different embryoids. The proportions of the 3 populations are shown.

i. Expression of SOX17 and BLIMP1 (RT-qPCR) in d4 embyoids in response to different SOX17 dosage.

Europe PMC Funders Author Manuscripts



Extended Data Fig.10. Response of *SOX17* KO Comp-hPSC to SOX17 and BLIMP1, and changes in epigenetic modifier expression after overexpression of SOX17/BLIMP1

a. Overexpression of Dex-inducible SOX17 (iS) and Dox-inducible BLIMP1 (iB) in SOX17 KO Comp-hPSC.

b. Immunostaining of Comp-hPSC 1 day after induction of SOX17 (Dex) or BLIMP1 (Dox), or both, and d2 embryoids following Dex-induced SOX17 with endogenous or Dox-induced BLIMP1. Scale bar: 50 μm.

c. Gene expression (RT-qPCR) in day 4 embryoids following induction by Dex (+SOX17), Dox (+BLIMP1), or Dex+Dox (+SOX17/BLIMP1). Bulk cells of embryoids induced from SOX17 KO and WT Comp-hPSC with Cytokines were used as controls.

d. Upon specification, hPGCs become refractory to Activin or Wnt signaling. Left schematic shows the experimental design. The embryoids were transferred to the medium with or without GSK3i (3μM) or ActivinA (100ng/ml) at day 0, 1 or 2 to see the effect on PGC induction. FACS patterns (right) show the induction efficiency of hPGC (%; NT/AP +ve) at day 4.

e. Day4 embryoids induced by cytokines with or without ActivinA or GSK3i.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank R. Campbell and C. Lee for help with animals and hPSCs, A. Riddell for FACS and T.Otani for advice on cmPSC. T.K. was supported by JSPS, the Uehara and Kanae Foundations; H.Z. by CSC; D.A.C. by CONACYT. The work was funded by BBSRC grant to R.A., C.A. and M.A.S (BB/M001466/1). M.A.S is a Wellcome Investigator.

References

1. Tang WW, Kobayashi T, Irie N, Dietmann S, Surani MA. Specification and epigenetic programming of the human germ line. Nat Rev Genet. 2016; 17:585–600. DOI: 10.1038/nrg.2016.88 [PubMed: 27573372]
2. Sugawa F, et al. Human primordial germ cell commitment in vitro associates with a unique PRDM14 expression profile. EMBO J. 2015; 34:1009–1024. DOI: 10.15252/embj.201488049 [PubMed: 25750208]
3. Sasaki K, et al. Robust In Vitro Induction of Human Germ Cell Fate from Pluripotent Stem Cells. Cell Stem Cell. 2015; 17:178–194. DOI: 10.1016/j.stem.2015.06.014 [PubMed: 26189426]
4. Irie N, et al. SOX17 Is a Critical Specifier of Human Primordial Germ Cell Fate. Cell. 2015; 160:253–268. DOI: 10.1016/j.cell.2014.12.013 [PubMed: 25543152]
5. 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. 2011; 146:519–532. DOI: 10.1016/j.cell.2011.06.052 [PubMed: 21820164]
6. Wu J, Izpisua Belmonte JC. Stem Cells: A Renaissance in Human Biology Research. Cell. 2016; 165:1572–1585. DOI: 10.1016/j.cell.2016.05.043 [PubMed: 27315475]
7. Rossant J. Mouse and human blastocyst-derived stem cells: vive les differences. Development. 2015; 142:9–12. DOI: 10.1242/dev.115451 [PubMed: 25516964]
8. Davidson KC, Mason EA, Pera MF. The pluripotent state in mouse and human. Development. 2015; 142:3090–3099. DOI: 10.1242/dev.116061 [PubMed: 26395138]
9. Tang WW, et al. A Unique Gene Regulatory Network Resets the Human Germline Epigenome for Development. Cell. 2015; 161:1453–1467. DOI: 10.1016/j.cell.2015.04.053 [PubMed: 26046444]
10. Guo F, et al. The Transcriptome and DNA Methylome Landscapes of Human Primordial Germ Cells. Cell. 2015; 161:1437–1452. DOI: 10.1016/j.cell.2015.05.015 [PubMed: 26046443]
11. Gkountela S, et al. DNA Demethylation Dynamics in the Human Prenatal Germline. Cell. 2015; 161:1425–1436. DOI: 10.1016/j.cell.2015.05.012 [PubMed: 26004067]
12. Klisch K, et al. The Sda/GM2-glycan is a carbohydrate marker of porcine primordial germ cells and of a subpopulation of spermatogonia in cattle, pigs, horses and llama. Reproduction. 2011; 142:667–674. DOI: 10.1530/rep-11-0007 [PubMed: 21896636]

13. Seki Y, et al. Cellular dynamics associated with the genome-wide epigenetic reprogramming in migrating primordial germ cells in mice. Development. 2007; 134:2627–2638. DOI: 10.1242/dev.005611 [PubMed: 17567665]
14. Hajkova P, et al. Epigenetic reprogramming in mouse primordial germ cells. Mech Dev. 2002; 117:15–23. [PubMed: 12204247]
15. Valdez Magana G, Rodriguez A, Zhang H, Webb R, Alberio R. Paracrine effects of embryo-derived FGF4 and BMP4 during pig trophoblast elongation. Dev Biol. 2014; 387:15–27. DOI: 10.1016/j.ydbio.2014.01.008 [PubMed: 24445281]
16. Yoshida M, et al. Conserved and divergent expression patterns of markers of axial development in eutherian mammals. Dev Dyn. 2016; 245:67–86. DOI: 10.1002/dvdy.24352 [PubMed: 26404161]
17. Aramaki S, et al. A mesodermal factor, T, specifies mouse germ cell fate by directly activating germline determinants. Dev Cell. 2013; 27:516–529. DOI: 10.1016/j.devcel.2013.11.001 [PubMed: 24331926]
18. Loh KM, et al. Efficient endoderm induction from human pluripotent stem cells by logically directing signals controlling lineage bifurcations. Cell Stem Cell. 2014; 14:237–252. DOI: 10.1016/j.stem.2013.12.007 [PubMed: 24412311]
19. D'Amour KA, et al. Efficient differentiation of human embryonic stem cells to definitive endoderm. Nat Biotechnol. 2005; 23:1534–1541. DOI: 10.1038/nbt1163 [PubMed: 16258519]
20. Murry CE, Keller G. Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. Cell. 2008; 132:661–680. DOI: 10.1016/j.cell.2008.02.008 [PubMed: 18295582]
21. Nakamura T, et al. A developmental coordinate of pluripotency among mice, monkeys and humans. Nature. 2016; 537:57–62. DOI: 10.1038/nature19096 [PubMed: 27556940]
22. Lim J, Thiery JP. Epithelial-mesenchymal transitions: insights from development. Development. 2012; 139:3471–3486. DOI: 10.1242/dev.071209 [PubMed: 22949611]
23. Sasaki K, et al. The Germ Cell Fate of Cynomolgus Monkeys Is Specified in the Nascent Amnion. Dev Cell. 2016; doi: 10.1016/j.devcel.2016.09.007
24. Kuckenberg P, Kubaczka C, Schorle H. The role of transcription factor Tcfap2c/TFAP2C in trophectoderm development. Reprod Biomed Online. 2012; 25:12–20. DOI: 10.1016/j.rbmo.2012.02.015 [PubMed: 22560121]
25. Robertson EJ, et al. Blimp1 regulates development of the posterior forelimb, caudal pharyngeal arches, heart and sensory vibrissae in mice. Development. 2007; 134:4335–4345. DOI: 10.1242/dev.012047 [PubMed: 18039967]
26. Carter AM, Enders AC. Placentation in mammals: Definitive placenta, yolk, sac, and paraplacenta. Theriogenology. 2016; 86:278–287. DOI: 10.1016/j.theriogenology.2016.04.041 [PubMed: 27155730]
27. Viotti M, Nowotschin S, Hadjantonakis AK. SOX17 links gut endoderm morphogenesis and germ layer segregation. Nat Cell Biol. 2014; 16:1146–1156. DOI: 10.1038/ncb3070 [PubMed: 25419850]
28. Murakami K, et al. NANOG alone induces germ cells in primed epiblast in vitro by activation of enhancers. Nature. 2016; 529:403–407. DOI: 10.1038/nature16480 [PubMed: 26751055]
29. Yamaji M, et al. Critical function of Prdm14 for the establishment of the germ cell lineage in mice. Nat Genet. 2008; 40:1016–1022. DOI: 10.1038/ng.186 [PubMed: 18622394]
30. Lin IY, et al. Suppression of the SOX2 neural effector gene by PRDM1 promotes human germ cell fate in embryonic stem cells. Stem Cell Reports. 2014; 2:189–204. DOI: 10.1016/j.stemcr.2013.12.009 [PubMed: 24527393]
31. Alberio R, Croxall N, Allegrucci C. Pig epiblast stem cells depend on activin/nodal signaling for pluripotency and self-renewal. Stem Cells Dev. 2010; 19:1627–1636. DOI: 10.1089/scd.2010.0012 [PubMed: 20210627]
32. Chen G, et al. Chemically defined conditions for human iPSC derivation and culture. Nat Methods. 2011; 8:424–429. DOI: 10.1038/nmeth.1593 [PubMed: 21478862]
33. Gafni O, et al. Derivation of novel human ground state naive pluripotent stem cells. Nature. 2013; 504:282–286. DOI: 10.1038/nature12745 [PubMed: 24172903]

34. Wang H, Luo X, Yao L, Lehman DM, Wang P. Improvement of Cell Survival During Human Pluripotent Stem Cell Definitive Endoderm Differentiation. Stem Cells Dev. 2015; 24:2536–2546. DOI: 10.1089/scd.2015.0018 [PubMed: 26132288]

35. Loh KM, et al. Mapping the Pairwise Choices Leading from Pluripotency to Human Bone, Heart, and Other Mesoderm Cell Types. Cell. 2016; 166:451–467. DOI: 10.1016/j.cell.2016.06.011 [PubMed: 27419872]

36. Ng ES, Davis R, Stanley EG, Elefanty AG. A protocol describing the use of a recombinant protein-based, animal product-free medium (APEL) for human embryonic stem cell differentiation as spin embryoid bodies. Nat Protoc. 2008; 3:768–776. DOI: 10.1038/nprot.2008.42 [PubMed: 18451785]

37. Kobayashi T, Alberio R, Surani MA. Simulating gastrulation development and germ cell fate in vitro using human and monkey pluripotent stem cells. Protoc Exch. 2017; doi: 10.1038/protex.2017.050

38. Magnusdottir E, et al. A tripartite transcription factor network regulates primordial germ cell specification in mice. Nat Cell Biol. 2013; 15:905–915. DOI: 10.1038/ncb2798 [PubMed: 23851488]

39. Grabole N, et al. Prdm14 promotes germline fate and naive pluripotency by repressing FGF signalling and DNA methylation. EMBO Rep. 2013; 14:629–637. DOI: 10.1038/embor.2013.67 [PubMed: 23670199]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

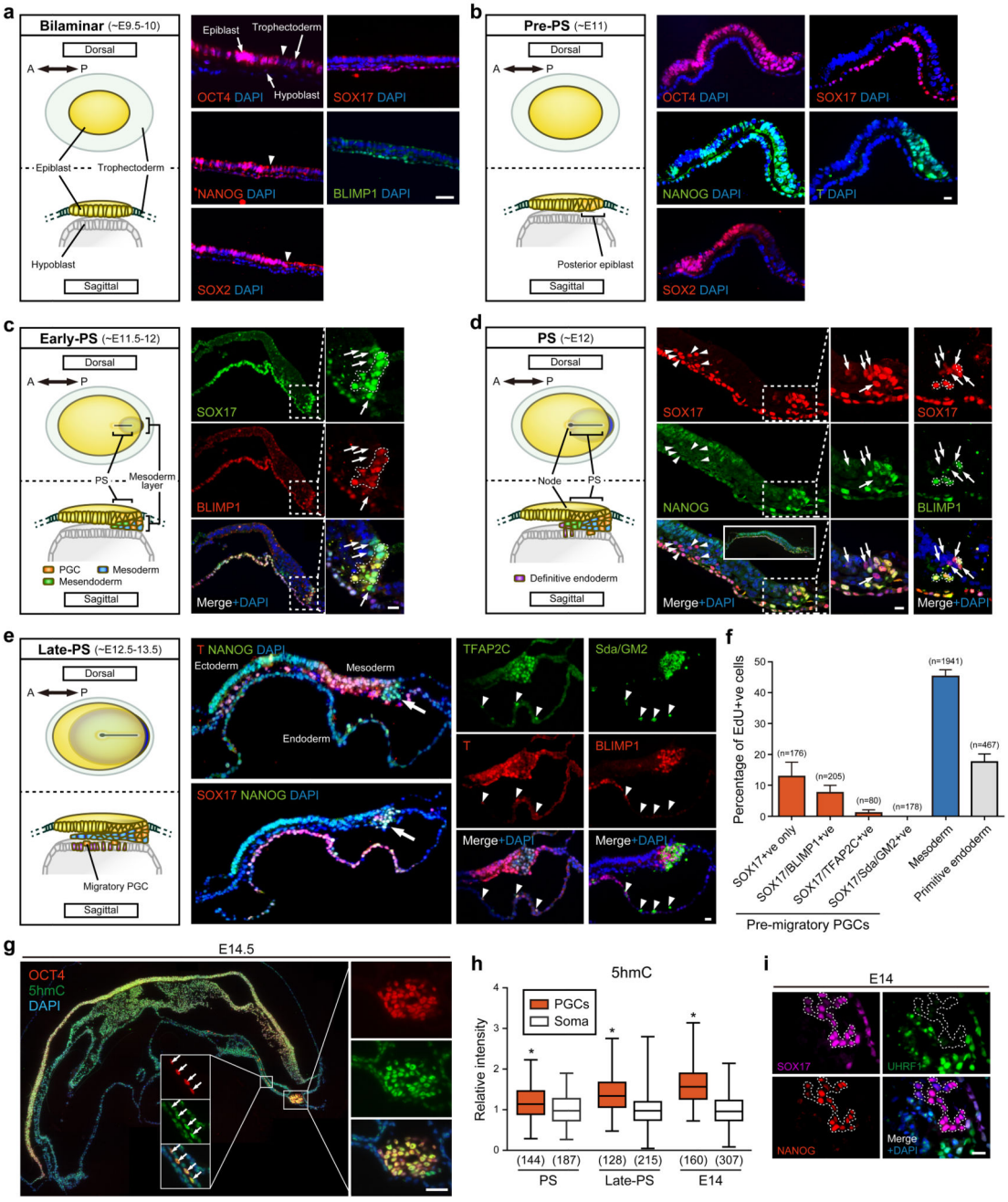


Fig.1. Specification of PGCs in gastrulating porcine embryos

Serial sections with immunostainings:

a. Bilaminar disc embryo (~E9.5-E10); Arrowhead marks the epiblast/trophectoderm boundary. Scale bar: 20 μm.

b. Pre-primitive streak embryo (Pre-PS; ~E11). Scale bar: 10 μm.

c. Early primitive streak embryo (Early-PS; ~E11.5-E12) with SOX17 and BLIMP1 expression. Close-up (dashed lines) shows four SOX17 +ve and BLIMP1 -ve cells (arrows).

Dashed lines highlight SOX17/BLIMP +ve cells. The hypoblast is SOX17/BLIMP1 +ve. Scale bar: 10 μm.

d. Primitive streak embryo (PS; ~E12) with a pPGC cluster showing SOX17 and NANOG expression. Four SOX17 +ve cells without NANOG in the most anterior pPGC cluster (arrows in middle image). The right most image (arrows) point to five SOX17 +ve and BLIMP1 -ve cells. Arrowheads show anterior PS with SOX17 +ve definitive endoderm cells. Dashed lines highlight SOX17/BLIMP +ve cells. Scale bar: 10 μm. Inset shows the whole embryo.

e. Late primitive streak embryo (Late-PS; ~E12.5-E13.5) with a pPGC cluster (arrow) showing NANOG, SOX17, TFAP2C, BLIMP1, T and Sda/GM2 expression. Arrowheads: early migratory pPGCs. Scale bar: 25 μm.

A<–>P; anterior-posterior axis

f. Quantification of EdU incorporation in pPGCs and somatic cells. Numbers denote analyzed cells.

g. Sagittal section of E14.5 embryo immunostained for OCT4 and 5hmC, and the pPGC cluster (white square). Arrows: migratory PGCs. Scale bar: 20 μm.

h. Quantification of 5hmC.in analyzed cells. (Mann-Whitney: * $p<0.01$).

i. Immunostaining for UHRF1 in E14 embryos. Dashed line delimits the pPGC cluster. Scale bar: 20 μm.

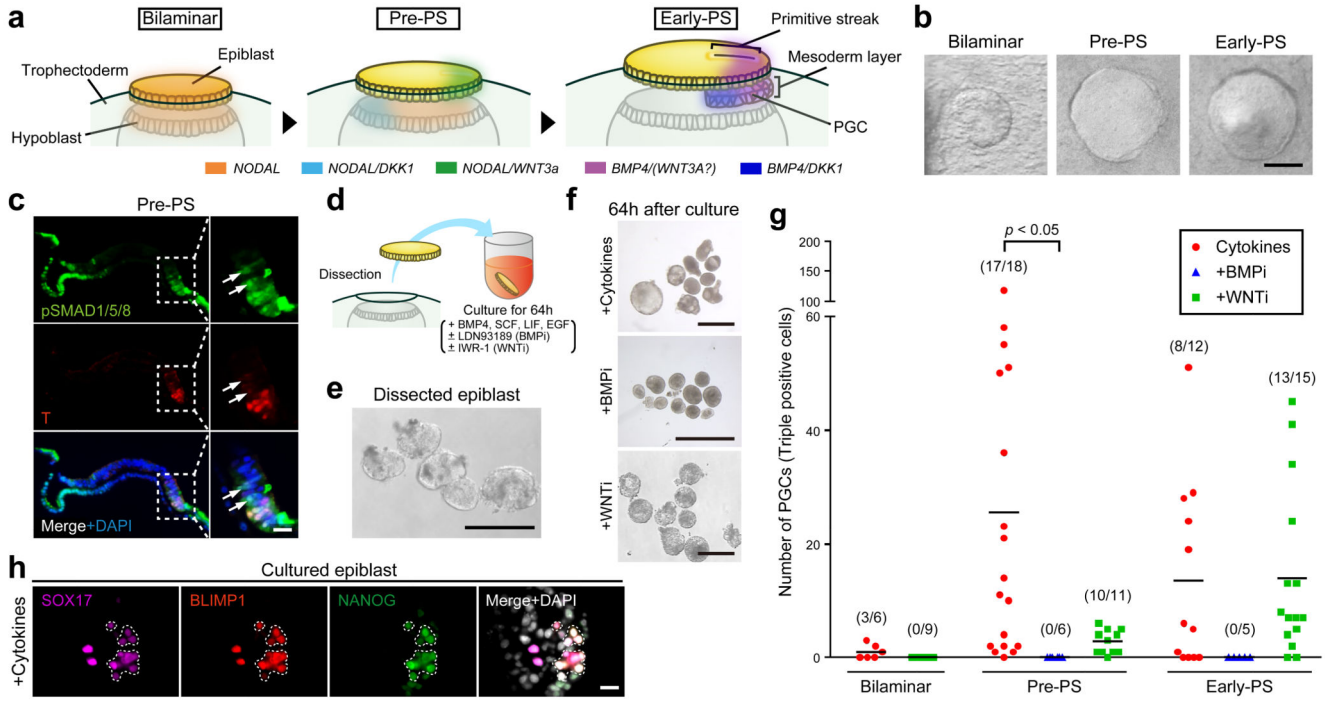


Fig.2. Competence for pPGC specification

a. Representation of porcine epiblasts and signaling for pPGC induction.

b. Dissected epiblasts. Scale bar: 0.25 mm.

c. Serial sections of a Pre-PS embryo showing pSMAD1/5/8 and T. Close-up (dashed lines) of posterior epiblast cells (arrows) with nuclear pSMAD but no T. Scale bar: 20 μm.

d. Epiblasts culture for 64 h.

e. Epiblasts prior to culture. Scale bar: 0.5mm.

f. Epiblasts after 64 h culture.

g. Scatter plot for the triple positive pPGCs staining.

h. Triple immunostaining of epiblasts. Scale bar: 20 μm.

WNTi (WNT inhibitor); BMPi (BMP inhibitor)

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

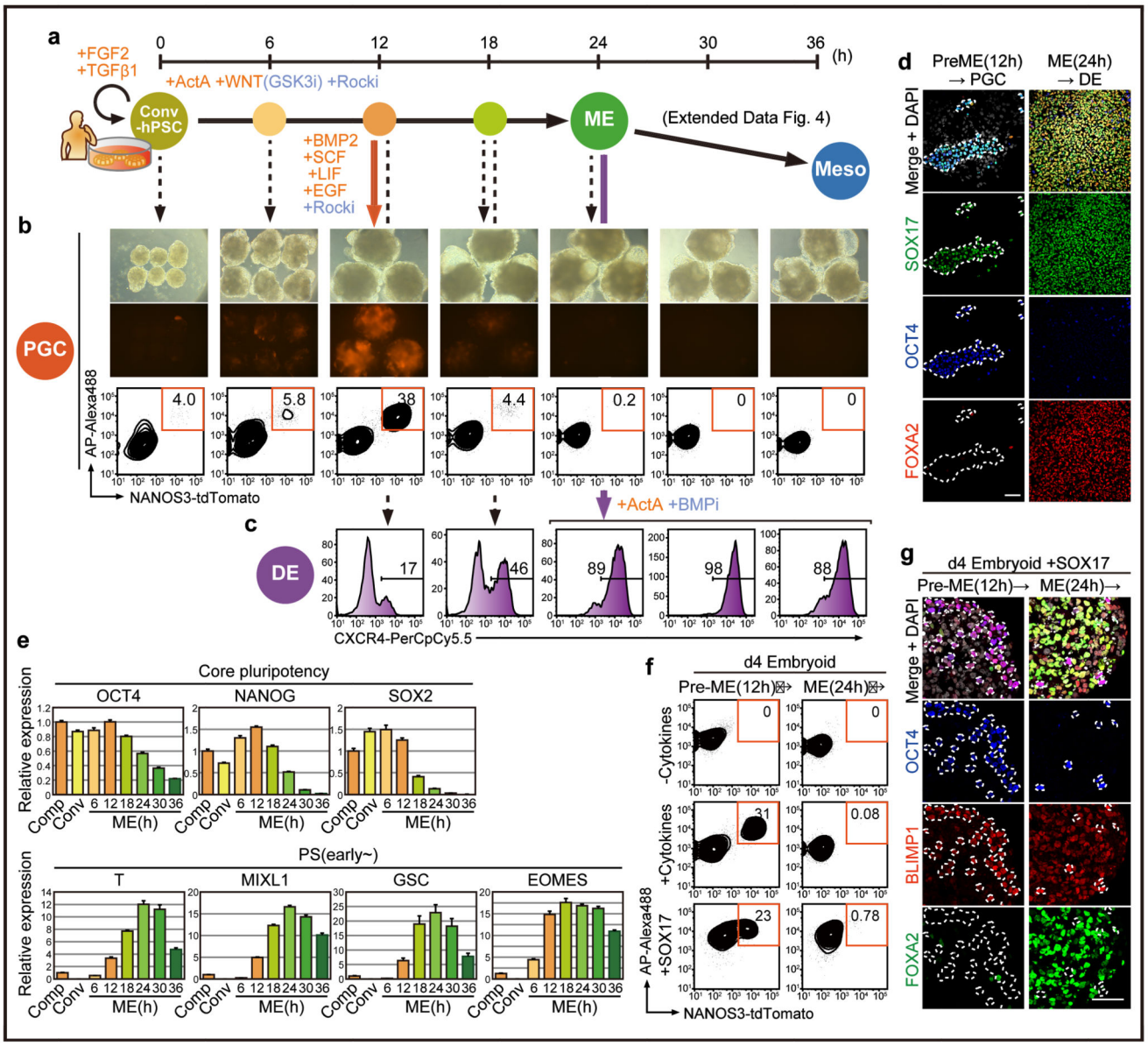


Fig.3. Simulating of human peri-gastrulation and hPGC competency

a. Schematics showing progressive gain and loss of competency for hPGCs.

b. Induction of hPGCs (% NT/AP +ve cells) in day 4-5 embryoids.

c. Induction of of definitive endoderm (DE: (% CXCR4 +ve cells)) at 12h–24h of during ME differentiation.

d. Immunostaining of hPGCs induced from Pre-ME(12h) and of DE from ME(24h). Dashed lines: SOX17/OCT4 +ve hPGCs. Scale bar: 50 μm.

e. Gene expression (RT-qPCR) changes during ME induction.

f. Induction of hPGCs (%NANOS3-tdTomato/AP +ve cells) from Pre-ME(12h) or ME(24h). Pre-ME are competent for hPGCs in response cytokines or SOX17, but ME are not.

g. Immunostaining of day4 embryoids from Pre-ME(12h) or ME(24h). SOX17 in Pre-ME induced OCT4/BLIMP1 +ve hPGCs (dashed line), but in ME, it induced DE (BLIMP1/FOXA2 +ve) cells. Scale bar: 50 μm.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

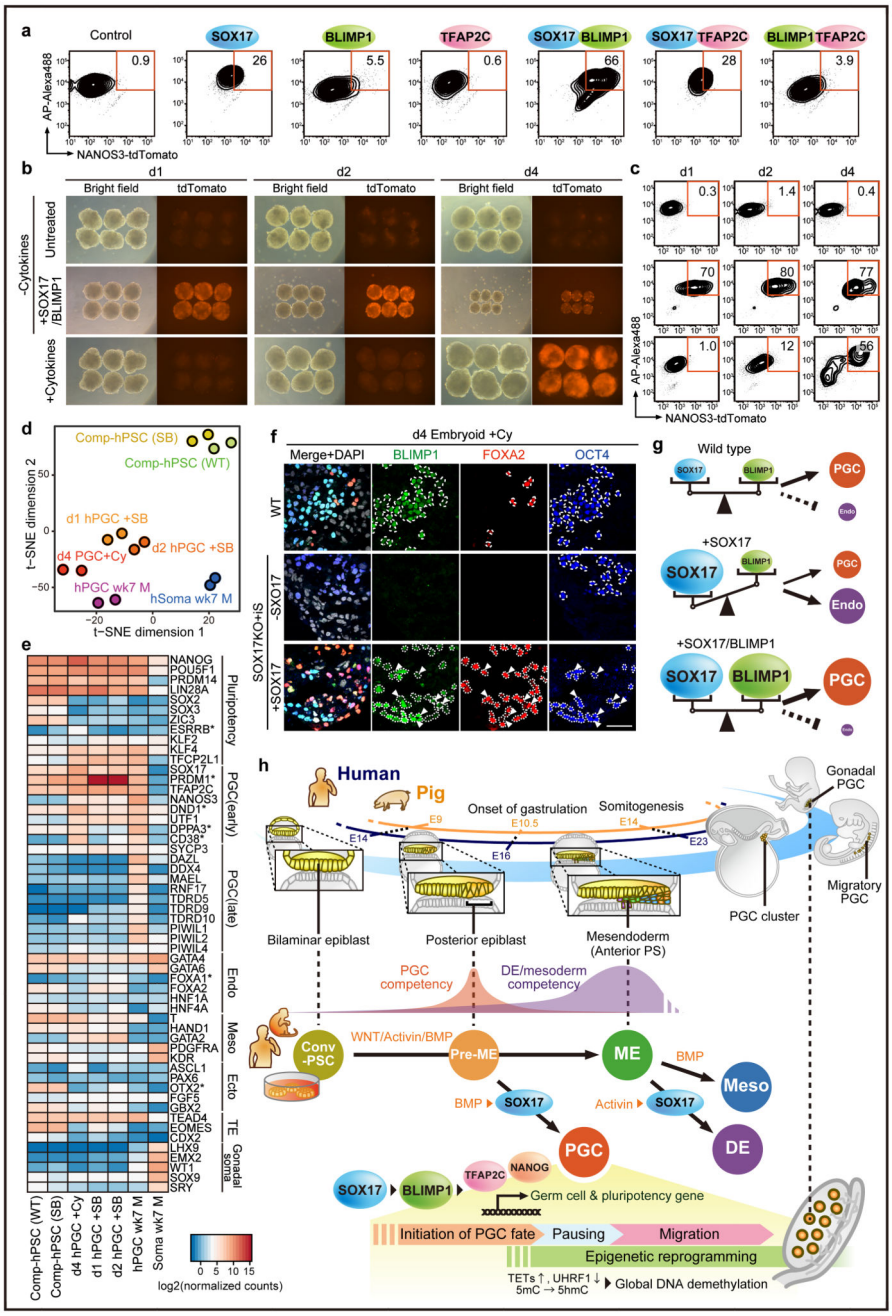


Fig.4. Induction of hPGCs by SOX17-BLIMP1 and combined representation of hPGCs and pPGCs specification

a. Induction of hPGCs by transcription factors in Comp-hPSCs: SOX17-BLIMP1 induce hPGCs

b. Day1–4 embryoids with NANOS3-tdTomato (NT) reporter respond to ectopic SOX17 and BLIMP1 with or without cytokines.

c. Induction of hPGCs (% NT/AP +ve cells) shown in Fig.4b.

d. t-SNE analysis of RNA-seq data.

e. Heat map of representative gene expression includes previous data4.

f. Day 4 embryoids induced by cytokines. Dashed line indicates positive signals. Arrowheads: OCT4/BLIMP1 +ve hPGC expressing FOXA2. Scale bar: 50 μm.

g. SOX17 and BLIMP1 gene dosage is critical during hPGC specification: BLIMP1 represses SOX17-induced endodermal genes.

h. Representation of human germ cell origin and program based on *in vitro* simulations from hPSCs and *in vivo* origin of pPGCs in porcine embryos.

Europe PMC Funders Author Manuscripts