Trophectoderm Potency is Retained Exclusively in Human Naïve Cells

Ge Guo[1,7,8] *, Giuliano Giuseppe Stirparo[1,7] *, Stanley Strawbridge[1], Jian Yang[5,6], James Clarke[1], Meng Amy Li[1], Sam Myers[1], Buse Nurten Özel [1], Jennifer Nichols[1,2] and Austin Smith[1,3,4]

[1] Wellcome-MRC Cambridge Stem Cell Institute, Jeffrey Cheah Biomedical Centre, University of Cambridge, Cambridge, CB2 0AW, United Kingdom.
[2] Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, CB2 1GA, United Kingdom.
[3] Department of Biochemistry, University of Cambridge, Cambridge, CB2 1QR, United Kingdom
[4] Living Systems Institute, University of Exeter, Exeter, EX4 4QD, United Kingdom
[5] Guangzhou Institutes of Biomedicine and Health (GIBH), Chinese Academy of Sciences, Guangzhou, China, 510530
[6] Present address: Key Laboratory of Arrhythmias, Ministry of Education, Shanghai East Hospital, Tongji University School of Medicine, Shanghai 200120, China.
[7] These authors contributed equally
[8] Lead contact

* Correspondence: gg251@cam.ac.uk (G.G.); jn270@cscr.cam.ac.uk (J.N.); austin.smith@exeter.ac.uk (A.S.)

SUMMARY

Classical mouse embryology has established a sequence of lineage bifurcations underpinning early mammalian development. Consistent with this paradigm, mouse embryonic stem cells have lost the capacity to generate extraembryonic trophectoderm. We show here, however, that human naïve epiblast stem cells readily produce this lineage. Inhibition of ERK signalling, fundamental to naïve cell propagation, is unexpectedly instrumental in trophectoderm induction. Transcriptome analyses authenticate trophoblast fate and expose a trajectory via reversion to inner cell mass (ICM). Nodal inhibition enhances differentiation and BMP signalling is not engaged. Strikingly, after formative transition primed stem cells cannot make trophectoderm but respond to BMP and form amnion. Gene perturbations in naïve cells reveal that YAP and TFAP2C promote trophectoderm as in mouse, while NANOG suppresses distinctively in human. Finally, ICMs from expanded human blastocysts efficiently regenerate trophectoderm. Thus in human, retained trophectoderm potential is an integral feature of emergent pluripotency that confers higher regulative plasticity.

INTRODUCTION

Trophectoderm is a cell lineage evolved in mammals to mediate blastocyst formation and uterine implantation, provide morphogenetic signals that pattern the early embryo, and later produce components of the placenta. In the early mammalian embryo, trophoblasts are the first specialized cell type to differentiate. Following fertilization and early cleavage divisions, blastomeres in eutherian embryos divide asymmetrically to produce both trophectoderm and undifferentiated inner cell mass (ICM). Delamination of the epithelial trophectoderm initiates blastocyst formation. Classic studies in mouse embryos have established that the topological segregation of trophectoderm and ICM is rapidly followed by fate restriction, such that by the mid-blastocyst (late 32-cell) stage ICM cells can no longer make trophectoderm (Gardner, 1983; Nichols and Gardner, 1984). Lineage restriction is reflected in the consolidation of distinct molecular identities (Posfai et al., 2017). Subsequently, a second binary fate decision resolves the ICM into epiblast and hypoblast (Chazaud et al., 2006; Gardner and Rossant, 1979; Plusa et al., 2008; Saiz et al., 2016). These observations have given rise to a textbook model of sequential lineage bifurcations at the onset of mammalian embryo development.

Mouse embryonic stem (ES) cells are cell lines derived directly from the naïve pre-implantation epiblast (Brook and Gardner, 1997; Evans and Kaufman, 1981; Martin, 1981; Nichols et al., 2009). Over prolonged expansion in vitro they retain global transcriptome proximity to their tissue stage of origin (Boroviak et al., 2014). Functionally, they can contribute massively to all embryo tissues in chimaeras, but do not make appreciable contributions to trophoblast derivatives (Beddington and Robertson, 1989; Bradley et al., 1984; Nagy et al., 1993), in line with the paradigm of early trophectoderm lineage restriction.

Trophectoderm versus ICM determination in the developing mouse blastocyst is underpinned by mutually exclusive and antagonistic expression of transcription factors Oct4 and Cdx2 (Strumpf et al., 2005). ES cells can be made to transdifferentiate into trophectoderm-like cells by either forced expression of *Cdx2* or deletion of *Oct4* (Niwa et al., 2005). Expression of other trophectoderm lineage transcription factors such as Tfap2c (Adachi et al., 2013), or demethylation and up-regulation of Elf5 (Ng et al., 2008), also provoke transdifferentiation into trophectoderm. Detailed characterization, however, has revealed that while cells with morphological features and markers of trophoblast are readily generated, functional trophectoderm identity is not fully established in these settings (Cambuli et al., 2014). Thus trophectoderm lineage restriction may be breached in mouse ES cells but is not readily erased. Moreover, depletion of Nanog, a central component of the ES cell transcription factor network, destabilises naive identity but results in primitive endoderm (Chambers et al., 2007; Mitsui et al., 2003), indicating that trophectoderm is not a “default” programme.

Human pluripotent stem cells (hPSC) (Takahashi et al., 2007; Thomson et al., 1998; Yu et al., 2007) differ from mouse embryonic stem cells and are considered to represent post-implantation epiblast, primed for embryonic germ layer differentiation (Rossant, 2015). Primed hPSC have been reported to differentiate into trophoblast upon treatment with bone morphogenetic protein (BMP) (Amita et al., 2013; Xu et al., 2002). However, generation of a pre-implantation lineage by stem cells that have post-implantation identity (Nakamura et al., 2016; O'Leary et al., 2012) is surprising and without apparent developmental logic. Furthermore, there are no reports that BMP is involved in trophoblast specification in the embryo and the BMP-induced cells in vitro do not fulfil stringent criteria for human

trophoblast identity (Bernardo et al., 2011; Lee et al., 2016). More recently, extended potential hPSCs (hEPSCs) have been described and reported to differentiate more readily into trophoblast-like cells in a BMP-dependent manner (Gao et al., 2019; Yang et al., 2017). However, the developmental authenticity, either of hEPSCs or of their trophoblast-like progeny, has yet to be ascertained.

Culture conditions have now been developed (Guo et al., 2017; Takashima et al., 2014; Theunissen et al., 2014) that support self-renewal of human stem cells with transcriptome proximity to pre-implantation epiblast (Bredenkamp et al., 2019a; Nakamura et al., 2016; Stirparo et al., 2018) and other naïve features (Dong et al., 2019). Consistent with early epiblast identity, naïve stem cells do not respond directly to germ layer inductive stimuli, but must first acquire competence through a process of capacitation (Guo et al., 2017; Rostovskaya et al., 2019). This formative transition reflects the dramatic change from pre- to early post-implantation epiblast in the embryo (Rossant and Tam, 2017; Smith, 2017).

In the developing human blastocyst CDX2 protein appears later than in the mouse and OCT4 is present in the initial trophectoderm layer (Niakan and Eggan, 2013). Similar observations have been reported in bovine (Berg et al., 2011), indicating that the mechanics of lineage segregation defined in mouse may vary in other mammals. The availability of naïve stem cells provides the opportunity for experimental interrogation of lineage restriction in human early epiblast. Here we investigate naïve cell propensity to produce trophectoderm and show how this differs from primed hPSC.

RESULTS

Human naïve stem cells can enter the trophectoderm lineage

Mouse ES cells self-renew efficiently in the presence of LIF and the MEK inhibitor PD0325901 (PD) alone (Dunn et al., 2014; Ying et al., 2008), whereas stable propagation of human naïve stem cells additionally requires the atypical protein kinase C inhibitor Gö6983 and blockade of the Wnt pathway (Bredenkamp et al., 2019a). We observed that when plated without feeders in PD only, a large fraction of human naïve cells differentiated within 3 days into flattened epithelial cells (Figure 1A). To determine the character of this differentiation we inspected early lineage marker expression (Figure 1B). Post-implantation epiblast markers, *SOX11* and *CDH2*, were not up-regulated, consistent with blockade of the formative pluripotency transition by PD. Nor were *GATA4* and *SOX17*, markers of primitive endoderm, induced (Figure S1A). However, *GATA2* and *GATA3*, transcription factors diagnostic of trophectoderm, were highly up-regulated. Significantly, Gö6983 and the Wnt inhibitor XAV939 individually reduced, and together completely blocked, induction of *GATA2* and *GATA3* (Figure S1B).

To evaluate whether expression of *GATA2* and *GATA3* might signify embarkation into the trophectoderm lineage we investigated additional markers. We observed induction of trophectoderm-associated genes *TEAD3* and *DAB2* over 3 days in PD only (Figure 1C). We confirmed that PD culture resulted in up-regulation of these markers in three other cell lines, including embryo derived HNES1 and chemically reset naïve cells (Figure S1C).

Immunostaining showed that many cells became positive for GATA3 protein after 3 days in PD (Figure 1D). Presence of GATA3 was exclusive with naïve pluripotency factors KLF17

and NANOG, which persisted in localized areas of the culture. To facilitate monitoring and quantification of trophectoderm induction we generated a *GATA3:mKO* knock in reporter line by CRISPR/Cas9 mediated homologous recombination in embryo-derived HNES1 naïve stem cells (Figure 1E). Reporter expression was undetectable in self-renewing cells in PXGL or on transfer to N2B27, but became clearly visible by fluorescence microscopy upon culture in PD only. Flow cytometric analyses quantified mKO bright cells on Day 3, ranging from 50-70% between experiments (Figure 1F).

The indication of trophectoderm differentiation from human naïve stem cells is unexpected because mouse ES cells do not generate this lineage at any appreciable frequency without genetic or epigenetic manipulation. We cultured mouse ES cells in PD only and did detect induction of *GATA3* (Figure S1D). However, neither *GATA2* nor other trophectoderm genes were up-regulated, consistent with inability to enter the lineage and indicating that the response of human naïve cells to PD is species-specific.

Nodal pathway inhibition facilitates trophectoderm induction

We previously reported that, in contrast to conventional primed hPSC, human naïve cultures do not collapse acutely in the presence of A83-01 (A83), a potent inhibitor of Nodal and Activin signaling (Takashima et al., 2014). However, colony-forming efficiency was reduced and cultures were not maintained long term. Using the *GATA3:mKO* cells we observed reporter activation in a significant fraction of cells after three or more passages in PXGL plus A83, coincident with emergence of spread cell morphology (Figure 2A, B). We also detected progressive increase in *GATA2* and *GATA3* mRNA (Figure S2A). Human naïve cells are cultured without exogenous activin or other TGFβ family ligand that can activate SMAD2 and SMAD3 effectors. Nonetheless, we found that SMAD2 phosphorylation is abundant in naïve cells (Figure 2C). This is presumably due to autocrine activation by NODAL, which is expressed in naïve cells together with the convertase FURIN (Figure S2B)(Rostovskaya et al., 2019). A83 abolished SMAD2 phosphorylation (Figure 2D). We investigated whether culture in A83 alone might induce trophectoderm similarly to PD but observed only rare activation of the *GATA3:mKO* reporter (Figure 2E). However, we found that A83 enhanced the response to PD such that mKO bright cells appeared earlier and in greater numbers (Figure 2F). mRNAs for both *GATA3* and *GATA2* were accordingly up-regulated more rapidly (Figure 2G). We performed live cell imaging taking advantage of the mKO reporter. Time-lapse still images (Figure 2H) and movie (Supplemental Movies 1 and 2) show conversion over 60 hours in PD+A83 into a mKO positive epithelial monolayer of trophectoderm-like cells.

These results indicate that blockade of autocrine activity of the NODAL pathway in human naïve cells acts together with inhibition of ERK signaling releases robust trophectoderm lineage potential.

Transcriptome dynamics during naïve cell differentiation

For assessment of global similarity with embryo trophoblast we first reanalyzed human embryo transcriptome data (Petropoulos et al., 2016) to extract profiles for early (E5) and late (E6,E7) trophoblast (Figure S3A-F). We then undertook whole transcriptome sequencing over the time course of naïve cell differentiation in N2B27 medium alone or supplemented with PD, A83 or PD plus A83 (PD+A83). We prepared RNAseq libraries in duplicate from both embryo-derived HNES1 and reset cR-H9 cells. Dimensionality reduction

by principal component analysis (PCA) aligned samples according to treatment and time along two distinct trajectories (Figure 3A). One path was common to N2B27 and A83, and the second to PD and PD+A83. We superimposed on the PCA the density of contribution of genes with enriched expression in human embryo late trophoblast and found that this overlaid the endpoint of the PD and PD+A83 trajectory (Figure 3A). In contrast, the final day of N2B27 and A83 cultures lay adjacent to the maximum density area for genes up-regulated in primed hPSC. Hierarchical clustering using differentially expressed genes between epiblast and trophoblast further substantiated conversion of HNES1 and cR-H9 naïve cells in PD or PD+A83 into late blastocyst stage trophoblast, whereas trophoblast genes were not up-regulated in N2B27 or A83 alone (Figure 3B, S3G).

We identified 6 clusters of dynamically expressed genes during differentiation in PD+A83 (Figure 3C). Comparison with the human embryo transcriptome profiles revealed that epiblast-enriched genes are over-represented in clusters 1 and 2 (Figure 3D) whereas genes up-regulated in trophoblast are predominantly in clusters 5 and 6 (Figure 3E). The distribution of gene expression in the human embryo confirmed that clusters 1 and 2 are mostly expressed in epiblast while clusters 5 and 6 are enriched in trophoblast of the mature blastocyst (Figure 3F, S3H).

As an independent assessment of correspondence with primate embryo development we took advantage of single cell transcriptome data from *Macaca fascicularis* embryos (Nakamura et al., 2016). We averaged the scRNA-seq embryo data according to developmental tissue and stage and computed the PCA together with our population RNA-seq timecourse. The N2B27 and A83 samples gained similarity to post-implantation epiblast whereas the PD and PD+A83 differentiation trajectory related to trophectoderm progression (Figure S3I).

Examples of expression dynamics in vitro and in the human embryo for genes from clusters 5 and 6 are shown in Figure 3G. Up-regulation of *GATA2*, *GATA3* and *TBX3* begins from day 1, other markers from day 2 or day 3. *TFAP2C*, *TBX3* and *HAVCR1* are expressed in naïve cells and further up-regulated in PD and PD+A83 but unchanged or down-regulated in N2B27 or A83 alone. They are lowly expressed or absent in primed hPSCs (Figure S3J)

These transcriptome analyses confirm that human naïve stem cells in PD+A83 do not undergo formative transition but differentiate via a distinct path into trophectoderm.

Reversion from naïve epiblast to early ICM

We explored further the trajectory towards trophoblast adopted by naïve cells. We focused on clusters 3 and 4, comprising genes that are transiently up-regulated in the PD+A83 time course. We examined distribution in the human embryo and found enrichment not in early trophectoderm, but in ICM (Figure 4A). To confirm this finding we carried out a reciprocal analysis by clustering dynamically expressed genes in the human embryo (Figure 4B) and examining overlap with the PD+A83 clusters (Figure 4C). Embryo cluster 3, enriched in ICM, shows the highest overlap with clusters 3 and 4 from in vitro differentiation. To further inspect the expression of human embryo cluster 3 genes in the in vitro differentiation time course, we implemented bootstrap Spearman correlation analysis. This confirmed transient up-regulation, peaking on day 2 (Figure 4D). Interestingly, when we performed cluster comparison with the early mouse embryo we observed no overlap between the ICM and

human in vitro differentiation (Figure S4A-D). Furthermore, cluster 3 and cluster 4 in the PD+A83 time course contain multiple transcription regulators that are specifically up-regulated in human ICM but show no specific expression in mouse ICM (Figures S4E,F).
We identified 1478 down-regulated and 1017 up-regulated genes in ICM compared with trophectoderm and EPI and examined representation in the in vitro differentiation clusters (Figure 4E). Cluster 3/4 contains the fewest genes down-regulated in ICM and is enriched for ICM up-regulated genes. Mean expression values confirm that genes with reduced expression in ICM are also down-regulated across the in vitro time course until day 5 and conversely that genes over-represented in ICM are up-regulated on day 1 and 2 in vitro and down-regulated on day 5 (Figure 4E).

We noted that PD+A83 clusters 3 and 4 include *GATA6, ESRRB* and *CDX2* (Figure 4F). GATA6 is associated with primitive endoderm, but is initially widely expressed in ICM, then down-regulated in epiblast in both mouse and human (Boroviak et al., 2018). ESRRB is of interest because in the human embryo, unlike mouse, *ESRRB* mRNA is barely detectable in naïve epiblast cells (Stirparo et al., 2018; Takashima et al., 2014) but it is present in ICM and early trophoblast cells. *CDX2* is a mouse trophoblast determinant but in human is expressed in ICM and early trophectoderm, then down-regulated in both late trophoblast and epiblast. By RT-qPCR we confirmed transient up-regulation of *ESRRB*, *GATA6* and other ICM-enriched genes on day 1-3 of PD+A83 treatment (Figure S4G).

These analyses suggest that naïve stem cells cultured in PD or PD+A83 may first revert to an ICM-like state from which the trophectoderm lineage is directly accessible. We reasoned that the two inhibitors may not be required continuously if their primary effect is to mediate such reversion. Indeed we found that 48 hours in PD+A83 resulted in efficient induction of GATA3:mKO positive cells by day 3 (Figure 4G). RT-qPCR confirmed expression of trophectoderm markers (Figure S4H). GATA3 positive cells were also induced following only 24 hours in inhibitors, albeit with reduced efficiency. Since culture in A83 alone has minimal capacity to induce trophectoderm (Figure 2E), we examined treatment with PD only for 24 hours followed by culture in N2B27. After 3 days we observed a minor fraction of mKO-high cells (Figure 4H). However, addition of A83 from 24 hours increased the yield of mKO positive cells to over 60%. Conversely, exposure to activin almost entirely suppressed the emergence of positive cells (Figure 4H, S4I). Immunostaining of cultures treated with PD for 24 hours then for 3 days with A83 confirmed the presence of GATA3 protein in many cells, exclusive with persisting OCT4 or NANOG expression (Figure 4I). We also detected a minority of cells that stained exclusively for SOX17 or for FOXA2, suggestive of primitive endoderm identity.

Collectively, these observations suggest that culture in PD alone causes naïve stem cells to revert to an ICM-like cell population that can differentiate into trophectoderm. The data also suggest that the enhancing effect of Nodal inhibition by A83 is mainly to promote trophectoderm lineage entry after ICM reversion.

Bone morphogenetic protein induces amnion and not trophectoderm from primed hPSC

Previous reports have suggested that primed hPSCs may form trophoblast upon treatment with BMP (Amita et al., 2013; Xu et al., 2002). More recently, human extended potential stem cells (hEPSC) have been reported to produce trophoblast more efficiently, a response

enhanced by BMP (Gao et al., 2019). BMP signaling is barely detected in naïve cells (Figure S5A) and we found that inhibition of the BMP pathway does not impede their differentiation into trophectoderm (Figure S5B). We therefore sought to clarify the relationship between naïve and primed hPSC or hEPSC lineage trajectories.

We first examined the transcriptomes of undifferentiated hEPSCs reported by two groups (Gao et al., 2019; Yang et al., 2017) and found that they are distinct from naïve stem cells and more similar to conventional primed hPSCs (Figure S5C). We took advantage of recent data from prolonged 3D culture of human embryos (Xiang et al., 2019) to re-examine the identity of naïve, conventional and extended potential hPSC. We computed the PCA for epiblast stages from the embryo together with the stem cell cultures (Figure 5A). The 3D plot corroborates the close relationship between naïve stem cells propagated in PXGL and E7 epiblast (Bredenkamp et al., 2019b; Stirparo et al., 2018). Further, the analysis demonstrates that hEPSCs are similar to conventional hPSCs with gene expression features of post-implantation epiblast from day 10 or later.

Then we interrogated data from hEPSC differentiation into putative trophoblast cells in a medium cocktail containing A83 (Gao et al., 2019). We first compared the hEPSC differentiation time course with naïve cell differentiation and averaged scRNA-seq samples from the human blastocyst (Petropoulos et al., 2016). The PCA computed using variable genes in the human embryo shows that conventional hPSCs and hEPSCs follow a common initial trajectory, with hEPSCs progressing further (Figure S5D). Notably, however, this path is distinct from naïve cell differentiation and does not converge on trophectoderm. We inspected genes differentially expressed in ICM and found no correspondence during hEPSC differentiation (Figure 5B), in marked contrast to the naïve cell time course (Figure 4F).

We speculated that hEPSC differentiation might represent another extraembryonic lineage, the amniotic epithelium, which forms from post-implantation epiblast. Recently it has been reported that conventional hPSC can differentiate into amnion-like cells in culture (Zheng et al., 2019). We extracted genes with enriched expression in amnion-like cells and examined their distribution in the PCA. The intensity overlay is concentrated in the area occupied by the endpoint of hEPSC differentiation (Figure S5E). We also performed an analysis using early and late amnion samples from *Macaca* embryo cultures (Ma et al., 2019). The intensity distributions were again concentrated in the region of the differentiated hEPSCs (Figure S5F). Consistent with these observations, standardized expression of amnion enriched genes (Zheng et al., 2019) was higher in differentiating hEPSC than in naïve cells treated with PD+A83 (Figures S5G).

We next recomputed the PCA using the top 100 differentially expressed genes from the human embryo culture transcriptome (Xiang et al., 2019) (Figure 5C). This analysis again resolved distinct trajectories for naïve cells compared to conventional hPSCs and hEPSCs. Naïve cell differentiation in PD+A83 proceeded via similarity to ICM and culminated in close proximity to early cytotrophoblast. In contrast conventional hPSCs and hEPSCs followed a common trajectory to transitional amnion (EPI to AME) (Xiang et al., 2019) with hEPSCs proceeding further towards amniotic epithelium. We examined global correlation between embryo stages and in vitro naïve or hEPSC differentiation by bootstrapped Spearman analysis. The traced heatmaps show that naïve cell differentiation progressed from high

initial correlation with pre-implantation epiblast via ICM to trophoblast (Figure 5D). In contrast hEPSCs proceeded from highest starting correlation with post-implantation epiblast via EPI to AME transition, then to amnion.

To obtain independent validation of distinct differentiation behaviors we first converted naïve HNES1 GATA3:mKO cells to primed status by in vitro capacitation in N2B27 plus XAV followed by expansion in activin, FGF and XAV (Rostovskaya et al., 2019). We compared the response of the isogenic naïve and primed reporter cells to PD, A83 or both inhibitors. mKO expression in naïve cells in PD became detectable on Day 2 and was almost uniformly high on Day 3. A83 accelerated these dynamics. In contrast, addition of BMP had no significant effect (Figure 5E). For the capacitated cells, reporter expression was negligible on day 2 in any condition, but the cultures became predominantly mKO positive by Day 3 in PD+A83 or by Day 5 in PD. Notably, however, expression of the reporter was at least 10-fold lower relative to differentiation from naïve cells. Addition of BMP modestly accelerated the dynamics but not the level of reporter activation. We also noted that A83 or BMP2 alone induced low expression of mKO in a fraction of primed reporter cells (Figure S5H), consistent with previous reports (Amita et al., 2013; Gao et al., 2019). RT-qPCR analysis confirmed modest up-regulation of *GATA3* and also *GATA2* (Figure 5F). However, other early and late trophectoderm markers were barely expressed. Instead we detected upregulation of a panel of amnion markers in primed but not naïve cell differentiation (Zheng et al., 2019) (Figure 5G). Finally, we compared requirements for BMP signalling. The BMP receptor inhibitor LDN had no effect on activation of GATA3:mKO and expression of trophectoderm markers in naïve cells, but completely blocked up-regulation of GATA factors and amnion markers in primed cells (Figure 5H,S5I). Instead, treatment with LDN together with PD+A83 diverted primed cells into the neural lineage (Figure S5J). Thus BMP is critical for the amnion fate from primed hPSC but is not involved in trophectoderm differentiation from naïve cells.

We generated RNA-seq data from capacitated GATA3:mKO cells and from conventional H9 cells treated with PD+A83, with or without BMP. For both cell lines in either condition, the PCA trajectory proceeds towards the area of epiblast-amnion transition with no proximity to ICM or trophoblast (Figure S5K).

Collectively these analyses demonstrate that naive and primed hPSCs differentiate along BMP-independent and BMP-dependent trajectories respectively into distinct trophectoderm and amnion fates.

OCT4, SOX2 and NANOG suppress trophectoderm lineage potential, which is enabled by YAP and TFAP2C

In mouse ES cells deletion of *Oct4* or *Sox2* results in up-regulation of *Cdx2* and differentiation to trophoblast-like cells. In contrast, deletion of *Nanog* provokes differentiation to primitive endoderm-like cells with no evidence of trophectoderm (Mitsui et al., 2003). To determine whether these relationships are conserved in human naïve cells we performed targeted mutagenesis using CRISPR/Cas9.

We first mutated *OCT4*, *SOX2* and *NANOG* in HNES1 cells by transfection with Cas9 and gRNA ribonucleoprotein complexes (RNP). Consistent with their expected essential roles, we observed reduced naïve colony numbers in PXGL for all three genes (Figure S6A). To

assess the fate of targeted cells, we performed immunostaining on cells either maintained in PXGL or exchanged to N2B27 after 24hrs. On day 5 after transfection, clusters of GATA3 positive cells were apparent in each case (Figure 6A,B). Notably, in PXGL GATA3 immunoreactivity was specific to those cells in which the targeted transcription factor was depleted. RT-qPCR confirmed up-regulation of early and late trophectoderm markers *GATA2*, *HAVCR1* and *SLC28A3* (Figure 6C). We repeated RNP targeting in the GATA3 reporter line and detected mKO bright cells in N2B27 (Figure 6D).

RNP transfection efficiency is limiting for mutation frequency. We therefore introduced a constitutive Cas9 expression construct into the *AAVS1* locus in naïve GATA3:mKO reporter cells. We then used piggybac (PB) transposition to integrate gRNA expression constructs containing a selectable marker. We designed distinct gRNAs from those used in the RNP transfections. After transfection and selection there was a massive reduction in colony numbers in PXGL for all three knockouts (Figure S6B) with activation of the mKO reporter in N2B27 (Figure 6E). mKO activation was further enhanced by A83. Notably, the *SOX2* knockout quantitatively recapitulated induction of mKO by PD+A83. We substantiated up-regulation of trophoblast markers by RT-qPCR (Figure S6C)

Trophectoderm differentiation in response to *NANOG* targeting is at variance with the phenotype in mouse ES cells, suggesting a function specific to human naïve cells. This prompted us to investigate whether NANOG can suppress trophectoderm formation. We introduced a doxycycline-inducible *NANOG* expression vector into GATA3:mKO naive cells. Treatment with DOX to induce NANOG prevented the appearance of mKO positive cells (Figure 6F,S6G) and suppressed up-regulation of trophectoderm markers in PD+A83 (Figure 6G).

Substantial expression of TFAP2C is a distinctive feature of human naïve stem cells and pre-implantation epiblast cells compared with mouse (Boroviak et al., 2018; Pastor et al., 2018; Stirparo et al., 2018). In mouse, TFAP2C is known as a trophoblast factor (Cao et al., 2015; Choi et al., 2012) and is barely expressed in ICM, EPI or ES cells. Forced expression, however, provokes trophoblast-like differentiation (Adachi et al., 2013; Kuckenberg et al., 2010). We targeted *TFAP2C* using Cas9/gRNA RNP. Transfected populations gave a decreased yield of mKO high cells (Figure 6H) and lower expression of trophectoderm markers in PD (Figure 6I). We evaluated *TFAP2C* knock out further using Cas9 expressing cells. We observed a reduction in colony numbers in PXGL (Figure S6B), consistent with a report that TFAP2C may be required for stable propagation of human naïve cells (Pastor et al., 2018). The impact on trophectoderm induction was marked. After 4 days in PD, fewer than 20% of cells showed appreciable expression of mKO (Figure 6J). In PD+A83 this effect was less, but still a two-fold reduction in the proportion of mKO bright cells. RT-qPCR detected diminished up-regulation of *GATA2* and *GATA3* with maintained expression of *NANOG* (Figure 6K).

TEAD1 and TEAD3 are both up-regulated during naïve cell differentiation into trophectoderm (Figure 1C, 3G) and their activator YAP plays a key role in mouse trophectoderm specification (Nishioka et al., 2009). We therefore targeted *YAP*. In contrast to other perturbations, we did not observe overt deficiency in naïve colony formation in PXGL. However, In PD *YAP* mutation reduced both mKO activation and up-regulation of *GATA3* and *GATA2* (Figure 6J,K). In PD+A83, mKO and trophectoderm markers were expressed

but to a lower level than in control transfectants. Dual targeting of YAP and TAZ strongly suppressed trophectoderm induction in either PD or PD+A83, although naïve cell propagation in PXGL was impaired.

Human ICM retains trophectoderm lineage plasticity

Ex vivo culture conditions may corrupt cell identities or alter developmental potential. We therefore examined whether the plasticity of naïve stem cells to generate trophectoderm is an authentic feature of ICM cells in expanded human blastocysts. Frozen blastocysts (E5 or E6) were thawed and cultured as necessary for formation of well-expanded blastocysts (E6 or E7) (Figure 7A). We immunostained a cohort of embryos with distinct ICMs for GATA3 and found intense nuclear staining localised to the outer layer of trophectoderm cells, as expected (Figure 7B). To obtain ICMs for explant culture we performed immunosurgery. Lysis of the entire trophectoderm layer and subsequent removal by trituration were monitored by light microscopy (Guo et al., 2016). Isolated ICMs were maintained intact and cultured on laminin-coated dishes in N2B27 medium. Large flattened cells with trophectoderm morphology outgrew within 48 hours. Immunofluorescence staining confirmed prevalence of GATA3 positive cells (Figure 7C). Trophectoderm generation from human naïve cells is BMP independent. Therefore, we applied the BMP inhibitor LDN and used time lapse imaging to monitor outgrowth (Supplemental Movies 3 and 4). ICMs proliferated and by 48 hours formed explants containing spread cells (Figure 7D). Immunofluorescence staining revealed GATA3 positive cells surrounding a cluster of SOX2 positive epiblast cells. We also observed SOX17 staining indicative of primitive endoderm cells adjacent to the epiblast cells. Expression of the three markers was exclusive, consistent with proper lineage segregation. We also cultured ICMs in N2B27 with PD+A83 and LDN. Although cell viability was lower in this condition, GATA3 positive outgrowths formed together with clusters of SOX2 and SOX17 expressing cells (Figure 7E).

These findings demonstrate that ICMs from expanded human blastocysts can readily regenerate trophoectoderm. Interestingly, trophectoderm differentiation occurred in N2B27 without the requirement for MEK inhibition manifest by naïve stem cells. This difference may reflect adaptation to self-renewal in the cultured cell lines. Alternatively, development within the human embryo may be asynchronous such that in expanded blastocyts some cells retain ICM features while others have progressed to epiblast. We inspected human embryo scRNA-seq data (Xiang et al., 2019) and indeed found co-existence of cells with ICM or epiblast profiles within individual embryos at E7 (Figure 7F).

These findings confirm that human ICM cells retain plasticity to form trophectoderm. Together with the stem cell analyses they support a model of pluripotent lineage progression in which the key change in competence occurs with the formative transition when amnion, germline and germ layer potency are gained and trophectoderm capacity is extinguished (Figure 7G). The coincident gain in BMP responsiveness is noteworthy.

DISCUSSION

In mammalian embryogenesis the initial formation of extraembryonic lineages must be accompanied by allocation and preservation of epiblast cells that can subsequently generate

the foetal lineages and germline. In mouse, fate mapping and molecular studies have substantiated a strict lineage bifurcation such that, even prior to specification of epiblast, ICM cells become refractory to trophectoderm specification. Our results expose a different scenario in human, whereby trophectoderm lineage potential is maintained as ICM and epiblast identities are established. Human naïve epiblast stem cells can differentiate with high efficiency into authentic trophectoderm and ICMs extracted from well-expanded human blastocysts robustly regenerate trophectoderm. However, this potential is exclusive to ICM and naïve epiblast. hPSCs representative of post-implantation epiblast have extinguished trophectoderm lineage capacity. Although they can undergo differentiation into epithelial cells superficially similar to trophoblast, global profiling points to a later extraembryonic lineage, amniotic epithelium. These findings resolve contradictory claims and re-establish a lineage hierarchy consistent with in vivo development. More broadly, our results identify developmental plasticity in the human ICM and naïve epiblast that is not precedented by studies in mouse, and may confer high regulative capacity in the human embryo.

Development is a continuous unidirectional process in which cells progress through transitions that are generally considered irreversible. The propensity uncovered here for dedifferentiation of human naïve epiblast stem cells to the preceding ICM state is therefore not predicted. Nonetheless, reversion of identity is strongly evidenced by global transcriptome analysis with 1017 and 1478 genes showing significant up- or down-regulation consistent with a switch from E6/7 epiblast to E5 ICM. Dedifferentiation occurs within 24 hours and appears essential for trophectoderm lineage potential because GATA3 positive cells are barely generated without this period of PD treatment. The early ICM does not feature general up-regulation of trophoblast-associated genes. TFAP2C is expressed, although at a similar level to naïve epiblast. Interestingly, CDX2 mRNA is also expressed in human ICM and reverted naïve epiblast cells. A further feature of note is that *ESRRB* is activated upon reversion to ICM status while *NANOG* is down-regulated (Stirparo et al., 2018). This relationship is opposite to that in mouse, where *Esrrb* is up-regulated in naïve epiblast cells directly downstream of *Nanog* (Festuccia et al., 2012). In general, human and mouse ICM show considerable differences in gene expression.

We find that human ICM cells from mature blastocysts can readily generate trophectoderm, unlike mouse ICM, which loses this potency at the mid-blastocyst stage. Human ICM differentiation to trophectoderm does not require MEK inhibition. This may indicate that trophoblast regenerates from cells persisting in a pre-epiblast state at E6/7, consistent with the observed asynchronous ICM to epiblast developmental progression within the human blastocyst (Xiang et al., 2019). Whether epiblast to ICM reversion may be specific to the in vitro stem cell context or can also occur in vivo remains to be determined. Nonetheless, whether from persisting earlier ICM cells or from naïve epiblast, mature human blastocysts have prolonged potential to reform trophoblast, and potentially to reconstitute an intact embryo after loss of trophoblast.

Human naïve stem cells are generated and maintained most robustly in PXGL medium (Bredenkamp et al., 2019b; Guo et al., 2017), which comprises three small molecule inhibitors of MEK/ERK, Wnt, and aPKC signaling respectively. These requirements, which differ from those for mouse ES cell propagation (Dunn et al., 2014; Ying et al., 2008), can be better understood in light of our findings. In both species, MEK/ERK inhibition prevents formative pluripotency transition to post-implantation epiblast (Smith, 2017). In human,

however, PD also promotes dedifferentiation from naïve epiblast to ICM identity. Wnt and aPKC inhibition block this reversion and consequent access to trophectoderm, thereby consolidating self-renewal. Thus, triple blockade of signaling input into the gene regulatory network is required to constrain cells specifically in the naïve epiblast state. In addition, autocrine activation of SMAD2 and SMAD3 by Nodal contributes to long-term stability, a further distinction from mouse ES cells dictated by the requirement to suppress trophectoderm lineage potential. Of note, Nodal and GDF3 are highly expressed in naïve epiblast in the human embryo (Stirparo et al., 2018).

Deletion of core pluripotency factors reveals shared and species-specific features. *OCT4* and *SOX2* knockouts trigger trophoblast formation from naïve stem cells in mice and human. Interestingly *NANOG* deletion also releases trophectoderm differentiation in human naïve cells, whereas mouse ES cell mutants either remain undifferentiated or become primitive endoderm (Chambers et al., 2007; Mitsui et al., 2003). NANOG may have a conserved function to establish and consolidate naïve epiblast identity, but greater permissiveness for dedifferentiation results in ICM and thence trophectoderm upon mutation in human. Of note, zygotic deletion of mouse *Nanog* results in failure of ICM progression to epiblast and retained capacity for trophectoderm differentiation (Silva et al., 2009).

Requirements for YAP and TFAP2C in mediating trophectoderm lineage segregation appear conserved. However, whereas TFAP2C is trophoblast specific in the mouse blastocyst, in human it is expressed in ICM and epiblast and plays a role in maintenance of naïve stem cell self-renewal (Pastor et al., 2018). Forced expression of *Tfap2c* in mouse ES cells provokes trophoblast-like differentiation (Adachi et al., 2013; Kuckenberg et al., 2010). We speculate that the dual functionality of TFAP2C may be central to the lineage plasticity of human naïve cells.

There have been several reports that conventional hPSCs and hEPSCs can differentiate into epithelial cells that exhibit some markers of trophectoderm (Amita et al., 2013; Gao et al., 2019; Xu et al., 2002; Yang et al., 2017). However, global transcriptome assessment indicates a trajectory unrelated to pre-implantation development. Instead, the path is towards a post-implantation extraembryonic lineage, amnion. Significantly, BMP-signaling is essential for amnion-like differentiation from primed stem cells, whereas trophectoderm induction from ICM or naïve stem cells is insensitive to this pathway. Indeed, naïve cells show negligible SMAD activation when exposed to BMP, indicating that responsiveness to this pathway is only acquired during capacitation.

Recently advances have been made in human embryo culture (Deglincerti et al., 2016; Shahbazi et al., 2016; Xiang et al., 2019). However, availability of human embryos is limited and quality variable. Our study illustrates that human naïve stem cells present a complementary, experimentally convenient, model for delineating the molecular mechanisms of early lineage segregation and uncovering species-specific features. Considerable current interest also focuses on the production of synthetic embryos or blastoids, achieved by combining mouse ES cells together with trophoblast stem cells (Rivron et al., 2018; Sozen et al., 2019). We speculate that the competency to produce all three primary lineages may enable reconstitution of a blastocyst entity with full developmental potential wholly from naïve stem cells. In principle, this could provide a simple system for elucidating principles of

self-organisation, taking advantage of high throughput genome engineering in clonogenic stem cells.

In summary, our results establish that the mouse paradigm of early lineage segregation is not strictly adhered to in human, and that human naïve cells have unprecedented plasticity with intrinsic potential for dedifferentiation. Interestingly, it has been reported that human trophectoderm at E5 can regenerate an ICM population (De Paepe et al., 2013). High regulative flexibility may be an important mechanism for safeguarding human embryos. In the context of assisted conception, this could explain why viable pregnancies can ensue from embryos that are judged morphologically to be of lower quality, or incur cell loss during blastocyst freezing and thawing. It will also be of interest to determine whether regenerative plasticity is a more widespread feature in mammalian embryos, that may have eroded in rodents due to their high and rapid reproduction with associated reduced investment in an individual embryo.

Acknowledgments
We are grateful to Kosuke Yusa and Bahar Mirshekar for the AAVS1-Cas9 targeting vector. We thank Rosalind Drummond and Tao Huang for technical support. Nicholas Bredenkamp, Reo Shoshi and Chonghyun Cha contributed to cell line modification. Maike Paramor generated sequencing libraries. Peter Humphreys and Darran Clements supported imaging. This research was funded by the Medical Research Council (MRC) of the United Kingdom. The Wellcome-MRC Cambridge Stem Cell Institute receives core support from Wellcome and MRC. AS is a Medical Research Council Professor

Author Contributions
Conceptualization, GG, AS; Methodology, GG; Investigation, GG, SS, JY, SM, JC, MAL, BNÖ, JN; Formal analysis, GGS; Writing, AS, GG; Supervision JN, GG, AS

Declaration of Interests
AS and GG are inventors on a patent application relating to human naïve stem cells filed by the University of Cambridge.

FIGURE LEGENDS

Figure 1. Human naïve stem cells differentiate to trophectoderm

A. Images of naïve stem cells differentiating in N2B27 with PD.
B. RT-qPCR assay for core pluripotency, post-implantation epiblast, and trophectoderm markers after 5 days culture in indicated conditions.
C. Time course of trophectoderm marker expression in presence of PD.
D. Immunostaining of cells cultured for 3 days in PD.
E. GATA3:mKO reporter fluorescence after 3 days in PD.
F. Flow cytometry analysis of GATA3:mKO expression after 3 days in N2B27 or with PD. Scale bars, 100μM

Figure 2. Inhibition of Nodal enhances trophectoderm differentiation

A. Flow cytometry analysis of GATA3:mKO cells in PXGL with or without A83 for 4 passages.
B. Fluorescence microscopy of GATA3:mKO positive cells in PXGL plus A83. Scale bar, 100μM
C. Phospho-Smad2/3 immunoblot on conventional H9 hPSC and naïve cells in indicated conditions.
D. Phospho-Smad2/3 immunoblot on naïve cells in absence or presence of A83-01
E. Flow cytometry analysis of GATA3:mKO cells in N2B27 or with A83 for 3 days.
F. Flow cytometry analysis of GATA3:mKO cells in PD or PD+A83 for 3 days.
G. RT-qPCR assay of *GATA2* and *GATA3* expression over time in indicated conditions.
H. Phase contrast and fluorescence time lapse stills of GATA3:mKO cells in PD+A83. (See also Supplemental Movie 1). Scale bar, 100μM

Figure 3. Global transcriptome evaluation of trophectoderm differentiation

A. Principal component analysis computed with all expressed protein coding genes (log2 expression >0, n=19450). Two-dimensional kernel density estimation of genes contribution for primed hPSC enriched genes, red dotted lines (primed VS naïve, Stirparo et al., 2018, log2FC > 2, n=1778) and trophectoderm (TE) enriched genes, purple dotted lines (lateTE vs EPI log2FC > 2, n=409).
B. One-way hierarchal cluster of differentially expressed genes between late TE (blue) and epiblast (EPI) (red) (top 200 UP and DOWN ranked genes, rank is the product of –log10padj and FC) in HNES1.
C. Heatmap of zscore centered value for clusters identified in PD+A83 time-course (cluster1: 1787, cluster2: 2594, cluster3: 1872, cluster4: 1227, cluster5: 2055, cluster6: 1839).
D. Venn diagram with percentage of EPI UP-regulated genes (EPI vs late TE, n=1632) for each cluster.
E. Venn diagram with percentage of TE UP-regulated genes (EPI vs late TE, n=910) for each cluster.
F. Embryo standardized score distribution for genes in cluster1/2 and cluster 4/5.
G. log2 expression value for TE enriched genes in A83, N2B27, PD and PD+A83 conditions.

Figure 4. Naïve cell reversion to an ICM-like state

A. Embryo standardized score distribution for genes in cluster3/4 in PD+A83 time-course.

B. Heatmap of zscore centred value for clusters identified in EPI, ICM, early TE, late TE embryo stages.
C. Heatmap with overlapping genes for each cluster identified in PD+A83 time-course (PXGL, d1,d2,d3,d5, log2expression > 1) and human embryo (EPI, ICM, earlyTE,lateTE, log2expression >1). Only genes in cluster with a membership > 0.3 (fuzzy cluster) were considered.
D. Bootstrapped Spearman correlation (100 iterations, number of genes=1000) between PD+A83 time-course (PXGL, d1,d2,d3,d5, log2 expression > 1) and embryo early ICM.
E. Top: Venn diagrams of differentially expressed genes between EPI/ICM and TE/ICM. Bottom: Cluster enrichment for genes up/down regulated in ICM, and zscores of mean expression in PD+A83 time course
F. log2 expression value for ICM enriched genes in PD+A83 condition (cluster 3/4 PD+A83 and cluster3 in human embryo). Log2 expression in EPI, ICM, early TE, late TE embryo stages.
G. Flow cytometry analysis of GATA3:mKO cells exposed to PD+A83 for indicated periods then in N2B27.
H. Flow cytometry analysis of GATA3:mKO cells in PD for 24 hours and followed by culture with A83 or Activin A for 72 hours.
I. Immunostaining for indicated transcription factors of cells cultured in PD03 for 24hrs followed by 72 hours in A83.

Figure 5. Alternative fates of naïve or primed stem cells

A. 3D PCA of human naïve and conventional primed cells (Stirparo et al., 2018), naïve cells from this study, hEPSCs (Gao et al., 2019; Yang et al., 2017), and human embryo in vitro development (Xiang et al., 2019), computed with variable genes in embryo development, n=1517.
B. Top: venn diagrams of differentially expressed genes between EPI/ICM and TE/ICM. Bottom: Cluster enrichment for genes up/down regulated in ICM and zscore in hiEPSC timecourse
C. PCA of A83, N2B27, PD and PD+A83 time courses for naïve cells (HNES1 and cR-H9), average of human embryo stages, hEPSC (d0,d2,d4,d6,d9,d12), primed hPSC (H9, HNES1) time courses in N2B27, PD+A83 and PD+A83+BMP2, and human embryo in vitro development (top 100 most variable genes in human development).
D. Traced heatmap computed with median of bootstrapped Spearman correlation (iteration 100, number of genes=50).
E. Flow cytometry analysis of naïve and primed GATA3:mKO cells in PD03, PD03+A83 or PD03+A83+BMP2.
F. RT-qPCR assay of GATA2 and GATA3 expression over time in indicated conditions comparing naïve and primed cells.
G. RT-qPCR assay of amnion and trophectoderm markers in naïve (N) and primed (P) cells after 5 days culture in PD03+A83 with BMP2
H. Flow cytometry analysis of naïve and primed GATA3:mKO cells in PD03+A83 with or without BMP inhibitor LDN.

Figure 6. Effect of genetic perturbations on trophectoderm lineage potential

A. Immunostaining for indicated markers after Cas9 RNP targeting of *OCT4*, *SOX2* or *NANOG* in HNES1 cells and maintenance in PXGL.
B. As A, but culture in N2B27

C. RT-qPCR assay of trophectoderm marker expression after Cas9 RNP targeting of indicated genes in HNES1 cells.
D. Flow cytometry analysis of GATA3:mKO cells after Cas9 RNP targeting and culture for 4 days in N2B27.
E. Percentage of mKO positive cells gated as in D after targeting indicated genes in Cas9 expressing cells.
F. Flow cytometry histogram of GATA3:mKO expression in N2B27 or PD+A83 with or without DOX induction of *NANOG*.
G. Heatmap of RT-qPCR gene expression values with and without DOX induction of NANOG in N2B27 or PD+A83.
H. GATA3:mKO flow cytometry plots after *GFP* (control) or *TFAP2C* targeting with Cas9 RNP and culture for 4 days in PD.
I. RT-qPCR assay of marker expression after *GFP* or *TFAP2C* targeting and culture as in H.
J. Reduction in GATA3:mKO activation measured by flow cytometry after targeting TFAP2C, YAP or YAP + TAZ in Cas9 expressing cells and culture for 4 days in indicated conditions. Relative to transfection with GFP gRNA vector.
K. RT-qPCR assay of marker expression after targeting and culture as in J.

Figure 7. Human embryonic ICM retains trophectoderm potential.
A. Phase images of human blastocysts classified as E5 on thawing and E6 after 24h in culture
B. Representative immunostaining of E6 human blastocyst for GATA3. White dashed line demarcates ICM population.
C. Phase image of ICM outgrowth after 72 hours in N2B27 and immunofluorescent images following fixation and staining.
D. Time lapse stills at beginning and end of 48 hour ICM outgrowth culture in N2B27 plus LDN. Immunostaining for GATA3, SOX17 and SOX2 at 48 hours.
E. ICMs cultured in N2B27 with PD+A83 and LDN, imaged and stained as in B.
F. Pie chart of number of cells and their associated lineage in individual E7 embryos (data from Xiang et al., 2019).
G. Schematic summary of findings showing that naïve cells can progress to formative pluripotency with gain of lineage competence including amnion, or may revert to ICM and thence differentiate to trophectoderm.

SUPPLEMENTAL FIGURE LEGENDS

Figure S1. Trophectoderm differentiation, Relates to Figure 1

A. RT-qPCR assay for primitive endoderm markers after 5 days culture in indicated conditions.
B. RT-qPCR assay for naïve and trophectoderm markers after 5 days culture in indicated conditions.
C. RT-qPCR assay for trophectoderm markers after 5 days culture in N2B27 or PD03 of three human naïve PSC lines; HNES1, cR-NCRM2 and cR-Shef6
D. RT-qPCR assay for core pluripotency and trophectoderm marker expression in mouse ES cells cultured for 3 days in indicated conditions.

Figure S2. Nodal signaling in human naïve stem cells, Relates to Figure 2

A. RT-qPCR assay for *GATA2* and *GATA3* expression in human naïve cells cultured in PXGL with A83. X-axis indicates passages in PXGL with A83.
B. Log2 expression value for Nodal and TGFβ ligands and receptors in human naïve and primed cells.

Figure S3. Global transcriptome evaluation of trophectoderm differentiation, Relates to Figure 3

A. One-way hierarchal clustering of E5 single cell (Petropoulos et al., 2016) computed with lineage marker genes (Stirparo et al., 2018).
B. One-way hierarchal clustering of E6 and E7 single cell (Petropoulos et al., 2016) computed with lineage marker genes (Stirparo et al., 2018).
C. Density distribution of each cells in each cluster along PCA dimension 1.
D. PCA computed for all the genes expressed in early ICM (cluster 4, FIG.S2A), early TE (cluster 5, Figure S2A), epiblast (Stirparo et al., 2018) and late TE (cluster 2, Figure S2 B, C, D). In red, cells expressing more than 6000 genes.
E. Principal component analysis for all the filtered cells at each stage (log2expression >0, n=18694)
F. PCA for all filtered cells, computed with differentially expressed genes in human embryo (n= 4507).
G. One-way hierarchical cluster of differentially expressed genes between late TE (blue) and EPI (red) (top 200 UP and DOWN ranked genes, rank is the product of –log10padj and FC) in cR-H9.
H. Binarized cumulative sum (if log2 expression > 5→ 1, otherwise 0) for genes in cluster 1/2 or cluster 5/6 in each embryo stage.
I. PCA computed with all orthologues (average macaque dataset, ICM, EPI, postE, postL, early TE, late TE and post PA.TE, Nakamura et al., 2016) of expressed protein coding genes (log2 expression in time-course > 0 & orthologues, n= 12992).
J. log2 expression of indicated genes in naïve and primed cells by RNA-seq.

Figure S4. Naïve cell reversion to an ICM-like state, Relates to Figure 4

A. Heatmap of z-score centered value for clusters identify in EPI, ICM, early TE, late TE embryo stages in mice
B. Heatmap with overlapping genes for each cluster identified in PD+A83 time-course (PXGL, d1, d2, d3, d5, log2expression > 1) and mouse embryo (EPI, ICM, early TE, late TE, log2 expression >1).
C. Mouse embryo standardized score distribution for genes in cluster 3/4 in PD+A83 time-course.
D. Bootstrap of Spearman correlation (100 iterations, number of genes=1000) between PD+A83 time-course (PXGL, d1, d2, d3, d5, log2 expression > 1) and mouse embryo early ICM.
E. Heatmap of standardized score (per condition) of genes in cluster 3 in PD+A83, cluster 3 in human embryo development and cluster 1-2-4-5-6 and not expressed (log2 expression >1) in mouse embryo development.
F. Heatmap of standardized score (per condition) of genes in cluster 4 in PD+A83, cluster 3 in human embryo development and cluster 1-2-4-5-6 and not expressed (log2 expression >1) in mouse embryo development. Only transcription factors and cofactors are visualized.
G. Heatmap of RT-qPCR expression of indicated genes enriched in cluster 3/4 in PD or PD+A83 time course.
H. RT-qPCR assay for trophectoderm markers of GATA3:mKO cells treated with PD+A83 followed by N2B27, as in Figure 4G
I. RT-qPCR assay for trophectoderm markers on GATA3:mKO cells treated with PD for 24 hours followed by culture with A83 or Activin A for 72 hours, as in Figure 4H.

Figure S5. Alternative fates of naïve or primed stem cells, Relates to figure 5

A. Phospho-Smad1/5 immunoblot on naïve cells and primed HNES1 cells in indicated conditions. P, PD03; B, BMP2, number indicates BMP2 concentration, ng/ml; L, human LIF.
B. Flow cytometry analysis of naïve GATA3:mKO cells in indicated culture conditions for three days. LDN, BMP inhibitor, LDN-193189.
C. Principal component analysis of human naïve and primed cells (Stirparo et al., 2018), naïve cells from this study, and hEPSC (Gao et al., 2019; Yang et al., 2017)(most variable genes, log2expression > 1, cv>0.5, n=3510).
D. PCA of A8301, N2B27, PD and PD+A83 conditions (HNES1 and cR-H9), average of human embryo stages (Petropoulos et al., 2016), and hEPSC time course (d0,d2,d4,d6,d9,d12), computed with differentially expressed genes in human embryo, n=4507.
E. Two-dimensional kernel density estimation of amnion genes in human differentiation in vitro (expression >1, Zheng et al., 2019).
F. Two-dimensional kernel density estimation of amnion genes in macaque in vivo (Ma et al., 2019).
G. Standardized expression of human amnion genes (Zheng et al., 2019) between PD+A83 time course (HNES1), hEPSC time course, and human early TE and late TE.
H. Flow cytometry analysis of naïve and primed GATA3:mKO cells in indicated culture conditions for 5 days.

I. RT-qPCR assay of GATA2, GATA3 and amnion markers in human naïve and primed cell differentiated in indicated conditions with or without LDN.
J. RT-qPCR assay for neural markers in human naïve and primed cells differentiated in indicated conditions for 5 days with or without LDN.
K. PCA of samples as in S5D plus capacitated HNES1 and conventional H9 hPSC time-courses in N2B27, PDA83 and PDA83/BMP2 (computed with differentially expressed genes in human embryo, n= 4507).

Figure S6. Effect of genetic perturbations on trophectoderm lineage potential, Relates to Figure 6

A. Images of alkaline phosphatase (AP) stained naïve cultures after indicated Cas9 RNP transfection and culture on PXGL on MEF for 4 days. Separate experiments shown were performed on HNES1 and HNES1-GATA3:mKO cells.
B. Images of AP stained human naïve cells after indicated gRNA transfection in Cas9 expressing GATA3:mKO cells. Cells were maintained in PXGL on MEF with puromycin selection for 7 days.
C. RT-qPCR assay following gRNA transfection in Cas9 expressing HNES1-GATA3:mKO cells for core pluripotency and trophectoderm markers after 4 days in indicated conditions.
D. Images of GATA3:mKO cells with or without DOX induction of *NANOG* in PD+A83 for 5 days.

Supplemental Movies 1 & 2

Phase contrast and fluorescence time lapse of GATA3:mKO cells cultured in PD+A83, 72hrs

Supplemental Movie 3

Time lapse of ICM explant in N2B27 with LDN, 48hrs

Supplemental Movie 4

Time lapse of ICM explant in PD+A83 with LDN, 48hrs

STAR METHODS

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Ge Guo, gg251@cam.ac.uk

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Ethics

Use of supernumerary human embryos in this research is approved by the Multi-Centre Research Ethics Committee, approval O4/MRE03/44, and licensed by the Human Embryology & Fertilisation Authority of the United Kingdom, research license R0178.

METHOD DETAILS

Human embryo culture

Supernumerary frozen blastocysts (E5 or E6) were thawed and cultured in N2B27 medium under mineral oil. Embryos that were well-expanded after thawing were judged to be E6 and processed immediately, while the majority of embryos were cultured for 24 hours for development to expanded blastocysts (E6 or E7). Immunosurgery was performed as described (Guo et al., 2016). In occasional cases in which complete lysis and removal of the trophectoderm could not be assured embryos were excluded from the study.

ICM explant culture

Isolated ICMs were placed on laminin-coated plates in N2B27 or N2B27 medium with chemical inhibitors for explant growth for 72 hours . For time-lapse imaging and confocal microscopy ICMs were cultured for 48 hours on Ibidis 24-well µ-plates coated with recombinant Laminin-111 E8 (Taniguchi et al., 2009)(kindly provided by Kiyotoshi Sekiguchi)

hPSC culture

Naïve and primed hPSCs were cultured in 5% O_2, 7% CO_2 in a humidified incubator at 37°C. Cell lines were maintained without antibiotics and confirmed free of mycoplasma contamination by periodic in-house PCR assay.

Naïve stem cells

Chemically reset (cR) and embryo-derived (HNES1) naïve stem cells were propagated in N2B27 with PXGL [1 µM PD0325901 (P), 2 µM XAV939 (X), 2 µM Gö6983 (G) and 10 ng/mL human LIF (L)] on irradiated MEF feeders as described (Bredenkamp et al., 2019b). ROCK inhibitor (Y-27632) and Geltrex (0.5µL per cm^2 surface area; hESC-Qualified, Thermo Fisher Scientific, A1413302,) were added to media during replating. Cultures were passaged by dissociation with Accutase (Biolegend, 423201) every 3-5 days.

Primed hPSC

Conventional primed hPSCs (H9, Shef6) were propagated on Geltrex in Essential 8 (E8) medium made in-house (Chen et al., 2011) or in AFX medium (N2B27 basal medium with 5ng/mL Activin A, 5ng/mL FGF2 and 2µM XAV).

Differentiation assay

Human naïve cells were plated in PXGL on Geltrex or Laminin at a 1:4 to 1:6 ratio. The next day, cultures were washed twice with PBS and medium exchanged to N2B27 with chemical inhibitors or cytokines. Medium was refreshed every day until assaying. Human primed cells were plated in AFX medium on Geltrex at 1:6 to 1:10 ratio and exchanged to assay conditions similarly to naïve cells. Medium was refreshed every day until assaying. Concentrations used in this assay: PD03 1µM, A83-01 1µM, BMP2 50ng/mL, LDN-193189 100nM, Activin A 20ng/mL.

Capacitation of human naïve cells
Cells were passaged once without feeders in PXGL medium then exchanged into N2B27 containing 2μM XAV for 10 days (Rostovskaya et al., 2019), followed by propagation in AFX medium.

Generation of GATA3:mKO reporter cell line
pGG195/GATA3 mKO targeting vector was designed to insert an *IresmKO-FRT-PGKNeobPA* cassette following the stop codon of *GATA3*. $1x10^6$ HNES1 naïve cells were transfected with 3μg *pGG195/GATA3mKO* and 3ug *px459/GATA3* gRNA using 100 μl Neon transfection kit. G418 (250ng/mL) selection was applied 2 days after transfection for 4 days. Cells were then harvested and transfected with *CAGGS-Flp* plasmid and plated in 2x10cm plates. Clones were picked 7 days after transfection and assayed for mKO expression in PXGL and PD03. Genomic DNA was prepared and correctly targeted heterozygous clones were confirmed by PCR amplification of the targeted junction and sequencing.

Inducible NANOG Expression
HNES1/GATA3mKO reporter cells were co-transfected with two *Piggybac* vectors carrying a *Tet3G*-inducible *NANOG* expression cassette and a *CAG-Tet3G-IresZeocin* cassette together with *PBase* plasmid. Two days after transfection, Zeocin (50 ug/mL) was applied for 5 days and individual clones were picked after 7 days. *NANOG* expression was induced with 10-20 ng/mL Doxycycline and assayed by qRT-qPCR.

CRISPR/Cas9 Knockout
Knockout by TrueGuide synthetic crRNA
True guide synthetic crRNA was purchased from Thermo Fisher Scientific, reconstituted and annealed with tracrRNA in RNA annealing buffer to generate double-stranded RNA duplex. The annealed RNA duplex was diluted in RNA storage buffer to 10μM stock. For each transfection, 1.2μl of the 10μM gRNA duplex was mixed with 300ng Cas9 protein and incubated at room temperature for 15 min before transfection. 10μL Neon transfection kit was used to transfect $1\text{-}1.5x10^5$ cells at 1150V, 30ms, 2 pulses. After transfection cells were plated without feeders in PXGL medium with ROCK inhibitor (Y-27632, 10μM). After 24 hours medium was exchanged to N2B27 or other differentiation assay medium for 4 days.
Knockout in Cas9 expressing naïve cells by Piggybac delivery of gRNA
gRNA oligos were synthesized and annealed to double-stranded DNA and cloned behind a U6 promoter (CML32) into a *Piggybac (PB*) vector containing a puromycin resistance gene. gRNA-expression plasmid was transfected together with *PBase* plasmid into HNES1 GATA3:mKO cells that had been engineered to constitutively express *Cas9* from the *AAVS1* genomic locus. Following transfection, cells were plated without feeders in PXGL with Y-27632 for 2 days then exchanged to medium for differentiation assay. Puromycin (0.5 μg/mL) was applied for at least 3 days to select cells with stable PB plasmid integration.

Reverse transcription and real-time PCR
Total RNA was extracted using ReliaPrep kit (Promega, Z6012) and cDNA synthesized with GoScript reverse transcriptase (Promega, A5004) and oligo(dT) adapter primers. TaqMan assays (Thermo Fisher Scientific) and Universal Probe Library (UPL) probes (Roche Molecular Systems) were used to perform gene quantification.

Immunostaining
Cells were fixed with 4% PFA for 10 min at room temperature and blocked/permeabilised in PBS with 0.1% Triton X-100, 5% Donkey serum for 30 min. Incubation with primary antibodies was overnight at 4°C. Wash was in 0.1% Triton X-100 twice, 10 min each time. Secondary antibodies were added for 1 h at room temperature. Whole embryo and embryo explant staining was performed as described (Guo et al., 2016)

Microscopy
Wide field images were taken using Leica DMI3000. Confocal images were taken using Leica SP-2 system. Time-lapse images were taken using Leica DMI6000 Matrix system fitted with a chamber with controlled temperature and CO2. Images were analyzed with Image J software.

Flow cytometry
Flow cytometry analysis was carried out on a CyAn ADP (Beckman Coulter) or BD LSR Fortessa instrument (BD Biosciences) with analysis using FlowJo software. DAPI staining was used to exclude the dead cell population.

Transcriptome sequencing
Total RNA was extracted from three biological replicate cultures of each cell line and time point using TRIzol/chloroform (Thermo Fisher Scientific, 15596018), and RNA integrity assessed by Qubit measurement and RNA nanochip Bioanalyzer. Ribosomal RNA was depleted from 1 µg of total RNA using Ribozero (Illumina kit). Sequencing libraries were prepared using the TruSeq RNA Sample Prep Kit (RS-122-2001, Illumina). Sequencing was performed on the Novaseq S1 or S2 platform (Illumina), according to the manufacturer's instructions.

QUANTIFICATION AND STATISTICAL ANALYSIS
Alignment was performed using the Genome build GRCh38 and STAR (Dobin et al., 2013) were used for aligning reads. Ensembl release 96 was used to guide gene annotation. After removal of inadequate samples, we quantified alignments to gene loci with htseq-count (Anders et al., 2014) based on annotation from Ensembl 96. Principal component, differential expression and cluster analyses were performed based on log2 expression values computed with custom scripts, in addition to the Bioconductor packages DESeq (Anders and Huber, 2010), FactoMineR (Lê et al., 2008) and MFuzz (Kumar and M, 2007). Gene density that contributed to PCA plot were calculated using kernel density estimation (MASS R package Venables, W. N. & Ripley, B. D. (2002))

For global analyses, we considered only genes with log2 expression > 0 (unless otherwise indicated) in at least one condition, not expressed genes were always omitted. Euclidean distance and average agglomeration methods were used for cluster analyses unless otherwise indicated. Human transcription factor and co-factors were downloaded from (http://bioinfo.life.hust.edu.cn/AnimalTFDB/). Time courses for H1.ESC, H1.EPSC and hiEPSC (Gao et al., 2019) were downloaded from array express (E-MTAB-7253) and re-aligned.

Cynomolgus monkey dataset (Nakamura et al., 2016) was kindly provided by Dr. Nakamura as an RPM table.

Expression values of ESC cultured in different human naïve/primed conditions and pre-implantation embryo single cell RNAseq were downloaded from (Stirparo et al., 2018). Embryo cells expressing less than 6000 genes with log2 expression > 1.5 were excluded. EPI, PrE and TE markers for cell stratification were selected according to Stirparo et al., 2018.

Human in vitro pre-gastrulating single cells were downloaded from Xiang et al., 2019 with GEO accession number GSE136447. High variable genes across EPI cells were computed according to the methods described (Boroviak et al., 2018; Stirparo et al., 2018). A non-linear regression curve was fitted between average log2 FPKM and the square of coefficient of variation (log CV^2); then, specific thresholds were applied along the x-axis (average $\log_2$ expression) and y-axis (log CV^2) to identify the most variable genes. Cells expressing less that 6000 genes with log2 expression > 1.5 were excluded from analysis.

Human amnion-like genes were downloaded from (Zheng et al., 2019) (Supplement Table 1, AMLC genes) and we selected only genes > 1 expression in amnion. Early and late AMLC genes in cynomolgus monkey were downloaded from (Ma et al., 2019)(Supplement Table 6).

DATA AND CODE AVAILABILITY

RNA-seq data from this study will be deposited in Gene Expression Omnibus

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Sox2	Santacruz	Cat#sc-365823
Nanog	R&D	Cat#AF1997
Oct4(C-10)	Santa Cruz	Cat#sc-5279
GATA3	abcam	Cat#ab199428
KLF17	Atlas Antibody	Cat#HPA024629
SOX17	R&D	Cat#AF1924
FOXA2	R&D	Cat#AF2400
pSmad 2	Cell Signalling	Cat#3101
pSmad 1/5/9	Cell Signalling	Cat#9511
b-tubulin	Abcam	Cat#ab6046
Bacterial and Virus Strains		
Biological Samples		
Chemicals, Peptides, and Recombinant Proteins		
MEK inhibitor PD0325901	ABCR	Cat#AB 253775
GSK3 inhibitor CHIR99021	ABCR	Cat#AB 253776
Tankyrase inhibitor XAV939	Cell Guidance Systems	Cat#SMS38-200
Gö6983	Bio-Techne	Cat#2285
ROCK inhibitor Y-27632	Merck Chemicals	Cat#688000-100MG
LIF	Made in-house	N/A
Activin A	Made in-house	N/A

Fgf2	Made in-house	N/A
BMP2	Made in-house	N/A
A83-01	Generon	Cat#A12358-50
LDN-193189	Axon Medchem	Cat#Axon 1509
TrueCut™ Cas9 Protein v2	ThermoFisher Scientific	Cat#A36498
Complete Culture Media and Cell Dissociation Reagent		
N2B27	Made in-house	N/A
AFX	Made in-house	N/A
E8	Made in-house	N/A
mTeSR1	Stemcell Technologies	Cat#05850
Accutase	Millipore	Cat#SCR005
Cell Attachment Proteins and Peptides		
Geltrex	Thermo Fisher Scientific	Cat#A1413302
Laminin	Millipore	Cat#CC095-5MG
Fibronectin	Millipore	Cat#FC010
Laminin 111-E8	From Kiyotoshi Sekiguchi	N/A
Critical Commercial Kits		
Alkaline Phosphatase Kit	Sigma-Aldrich	Cat#86R-1KT
Neon 10ul transfection kit	Thermo Fisher Scientific	Cat#MPK1096
Neon 100ul transfection kit	Thermo Fisher Scientific	Cat#MPK10096
Deposited Data		
Experimental Models: Cell Lines		
HNES1	Guo et al, 2016	N/A

HNES1_GATA3mKO	This study	N/A
cR-H9	Guo et al, 2017	N/A
cR-NCRM1	Guo et al, 2017	N/A
cR-Shef6	Guo et al, 2017	N/A
H9	WiCell	WA09
Experimental Models: Organisms/Strains		
Oligonucleotides		
TrueGuide™ tracrRNA	ThermoFisher Scientific	Cat#A35508
TrueGuide modified synthetic crRNA		
POU5F1	ThermoFisher Scientific	Cat#CRISPR777205_CR
SOX2	ThermoFisher Scientific	Cat#CRISPR1081382_CR
NANOG	ThermoFisher Scientific	Cat#CRISPR850052_CR
TFAP2C	ThermoFisher Scientific	Cat#CRISPR906394_CR
Recombinant DNA		
pGG195/GATA3mKO	This study	N/A
CML32	This study	N/A
px459_SpCas9-2A-Puro	Addgene	Cat#62988
PBase	Guo et al 2009	N/A
AAVS1-Cas9 targeting vector	Gift from Prof. Kosuke Yusa	N/A
pGG150-hNanog	This study	N/A

PB-CAG-Tet3G-IZ	This study	N/A
Software and Algorithms		
STAR	Dobin A et al., 2013	
htseq-count	Anders S et al., 2014	
DESeq	Anders S et al., 2010	
FactoMineR	Lê et al., 2008	
TF annotations	http://bioinfo.life.hust.edu.cn/AnimalTFDB/	
Samtools	Li et al., 2009	http://samtools.sourceforge.net/
MFuzz	Kumar L, Futschik M 2007	
R	R Core Team, 2017	https://www.R-project.org/
MASS	Venables, W. N. & Ripley, B. D. (2002)	
Genome and Genome annotation	Ensembl 96	http://apr2019.archive.ensembl.org/index.html
gplots	Gregory R. Warnes et al., 2019	
plot3D	Karline Soetaert (2017)	

Table S1 gRNA sequences for targeting in Cas9 expressing naïve cells

gRNA target	***Sequence***
POU5F1-1	GCCACCAAATAGAACCCCCA
POU5F1-2	GCCACCAAATAGAACCCCCA
SOX2-1	GTTATAAATACCGGCCCCGG
SOX2-2	GCAGGGCGCTCACGTCGTAG
NANOG-1	GGTTCACCAGGCATCCCTGG
NANOG-2	GGTCGCAAAAAAGGAAGACA
TFAP2C-1	ATATTCGGCGACTCCAGTGT
TFAP2C-2	GCTTAAATGCCTCGTTAC
YAP1	GACGTTCATCTGGGACAGCA
TAZ	TACGAGCTCATCGAGAAGCG
GFP	GAGCTGGACGGCGACGTAAA

REFERENCES

Adachi, K., Nikaido, I., Ohta, H., Ohtsuka, S., Ura, H., Kadota, M., Wakayama, T., Ueda, H.R., and Niwa, H. (2013). Context-dependent wiring of Sox2 regulatory networks for self-renewal of embryonic and trophoblast stem cells. Mol Cell *52*, 380-392.

Amita, M., Adachi, K., Alexenko, A.P., Sinha, S., Schust, D.J., Schulz, L.C., Roberts, R.M., and Ezashi, T. (2013). Complete and unidirectional conversion of human embryonic stem cells to trophoblast by BMP4. Proceedings of the National Academy of Sciences *110*, E1212-E1221.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol *11*, R106.

Anders, S., Pyl, P.T., and Huber, W. (2014). HTSeq — A Python framework to work with high-throughput sequencing data. bioRxiv.

Beddington, R.S.P., and Robertson, E.J. (1989). An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. Development *105*, 733-737.

Berg, D.K., Smith, C.S., Pearton, D.J., Wells, D.N., Broadhurst, R., Donnison, M., and Pfeffer, P.L. (2011). Trophectoderm lineage determination in cattle. Dev Cell *20*, 244-255.

Bernardo, A.S., Faial, T., Gardner, L., Niakan, K.K., Ortmann, D., Senner, C.E., Callery, E.M., Trotter, M.W., Hemberger, M., Smith, J.C.*, et al.* (2011). BRACHYURY and CDX2 mediate BMP-induced differentiation of human and mouse pluripotent stem cells into embryonic and extraembryonic lineages. Cell Stem Cell *9*, 144-155.

Boroviak, T., Loos, R., Bertone, P., Smith, A., and Nichols, J. (2014). The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. Nat Cell Biol *16*, 516-528.

Boroviak, T., Stirparo, G.G., Dietmann, S., Hernando-Herraez, I., Mohammed, H., Reik, W., Smith, A., Sasaki, E., Nichols, J., and Bertone, P. (2018). Single cell transcriptome analysis of human, marmoset and mouse embryos reveals common and divergent features of preimplantation development. Development *145*.

Bradley, A., Evans, M.J., Kaufman, M.H., and Robertson, E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature *309*, 255-256.

Bredenkamp, N., Yang, J., Clarke, J., Stirparo, G.G., von Meyenn, F., Baker, D., Drummond, R., li, D., Wu, C., Rostovskaya, M.*, et al.* (2019a). Efficient RNA-mediated reprogramming of human somatic cells to naïve pluripotency facilitated by tankyrase inhibition. bioRxiv, 636670.

Bredenkamp, N., Yang, J., Clarke, J., Stirparo, G.G., von Meyenn, F., Dietmann, S., Baker, D., Drummond, R., Ren, Y., Li, D.*, et al.* (2019b). Wnt Inhibition Facilitates RNA-Mediated Reprogramming of Human Somatic Cells to Naive Pluripotency. Stem Cell Reports.

Brook, F.A., and Gardner, R.L. (1997). The origin and efficient derivation of embryonic stem cells in the mouse. PNAS *94*, 5709-5712.

Cambuli, F., Murray, A., Dean, W., Dudzinska, D., Krueger, F., Andrews, S., Senner, C.E., Cook, S.J., and Hemberger, M. (2014). Epigenetic memory of the first cell fate decision prevents complete ES cell reprogramming into trophoblast. Nat Commun *5*, 5538.

Cao, Z., Carey, T.S., Ganguly, A., Wilson, C.A., Paul, S., and Knott, J.G. (2015). Transcription factor AP-2γ induces early Cdx2 expression and represses HIPPO signaling to specify the trophectoderm lineage. Development *142*, 1606-1615.

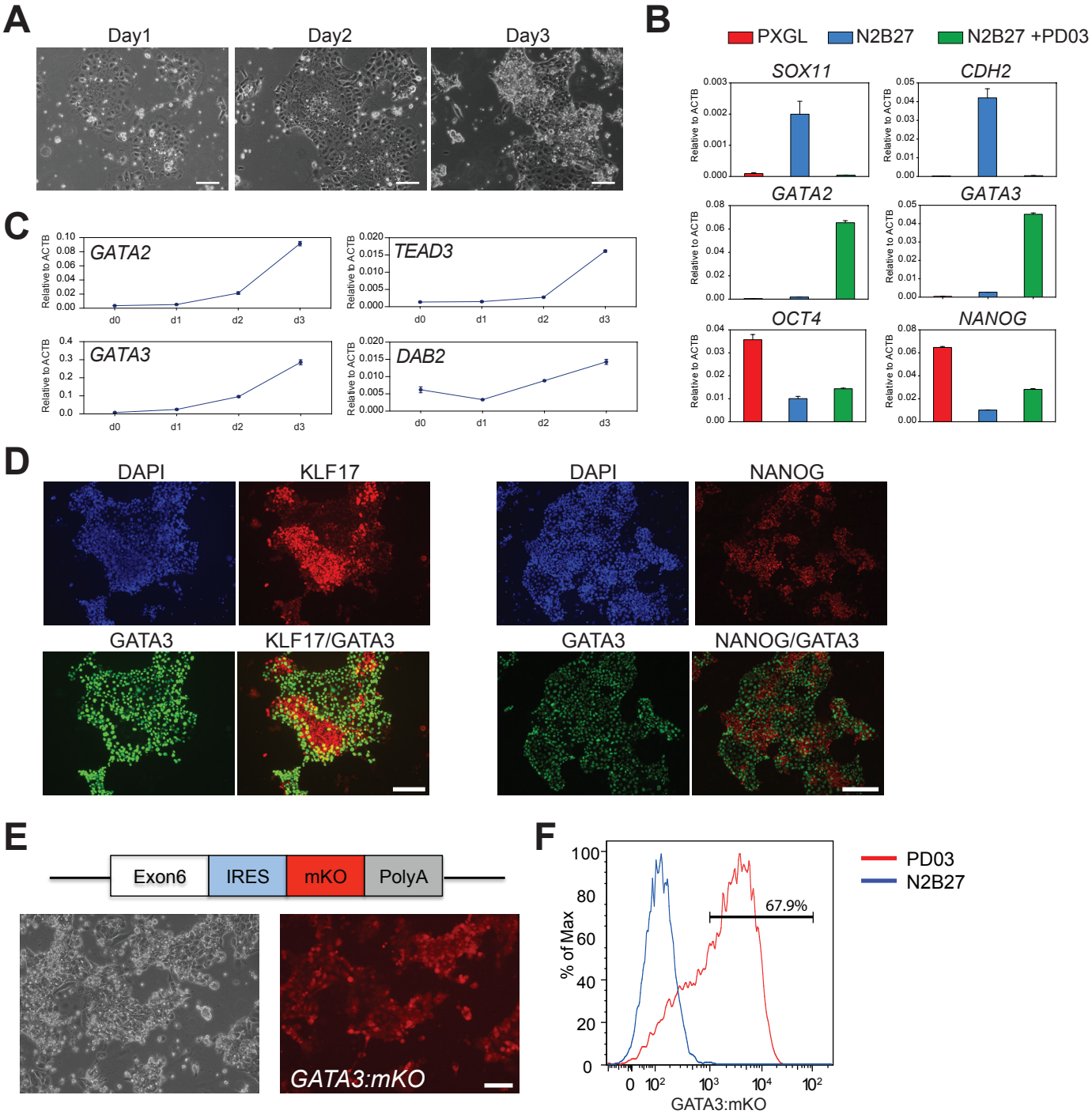
Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. Nature *450*, 1230-1234.

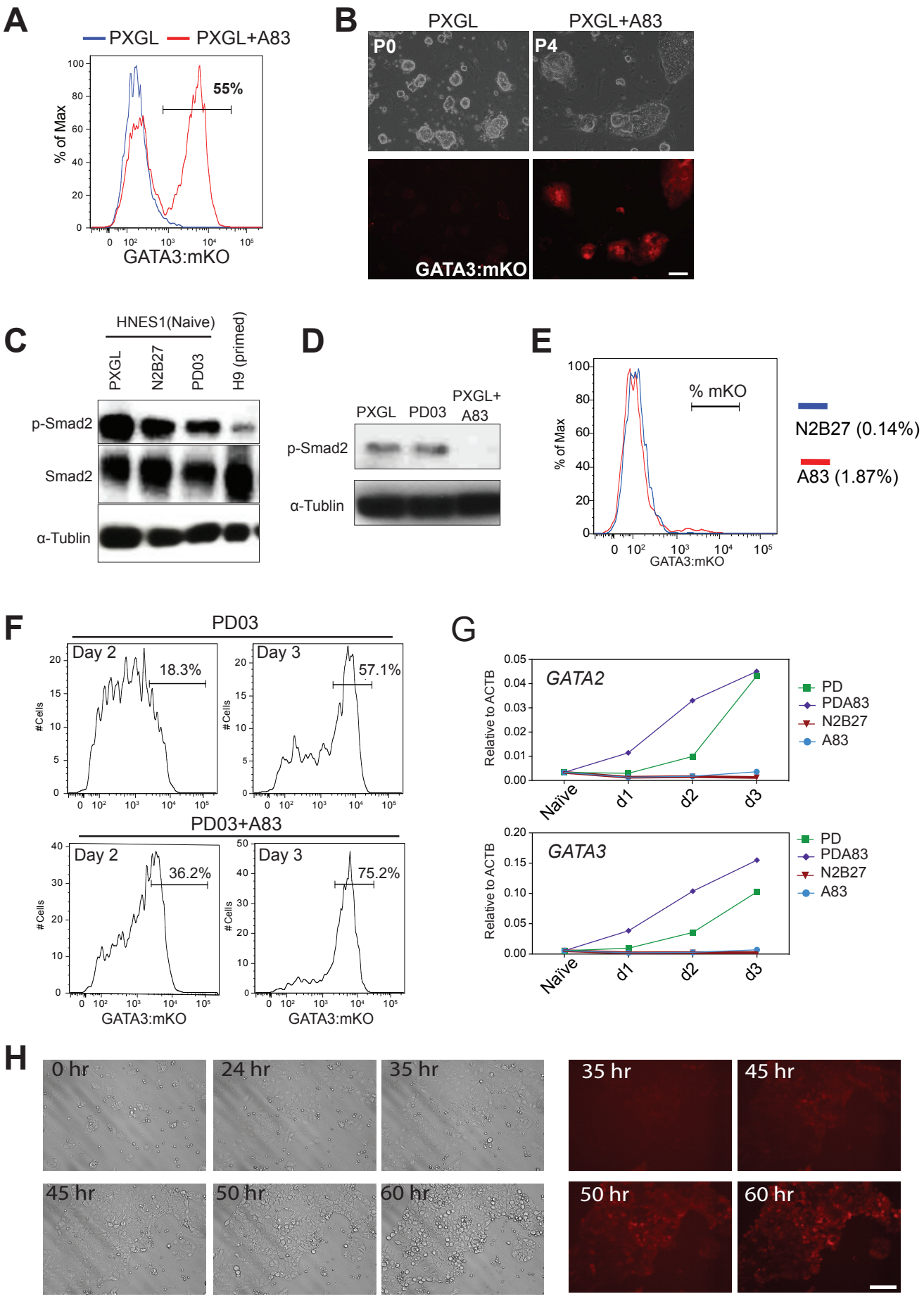
Chazaud, C., Yamanaka, Y., Pawson, T., and Rossant, J. (2006). Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. Dev Cell *10*, 615-624.

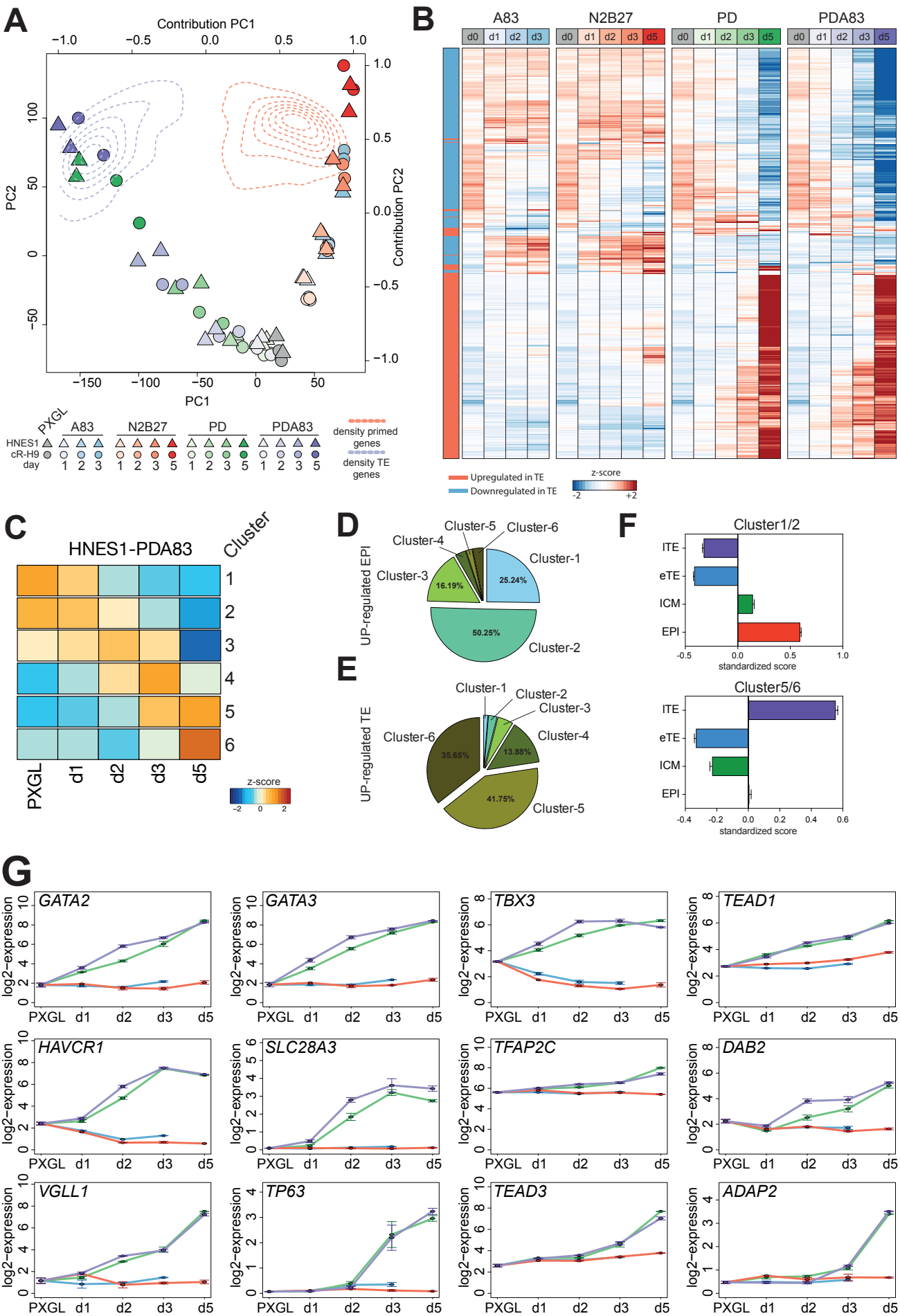
Chen, G., Gulbranson, D.R., Hou, Z., Bolin, J.M., Ruotti, V., Probasco, M.D., Smuga-Otto, K., Howden, S.E., Diol, N.R., Propson, N.E.*, et al.* (2011). Chemically defined conditions for human iPSC derivation and culture. Nat Methods *8*, 424-429.
Choi, I., Carey, T.S., Wilson, C.A., and Knott, J.G. (2012). Transcription factor AP-2gamma is a core regulator of tight junction biogenesis and cavity formation during mouse early embryogenesis. Development *139*, 4623-4632.
De Paepe, C., Cauffman, G., Verloes, A., Sterckx, J., Devroey, P., Tournaye, H., Liebaers, I., and Van de Velde, H. (2013). Human trophectoderm cells are not yet committed. Hum Reprod *28*, 740-749.
Deglincerti, A., Croft, G.F., Pietila, L.N., Zernicka-Goetz, M., Siggia, E.D., and Brivanlou, A.H. (2016). Self-organization of the in vitro attached human embryo. Nature *533*, 251-254.
Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.
Dong, C., Fischer, L., and Theunissen, T.W. (2019). Recent insights into the naive state of human pluripotency and its applications. Exp Cell Res, 111645.
Dunn, S.J., Martello, G., Yordanov, B., Emmott, S., and Smith, A.G. (2014). Defining an essential transcription factor program for naïve pluripotency. Science *344*, 1156-1160.
Evans, M.J., and Kaufman, M. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154-156.
Festuccia, N., Osorno, R., Halbritter, F., Karwacki-Neisius, V., Navarro, P., Colby, D., Wong, F., Yates, A., Tomlinson, Simon R., and Chambers, I. (2012). Esrrb Is a Direct Nanog Target Gene that Can Substitute for Nanog Function in Pluripotent Cells. Cell Stem Cell *11*, 477-490.
Gao, X., Nowak-Imialek, M., Chen, X., Chen, D., Herrmann, D., Ruan, D., Chen, A.C.H., Eckersley-Maslin, M.A., Ahmad, S., Lee, Y.L.*, et al.* (2019). Establishment of porcine and human expanded potential stem cells. Nat Cell Biol *21*, 687-699.
Gardner, R.L. (1983). Origin and differentiation of extra-embryonic tissues in the mouse. IntRevExpPath *24*, 63-133.
Gardner, R.L., and Rossant, J. (1979). Investigation of the fate of 4-5 day post-coitum mouse inner cell mass cells by blastocyst injection. J Embryol Exp Morphol *52*, 141-152.
Guo, G., von Meyenn, F., Rostovskaya, M., Clarke, J., Dietmann, S., Baker, D., Sahakyan, A., Myers, S., Bertone, P., Reik, W.*, et al.* (2017). Epigenetic resetting of human pluripotency. Development *144*, 2748-2763.
Guo, G., von Meyenn, F., Santos, F., Chen, Y., Reik, W., Bertone, P., Smith, A., and Nichols, J. (2016). Naive Pluripotent Stem Cells Derived Directly from Isolated Cells of the Human Inner Cell Mass. Stem Cell Reports *6*, 437-446.
Kuckenberg, P., Buhl, S., Woynecki, T., van Fürden, B., Tolkunova, E., Seiffe, F., Moser, M., Tomilin, A., Winterhager, E., and Schorle, H. (2010). The Transcription Factor TCFAP2C/AP-2γ Cooperates with CDX2 To Maintain Trophectoderm Formation. Mol Cell Biol *30*, 3310-3320.
Kumar, L., and M, E.F. (2007). Mfuzz: a software package for soft clustering of microarray data. Bioinformation *2*, 5-7.
Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: An R package for multivariate analysis. Journal of Statistical Software *25*, 1-18.
Lee, C.Q., Gardner, L., Turco, M., Zhao, N., Murray, M.J., Coleman, N., Rossant, J., Hemberger, M., and Moffett, A. (2016). What Is Trophoblast? A Combination of Criteria Define Human First-Trimester Trophoblast. Stem Cell Reports *6*, 257-272.
Ma, H., Zhai, J., Wan, H., Jiang, X., Wang, X., Wang, L., Xiang, Y., He, X., Zhao, Z.-A., Zhao, B.*, et al.* (2019). In vitro culture of cynomolgus monkey embryos beyond early gastrulation. Science *366*, eaax7890.
Martin, G.R. (1981). 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.

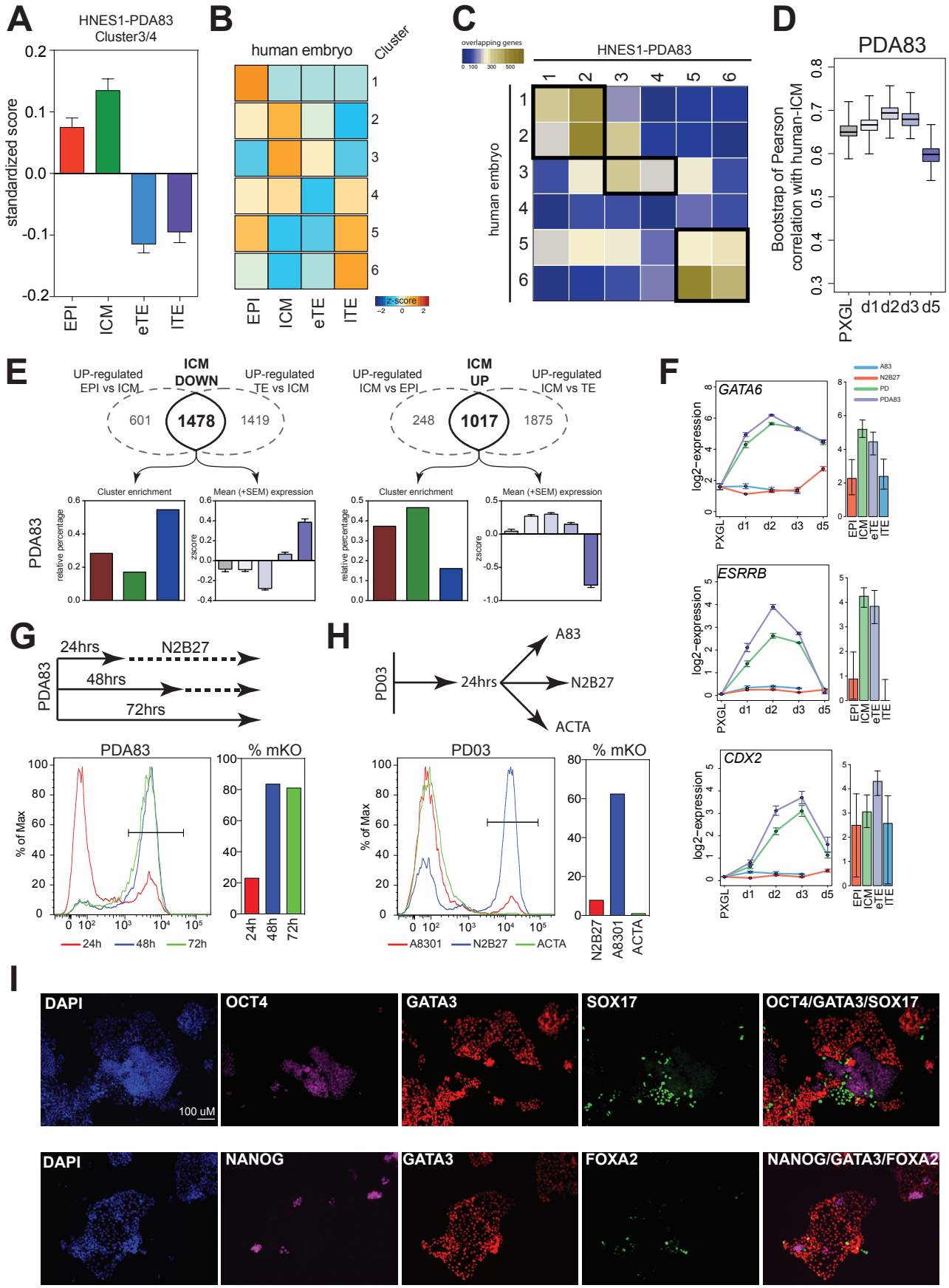
Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Marauyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell *113*, 631-642.
Nagy, A., Rossant, J., Nagy, R., Abramow-Newerly, W., and Roder, J.C. (1993). Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. Proc Natl Acad Sci U S A *90*, 8424-8428.
Nakamura, T., Okamoto, I., Sasaki, K., Yabuta, Y., Iwatani, C., Tsuchiya, H., Seita, Y., Nakamura, S., Yamamoto, T., and Saitou, M. (2016). A developmental coordinate of pluripotency among mice, monkeys and humans. Nature *537*, 57-62.
Ng, R.K., Dean, W., Dawson, C., Lucifero, D., Madeja, Z., Reik, W., and Hemberger, M. (2008). Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5. Nat Cell Biol.
Niakan, K.K., and Eggan, K. (2013). Analysis of human embryos from zygote to blastocyst reveals distinct gene expression patterns relative to the mouse. Dev Biol *375*, 54-64.
Nichols, J., and Gardner, R.L. (1984). Heterogeneous differentiation of external cells in individual isolated early mouse inner cell masses in culture. J Embryol Exp Morphol *80*, 225-240.
Nichols, J., Jones, K., Phillips, J.M., Newland, S.A., Roode, M., Mansfield, W., Smith, A., and Cooke, A. (2009). Validated germline-competent embryonic stem cell lines from nonobese diabetic mice. Nat Med *15*, 814-818.
Nishioka, N., Inoue, K., Adachi, K., Kiyonari, H., Ota, M., Ralston, A., Yabuta, N., Hirahara, S., Stephenson, R.O., Ogonuki, N.*, et al.* (2009). The Hippo signaling pathway components Lats and Yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass. Dev Cell *16*, 398-410.
Niwa, H., Toyooka, Y., Daisuke, S., Strumpf, D., Takahashi, K., Yagi, R., and Rossant, J. (2005). Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. Cell *123*, 917-929.
O'Leary, T., Heindryckx, B., Lierman, S., van Bruggen, D., Goeman, J.J., Vandewoestyne, M., Deforce, D., de Sousa Lopes, S.M., and De Sutter, P. (2012). Tracking the progression of the human inner cell mass during embryonic stem cell derivation. Nat Biotechnol *30*, 278-282.
Pastor, W.A., Liu, W., Chen, D., Ho, J., Kim, R., Hunt, T.J., Lukianchikov, A., Liu, X., Polo, J.M., Jacobsen, S.E.*, et al.* (2018). TFAP2C regulates transcription in human naive pluripotency by opening enhancers. Nat Cell Biol *20*, 553-564.
Petropoulos, S., Edsgard, D., Reinius, B., Deng, Q., Panula, S.P., Codeluppi, S., Plaza Reyes, A., Linnarsson, S., Sandberg, R., and Lanner, F. (2016). Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. Cell *165*, 1012-1026.
Plusa, B., Piliszek, A., Frankenberg, S., Artus, J., and Hadjantonakis, A.K. (2008). Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. Development *135*, 3081-3091.
Posfai, E., Petropoulos, S., de Barros, F.R.O., Schell, J.P., Jurisica, I., Sandberg, R., Lanner, F., and Rossant, J. (2017). Position- and Hippo signaling-dependent plasticity during lineage segregation in the early mouse embryo. Elife *6*, e22906.
Rivron, N.C., Frias-Aldeguer, J., Vrij, E.J., Boisset, J.C., Korving, J., Vivie, J., Truckenmuller, R.K., van Oudenaarden, A., van Blitterswijk, C.A., and Geijsen, N. (2018). Blastocyst-like structures generated solely from stem cells. Nature *557*, 106-111.
Rossant, J. (2015). Mouse and human blastocyst-derived stem cells: vive les differences. Development *142*, 9-12.
Rossant, J., and Tam, P.P.L. (2017). New Insights into Early Human Development: Lessons for Stem Cell Derivation and Differentiation. Cell Stem Cell *20*, 18-28.
Rostovskaya, M., Stirparo, G.G., and Smith, A. (2019). Capacitation of human naïve pluripotent stem cells for multi-lineage differentiation. Development *146*, dev172916.

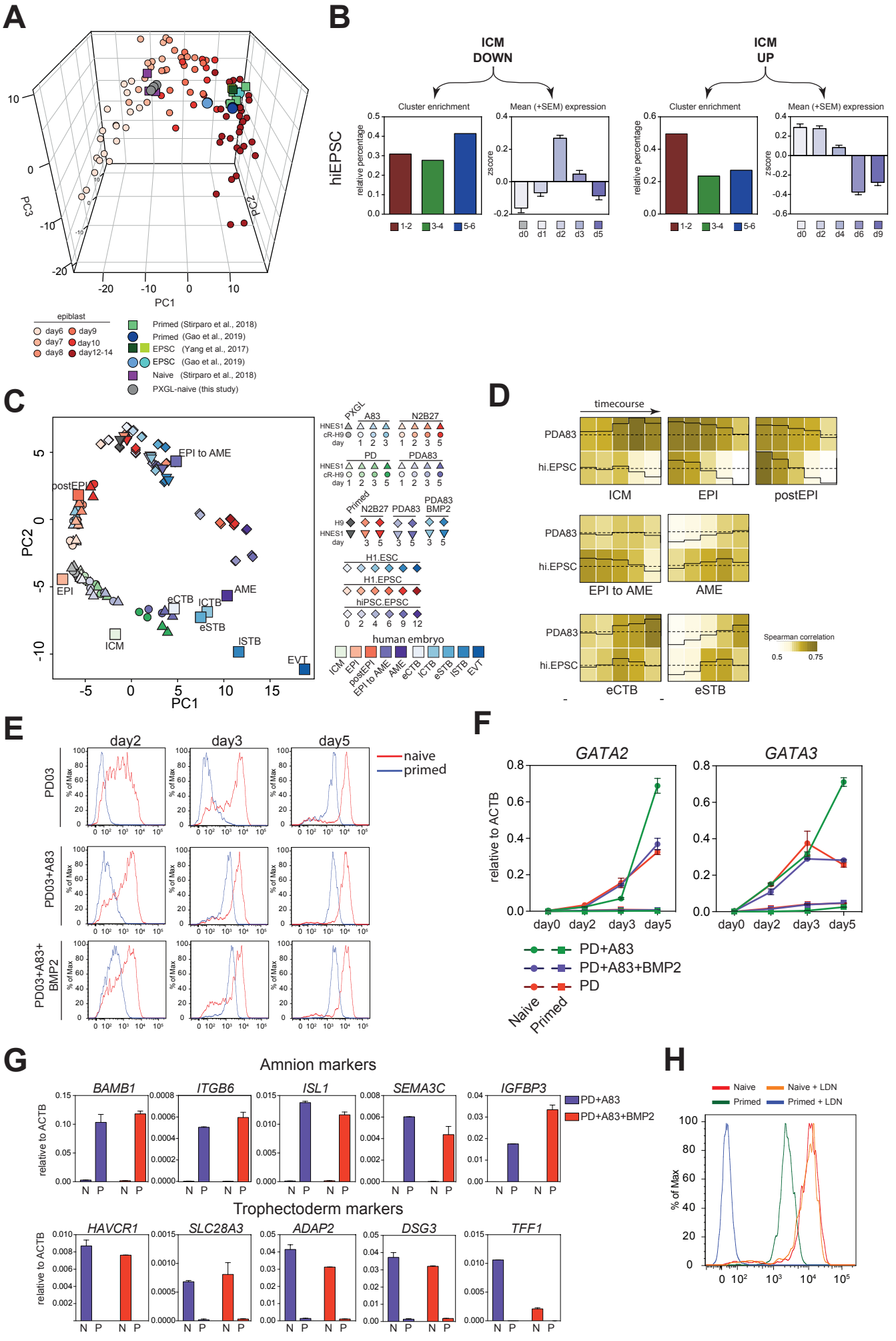
Saiz, N., Williams, K.M., Seshan, V.E., and Hadjantonakis, A.K. (2016). Asynchronous fate decisions by single cells collectively ensure consistent lineage composition in the mouse blastocyst. Nat Commun *7*, 13463.
Shahbazi, M.N., Jedrusik, A., Vuoristo, S., Recher, G., Hupalowska, A., Bolton, V., Fogarty, N.M., Campbell, A., Devito, L.G., Ilic, D.*, et al.* (2016). Self-organization of the human embryo in the absence of maternal tissues. Nat Cell Biol *18*, 700-708.
Silva, J., Nichols, J., Theunissen, T.W., Guo, G., van Oosten, A.L., Barrandon, O., Wray, J., Yamanaka, S., Chambers, I., and Smith, A. (2009). Nanog is the gateway to the pluripotent ground state. Cell *138*, 722-737.
Smith, A. (2017). Formative pluripotency: the executive phase in a developmental continuum. Development *144*, 365-373.
Sozen, B., Cox, A.L., De Jonghe, J., Bao, M., Hollfelder, F., Glover, D.M., and Zernicka-Goetz, M. (2019). Self-Organization of Mouse Stem Cells into an Extended Potential Blastoid. Dev Cell *51*, 698-712.e698.
Stirparo, G.G., Boroviak, T., Guo, G., Nichols, J., Smith, A., and Bertone, P. (2018). Integrated analysis of single-cell embryo data yields a unified transcriptome signature for the human pre-implantation epiblast. Development *145*.
Strumpf, D., Mao, C.A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F., and Rossant, J. (2005). Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. Development *132*, 2093-2102.
Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861-872.
Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficz, G., Krueger, F., Oxley, D., Santos, F., Clarke, J., Mansfield, W.*, et al.* (2014). Resetting Transcription Factor Control Circuitry toward Ground-State Pluripotency in Human. Cell *158*, 1254-1269.
Taniguchi, Y., Ido, H., Sanzen, N., Hayashi, M., Sato-Nishiuchi, R., Futaki, S., and Sekiguchi, K. (2009). The C-terminal region of laminin beta chains modulates the integrin binding affinities of laminins. J Biol Chem *284*, 7820-7831.
Theunissen, T.W., Powell, B.E., Wang, H., Mitalipova, M., Faddah, D.A., Reddy, J., Fan, Z.P., Maetzel, D., Ganz, K., Shi, L.*, et al.* (2014). Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. Cell Stem Cell *15*, 471-487.
Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145-1147.
Xiang, L., Yin, Y., Zheng, Y., Ma, Y., Li, Y., Zhao, Z., Guo, J., Ai, Z., Niu, Y., Duan, K.*, et al.* (2019). A developmental landscape of 3D-cultured human pre-gastrulation embryos. Nature.
Xu, R.H., Chen, X., Li, D.S., Li, R., Addicks, G.C., Glennon, C., Zwaka, T.P., and Thomson, J.A. (2002). BMP4 initiates human embryonic stem cell differentiation to trophoblast. Nat Biotechnol *20*, 1261-1264.
Yang, Y., Liu, B., Xu, J., Wang, J., Wu, J., Shi, C., Xu, Y., Dong, J., Wang, C., Lai, W.*, et al.* (2017). Derivation of Pluripotent Stem Cells with In Vivo Embryonic and Extraembryonic Potency. Cell *169*, 243-257 e225.
Ying, Q.L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519-523.
Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R.*, et al.* (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science *318*, 1917-1920.
Zheng, Y., Xue, X., Shao, Y., Wang, S., Esfahani, S.N., Li, Z., Muncie, J.M., Lakins, J.N., Weaver, V.M., Gumucio, D.L.*, et al.* (2019). Controlled modelling of human epiblast and amnion development using stem cells. Nature *573*, 421-425.

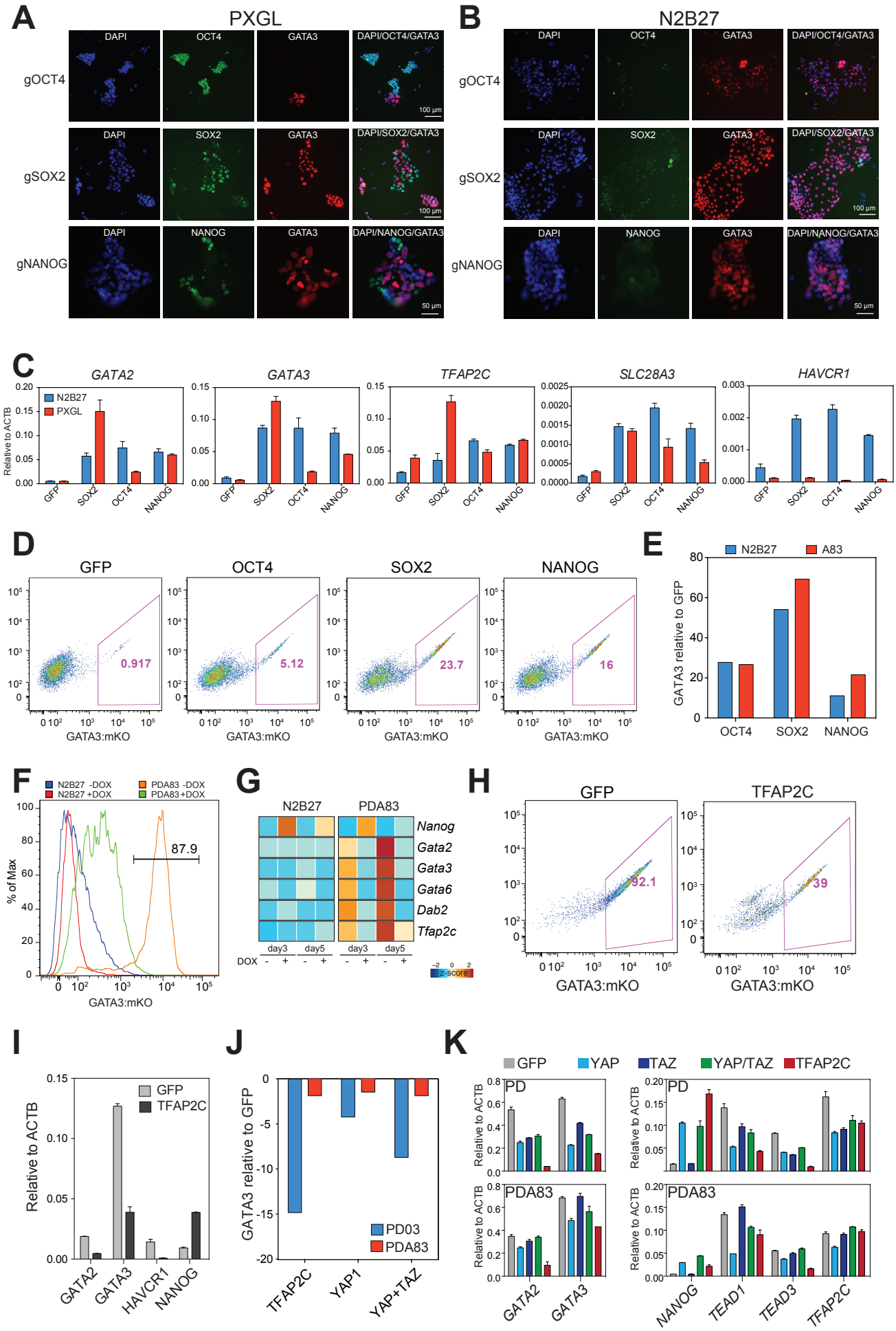

A
Day1
Day2
Day3
B
PXGL
N2B27
N2B27 +PD03
SOX11
CDH2
GATA2
GATA3
OCT4
NANOG
Relative to ACTB
C
GATA2
TEAD3
GATA3
DAB2
d0
d1
d2
d3
D
DAPI
KLF17
NANOG
GATA3
KLF17/GATA3
NANOG/GATA3
E
Exon6
IRES
mKO
PolyA
GATA3:mKO
F
PD03
N2B27
67.9%
% of Max
GATA3:mKO

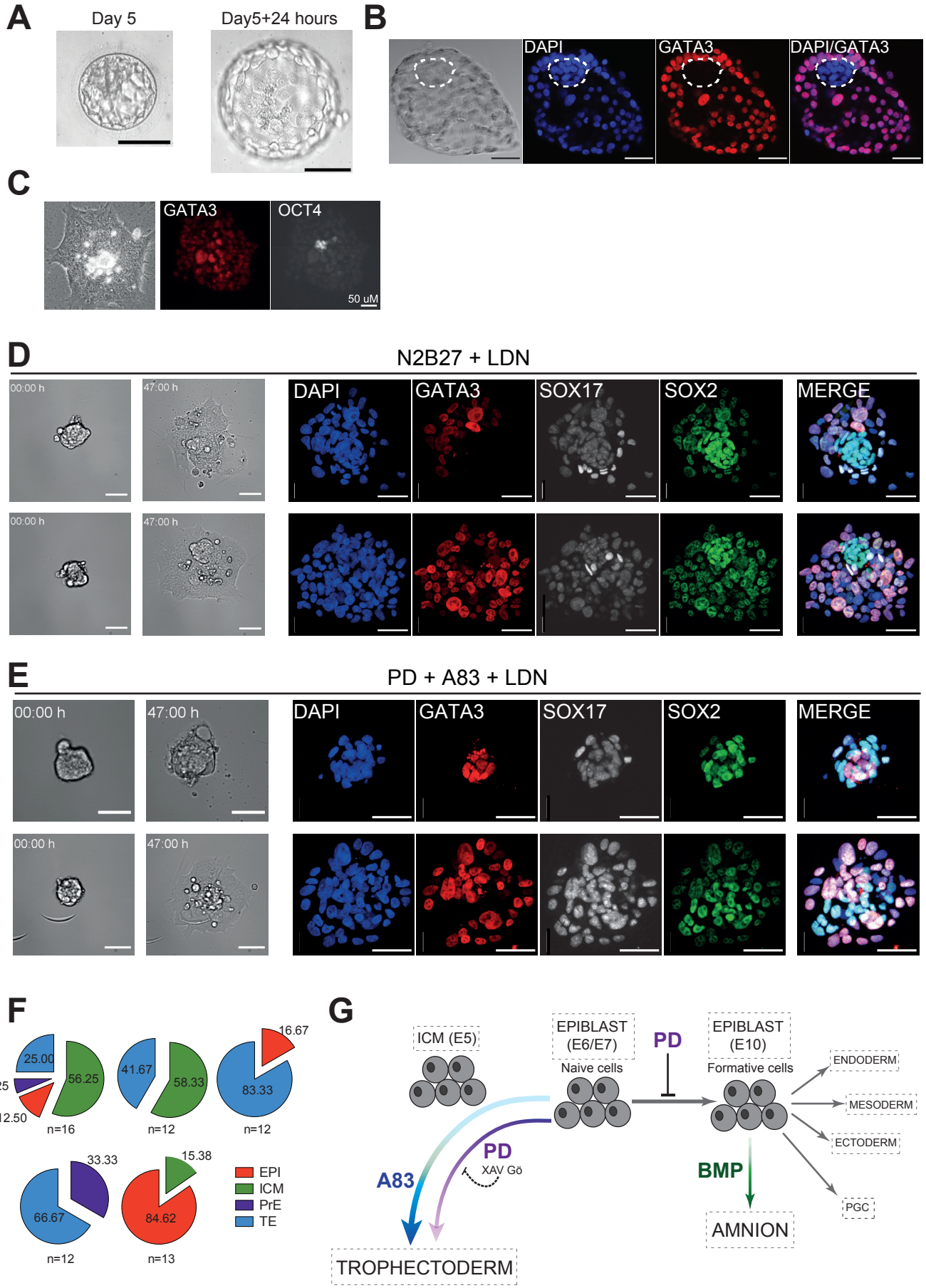

A
PXGL
PXGL+A83
55%
% of Max
GATA3:mKO
B
PXGL
PXGL+A83
P0
P4
GATA3:mKO
C
HNES1(Naive)
PXGL
N2B27
PD03
H9 (primed)
p-Smad2
Smad2
α-Tublin
D
PXGL
PD03
PXGL+
A83
p-Smad2
α-Tublin
E
% mKO
% of Max
GATA3:mKO
N2B27 (0.14%)
A83 (1.87%)
F
PD03
Day 2
18.3%
Day 3
57.1%
PD03+A83
Day 2
36.2%
Day 3
75.2%
#Cells
GATA3:mKO
G
GATA2
GATA3
Relative to ACTB
PD
PDA83
N2B27
A83
Naïve
d1
d2
d3
H
0 hr
24 hr
35 hr
45 hr
50 hr
60 hr

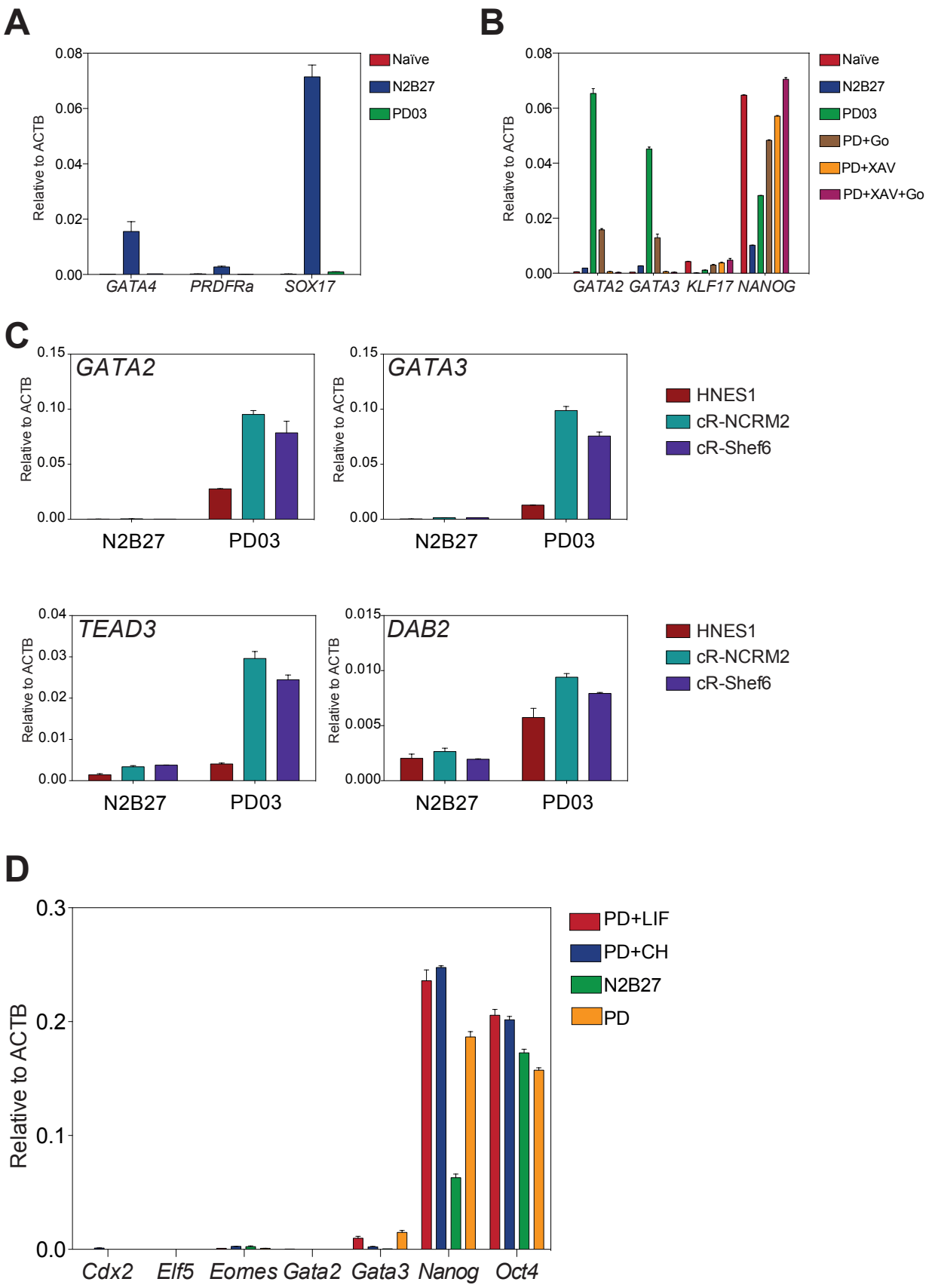

A
Contribution PC1
PC1
PC2
Contribution PC2
PXGL
A83
N2B27
PD
PDA83
HNES1
cR-H9
day
density primed genes
density TE genes
B
A83
N2B27
PD
PDA83
d0
d1
d2
d3
d5
Upregulated in TE
Downregulated in TE
z-score
C
HNES1-PDA83
Cluster
PXGL
d1
d2
d3
d5
z-score
D
UP-regulated EPI
Cluster-1
Cluster-2
Cluster-3
Cluster-4
Cluster-5
Cluster-6
25.24%
50.25%
16.19%
E
UP-regulated TE
Cluster-1
Cluster-2
Cluster-3
Cluster-4
Cluster-5
Cluster-6
35.65%
41.75%
13.88%
F
Cluster1/2
Cluster5/6
ITE
eTE
ICM
EPI
standardized score
G
GATA2
GATA3
TBX3
TEAD1
HAVCR1
SLC28A3
TFAP2C
DAB2
VGLL1
TP63
TEAD3
ADAP2
log2−expression
PXGL
d1
d2
d3
d5

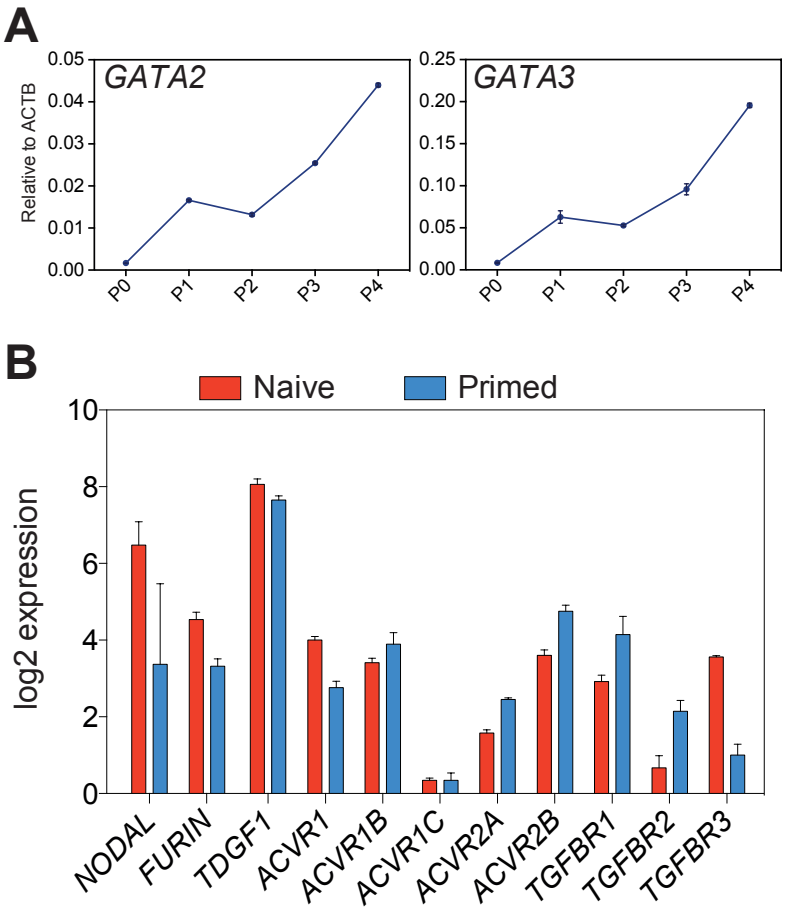

A
HNES1-PDA83
Cluster3/4
standardized score
EPI
ICM
eTE
ITE
B
human embryo
Cluster
z-score
C
overlapping genes
HNES1-PDA83
human embryo
D
PDA83
Bootstrap of Pearson correlation with human-ICM
PXGL
d1 d2 d3 d5
E
UP-regulated EPI vs ICM
ICM DOWN
UP-regulated TE vs ICM
601
1478
1419
UP-regulated ICM vs EPI
ICM UP
UP-regulated ICM vs TE
248
1017
1875
PDA83
Cluster enrichment
Mean (+SEM) expression
relative percentage
zscore
F
GATA6
ESRRB
CDX2
log2-expression
A83
N2B27
PD
PDA83
G
24hrs
N2B27
48hrs
72hrs
PDA83
% of Max
% mKO
24h
48h
72h
H
PD03
24hrs
A83
N2B27
ACTA
PD03
% mKO
A8301
N2B27
ACTA
I
DAPI
OCT4
GATA3
SOX17
OCT4/GATA3/SOX17
100 uM
DAPI
NANOG
GATA3
FOXA2
NANOG/GATA3/FOXA2

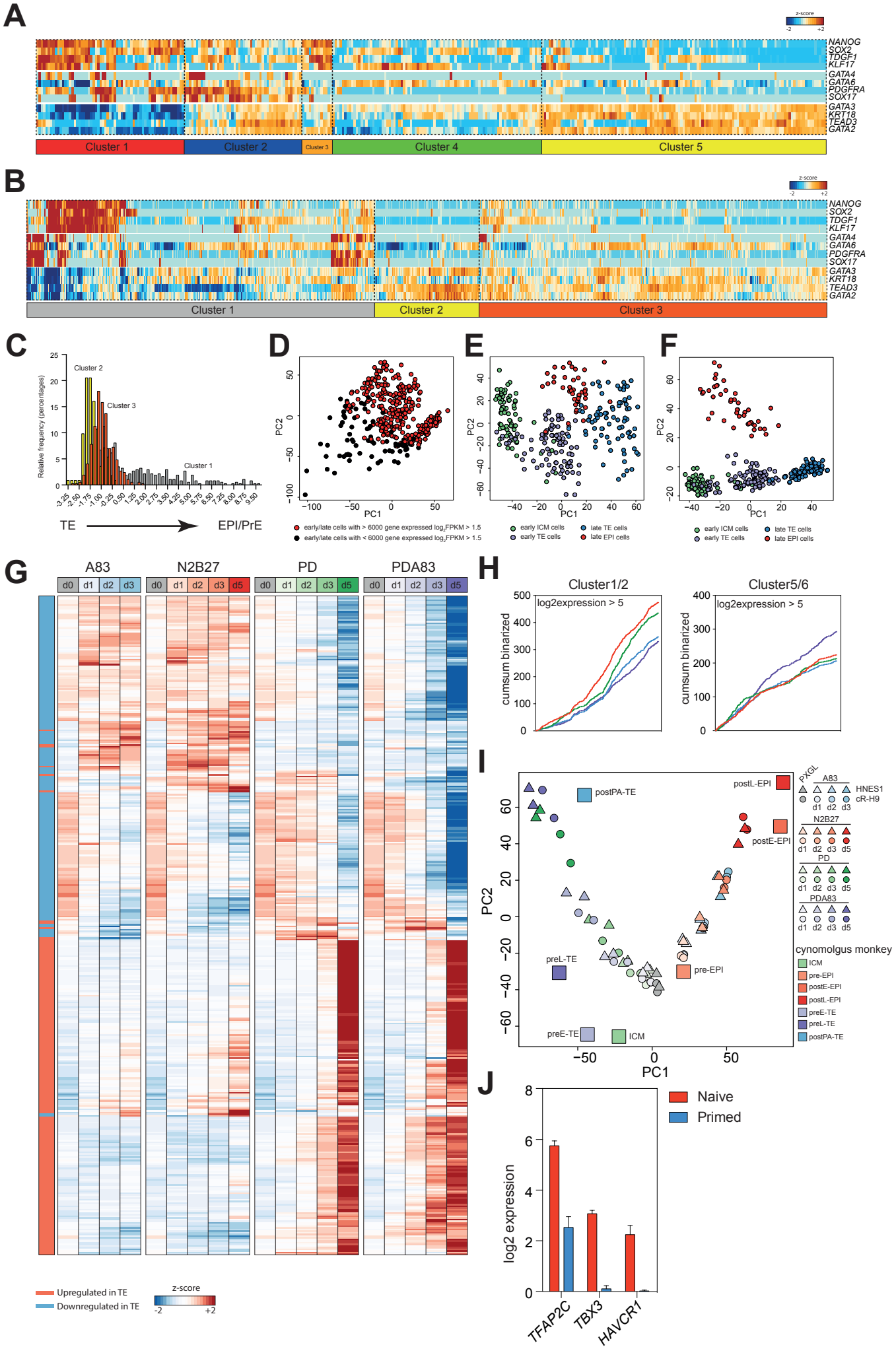

A
PC1
PC2
PC3
epiblast
day6
day7
day8
day9
day10
day12-14
Primed (Stirparo et al., 2018)
Primed (Gao et al., 2019)
EPSC (Yang et al., 2017)
EPSC (Gao et al., 2019)
Naive (Stirparo et al., 2018)
PXGL-naive (this study)
B
ICM DOWN
ICM UP
Cluster enrichment
Mean (+SEM) expression
hiEPSC
relative percentage
zscore
1-2
3-4
5-6
d0
d1
d2
d3
d5
d4
d6
d9
C
EPI to AME
postEPI
EPI
AME
eCTB
ICTB
eSTB
ICM
ISTB
EVT
PC1
PC2
PXGL
A83
N2B27
PD
PDA83
HNES1
cR-H9
day
Primed
BMP2
H9
H1.ESC
H1.EPSC
hiPSC.EPSC
human embryo
D
timecourse
PDA83
hi.EPSC
ICM
EPI
postEPI
EPI to AME
AME
eCTB
eSTB
Spearman correlation
0.5
0.75
E
day2
day3
day5
naive
primed
PD03
PD03+A83
PD03+A83+BMP2
% of Max
F
GATA2
GATA3
relative to ACTB
day0
day2
day3
day5
PD+A83
PD+A83+BMP2
PD
Naive
Primed
G
Amnion markers
BAMB1
ITGB6
ISL1
SEMA3C
IGFBP3
PD+A83
PD+A83+BMP2
relative to ACTB
N P
Trophectoderm markers
HAVCR1
SLC28A3
ADAP2
DSG3
TFF1
H
Naive
Naive + LDN
Primed
Primed + LDN
% of Max

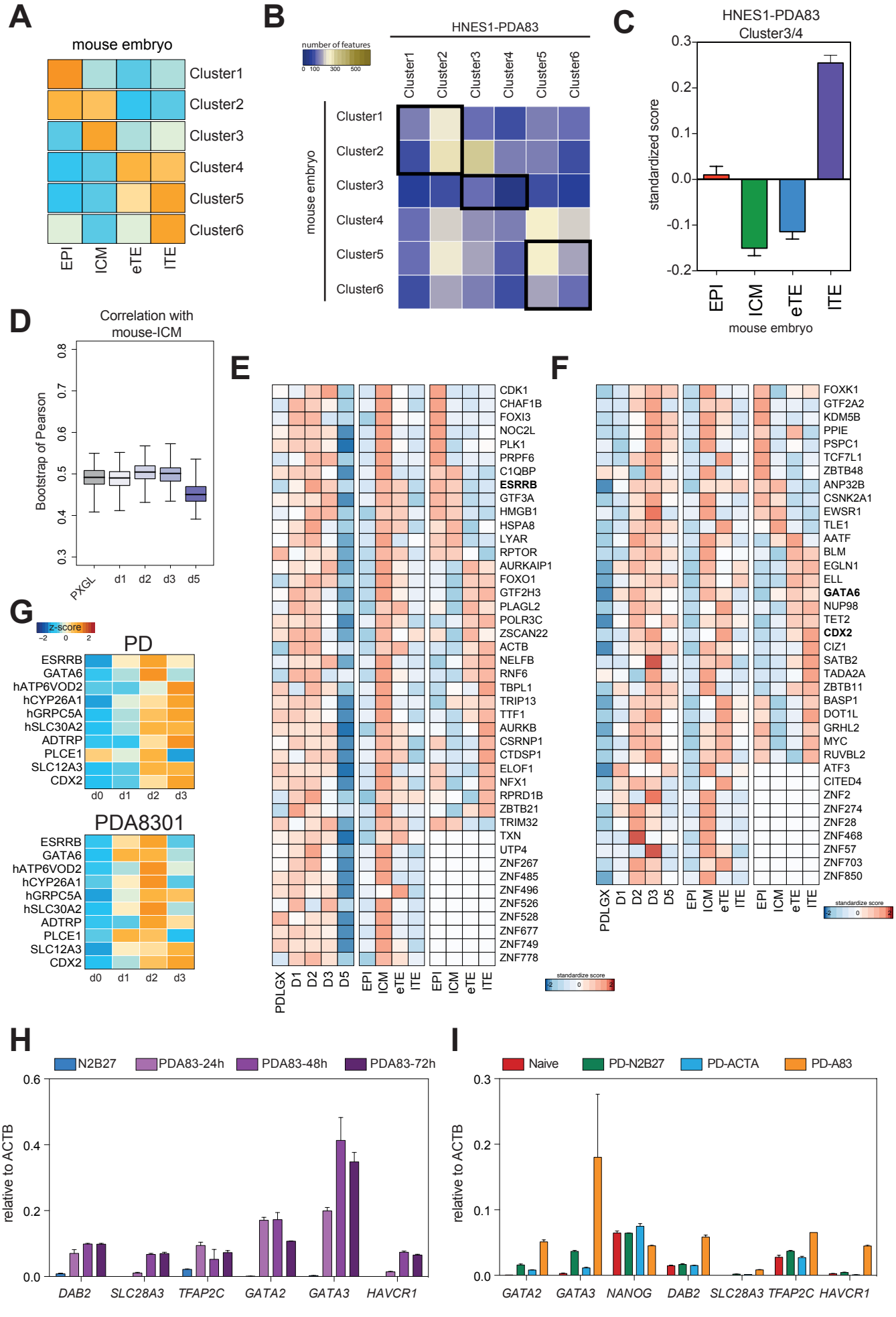

A
PXGL
gOCT4
gSOX2
gNANOG
DAPI
OCT4
GATA3
DAPI/OCT4/GATA3
SOX2
DAPI/SOX2/GATA3
NANOG
DAPI/NANOG/GATA3
100 µm
50 µm
B
N2B27
C
GATA2
GATA3
TFAP2C
SLC28A3
HAVCR1
N2B27
PXGL
Relative to ACTB
GFP
SOX2
OCT4
NANOG
D
GFP
OCT4
SOX2
NANOG
0.917
5.12
23.7
16
GATA3:mKO
E
N2B27
A83
GATA3 relative to GFP
OCT4
SOX2
NANOG
F
N2B27 -DOX
N2B27 +DOX
PDA83 -DOX
PDA83 +DOX
87.9
% of Max
GATA3:mKO
G
N2B27
PDA83
Nanog
Gata2
Gata3
Gata6
Dab2
Tfap2c
day3
day5
DOX
z-score
H
GFP
TFAP2C
92.1
39
GATA3:mKO
I
GFP
TFAP2C
Relative to ACTB
GATA2
GATA3
HAVCR1
NANOG
J
GATA3 relative to GFP
PD03
PDA83
TFAP2C
YAP1
YAP+TAZ
K
GFP
YAP
TAZ
YAP/TAZ
TFAP2C
PD
PDA83
Relative to ACTB
GATA2
GATA3
NANOG
TEAD1
TEAD3
TFAP2C

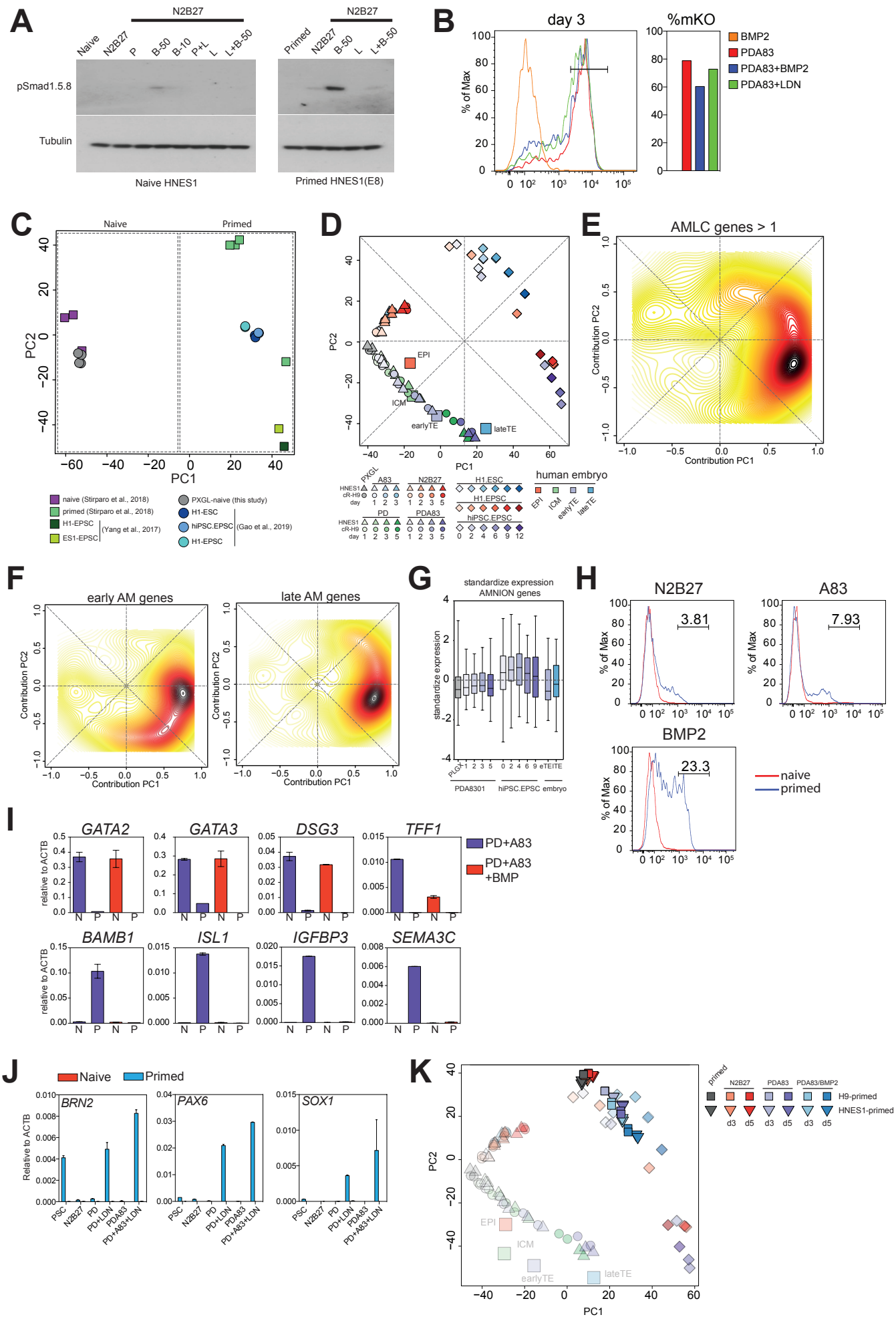

A
Day 5
Day5+24 hours
B
DAPI
GATA3
DAPI/GATA3
C
GATA3
OCT4
50 uM
D
N2B27 + LDN
00:00 h
47:00 h
DAPI
GATA3
SOX17
SOX2
MERGE
E
PD + A83 + LDN
00:00 h
47:00 h
DAPI
GATA3
SOX17
SOX2
MERGE
F
25.00
56.25
6.25
12.50
n=16
41.67
58.33
n=12
16.67
83.33
n=12
33.33
66.67
n=12
15.38
84.62
n=13
EPI
ICM
PrE
TE
G
ICM (E5)
EPIBLAST (E6/E7)
Naive cells
PD
EPIBLAST (E10)
Formative cells
ENDODERM
MESODERM
ECTODERM
PGC
A83
PD
XAV Gö
BMP
AMNION
TROPHECTODERM

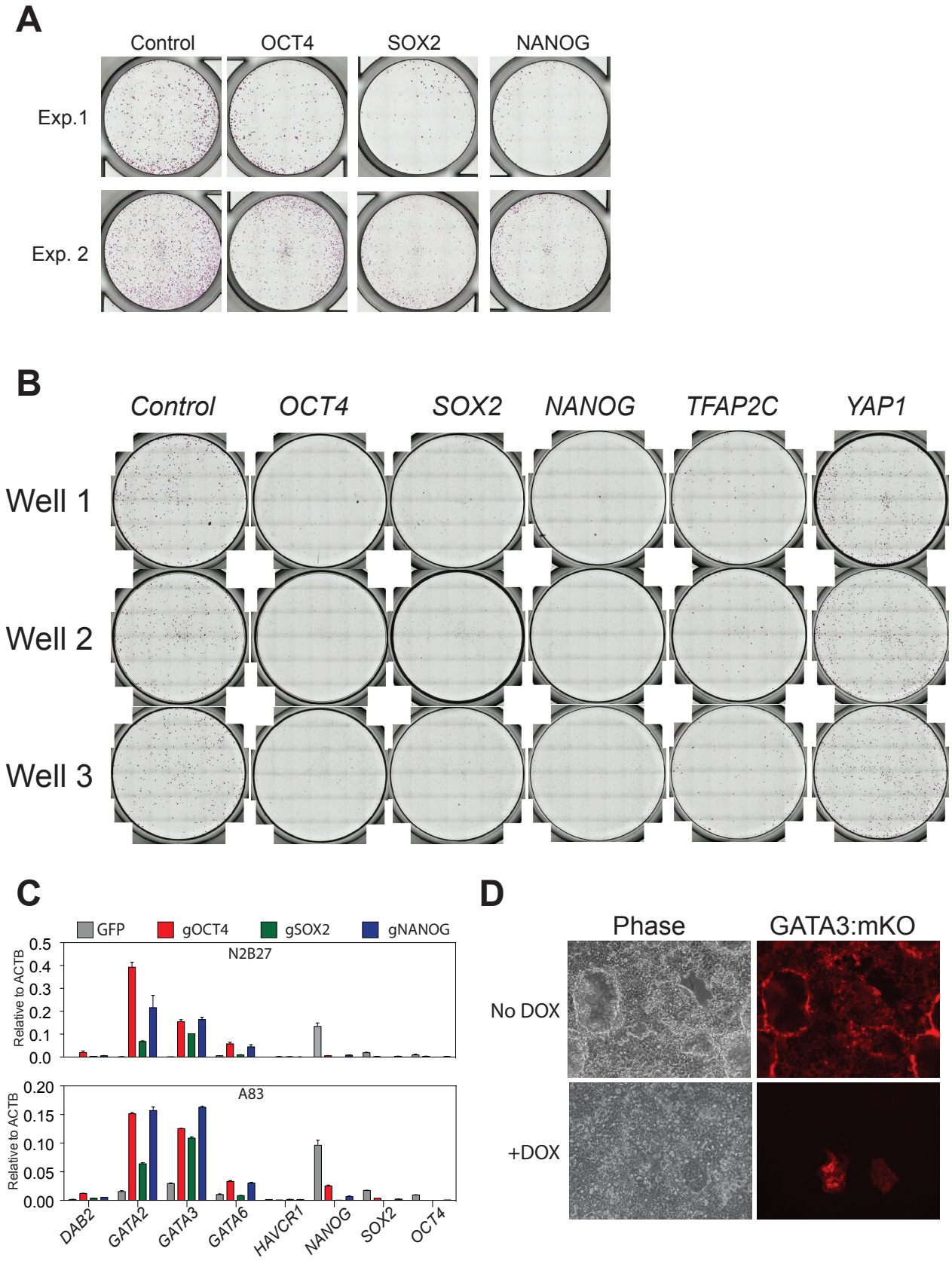

A
Relative to ACTB
0.08
0.06
0.04
0.02
0.00
GATA4
PRDFRa
SOX17
Naïve
N2B27
PD03
B
Relative to ACTB
GATA2
GATA3
KLF17
NANOG
Naïve
N2B27
PD03
PD+Go
PD+XAV
PD+XAV+Go
C
GATA2
GATA3
TEAD3
DAB2
Relative to ACTB
0.15
0.10
0.05
0.00
0.04
0.03
0.02
0.01
0.015
0.010
0.005
0.000
N2B27
PD03
HNES1
cR-NCRM2
cR-Shef6
D
Relative to ACTB
0.3
0.2
0.1
0.0
Cdx2
Elf5
Eomes
Gata2
Gata3
Nanog
Oct4
PD+LIF
PD+CH
N2B27
PD


A
GATA2
GATA3
Relative to ACTB
0.05
0.04
0.03
0.02
0.01
0.00
0.25
0.20
0.15
0.10
0.05
0.00
P0
P1
P2
P3
P4
B
Naive
Primed
log2 expression
10
8
6
4
2
0
NODAL
FURIN
TDGF1
ACVR1
ACVR1B
ACVR1C
ACVR2A
ACVR2B
TGFBR1
TGFBR2
TGFBR3


A
z-score
-2
+2
NANOG
SOX2
TDGF1
KLF17
GATA4
GATA6
PDGFRA
SOX17
GATA3
KRT18
TEAD3
GATA2
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
B
z-score
-2
+2
NANOG
SOX2
TDGF1
KLF17
GATA4
GATA6
PDGFRA
SOX17
GATA3
KRT18
TEAD3
GATA2
Cluster 1
Cluster 2
Cluster 3
C
Relative frequency (percentages)
Cluster 2
Cluster 3
Cluster 1
TE
EPI/PrE
D
PC2
PC1
early/late cells with > 6000 gene expressed log2FPKM > 1.5
early/late cells with < 6000 gene expressed log2FPKM > 1.5
E
PC2
PC1
early ICM cells
early TE cells
late TE cells
late EPI cells
F
PC2
PC1
early ICM cells
early TE cells
late TE cells
late EPI cells
G
A83
N2B27
PD
PDA83
d0
d1
d2
d3
d5
Upregulated in TE
Downregulated in TE
z-score
-2
+2
H
Cluster1/2
Cluster5/6
log2expression > 5
cumsum binarized
I
PC2
PC1
postPA-TE
postL-EPI
postE-EPI
preL-TE
pre-EPI
preE-TE
ICM
PXGL
A83
HNES1
cR-H9
N2B27
PD
PDA83
cynomolgus monkey
ICM
pre-EPI
postE-EPI
postL-EPI
preE-TE
preL-TE
postPA-TE
J
Naive
Primed
log2 expression
TFAP2C
TBX3
HAVCR1


A
mouse embryo
Cluster1
Cluster2
Cluster3
Cluster4
Cluster5
Cluster6
EPI
ICM
eTE
ITE
B
number of features
0 100 300 500
HNES1-PDA83
Cluster1
Cluster2
Cluster3
Cluster4
Cluster5
Cluster6
mouse embryo
C
HNES1-PDA83
Cluster3/4
standardized score
0.3
0.2
0.1
0.0
-0.1
-0.2
EPI
ICM
eTE
ITE
mouse embryo
D
Correlation with
mouse-ICM
Bootstrap of Pearson
0.8
0.7
0.6
0.5
0.4
0.3
PXGL
d1
d2
d3
d5
E
CDK1
CHAF1B
FOXI3
NOC2L
PLK1
PRPF6
C1QBP
ESRRB
GTF3A
HMGB1
HSPA8
LYAR
RPTOR
AURKAIP1
FOXO1
GTF2H3
PLAGL2
POLR3C
ZSCAN22
ACTB
NELFB
RNF6
TBPL1
TRIP13
TTF1
AURKB
CSRNP1
CTDSP1
ELOF1
NFX1
RPRD1B
ZBTB21
TRIM32
TXN
UTP4
ZNF267
ZNF485
ZNF496
ZNF526
ZNF528
ZNF677
ZNF749
ZNF778
PDLGX
D1
D2
D3
D5
EPI
ICM
eTE
ITE
EPI
ICM
eTE
ITE
standardize score
-2
0
2
F
FOXK1
GTF2A2
KDM5B
PPIE
PSPC1
TCF7L1
ZBTB48
ANP32B
CSNK2A1
EWSR1
TLE1
AATF
BLM
EGLN1
ELL
GATA6
NUP98
TET2
CDX2
CIZ1
SATB2
TADA2A
ZBTB11
BASP1
DOT1L
GRHL2
MYC
RUVBL2
ATF3
CITED4
ZNF2
ZNF274
ZNF28
ZNF468
ZNF57
ZNF703
ZNF850
PDLGX
D1
D2
D3
D5
EPI
ICM
eTE
ITE
EPI
ICM
eTE
ITE
standardize score
-2
0
2
G
z-score
-2
0
2
PD
ESRRB
GATA6
hATP6VOD2
hCYP26A1
hGRPC5A
hSLC30A2
ADTRP
PLCE1
SLC12A3
CDX2
d0
d1
d2
d3
PDA8301
ESRRB
GATA6
hATP6VOD2
hCYP26A1
hGRPC5A
hSLC30A2
ADTRP
PLCE1
SLC12A3
CDX2
d0
d1
d2
d3
H
N2B27
PDA83-24h
PDA83-48h
PDA83-72h
relative to ACTB
0.6
0.4
0.2
0.0
DAB2
SLC28A3
TFAP2C
GATA2
GATA3
HAVCR1
I
Naive
PD-N2B27
PD-ACTA
PD-A83
relative to ACTB
0.3
0.2
0.1
0.0
GATA2
GATA3
NANOG
DAB2
SLC28A3
TFAP2C
HAVCR1


A
N2B27
Naive
N2B27
P
B-50
B-10
P+L
L
L+B-50
Primed
N2B27
B-50
L
L+B-50
pSmad1.5.8
Tubulin
Naive HNES1
Primed HNES1(E8)
B
day 3
%mKO
% of Max
BMP2
PDA83
PDA83+BMP2
PDA83+LDN
C
Naive
Primed
PC1
PC2
naive (Stirparo et al., 2018)
primed (Stirparo et al., 2018)
H1-EPSC
ES1-EPSC
(Yang et al., 2017)
PXGL-naive (this study)
H1-ESC
hiPSC.EPSC
H1-EPSC
(Gao et al., 2019)
D
EPI
ICM
earlyTE
lateTE
PC1
PC2
PXGL
A83
N2B27
H1.ESC
H1.EPSC
hiPSC.EPSC
PD
PDA83
HNES1
cR-H9
day
human embryo
EPI
ICM
earlyTE
lateTE
E
AMLC genes > 1
Contribution PC1
Contribution PC2
F
early AM genes
late AM genes
Contribution PC1
Contribution PC2
G
standardize expression
AMNION genes
standardize expression
PDA8301
hiPSC.EPSC
embryo
H
N2B27
A83
BMP2
% of Max
3.81
7.93
23.3
naive
primed
I
GATA2
GATA3
DSG3
TFF1
BAMB1
ISL1
IGFBP3
SEMA3C
relative to ACTB
PD+A83
PD+A83 +BMP
J
Naive
Primed
BRN2
PAX6
SOX1
Relative to ACTB
PSC
N2B27
PD
PD+LDN
PDA83
PD+A83+LDN
K
primed
N2B27
PDA83
PDA83/BMP2
H9-primed
HNES1-primed
d3
d5
EPI
ICM
earlyTE
lateTE
PC1
PC2


A
Control
OCT4
SOX2
NANOG
Exp.1
Exp. 2
B
Control
OCT4
SOX2
NANOG
TFAP2C
YAP1
Well 1
Well 2
Well 3
C
GFP
gOCT4
gSOX2
gNANOG
N2B27
A83
Relative to ACTB
DAB2
GATA2
GATA3
GATA6
HAVCR1
NANOG
SOX2
OCT4
D
Phase
GATA3:mKO
No DOX
+DOX