© 2019. Published by The Company of Biologists Ltd | Development (2019) 146, dev180620. doi:10.1242/dev.180620

Naïve human pluripotent stem cells respond to Wnt, Nodal and LIF signalling to produce expandable naïve extra-embryonic endoderm

Madeleine Linneberg-Agerholm*, Yan Fung Wong*, Jose Alejandro Romero Herrera, Rita S. Monteiro, Kathryn G. V. Anderson and Joshua M. Brickman‡

ABSTRACT

Embryonic stem cells (ESCs) exist in at least two states that transcriptionally resemble different stages of embryonic development. Naïve ESCs resemble peri-implantation stages and primed ESCs the pre-gastrulation epiblast. In mouse, primed ESCs give rise to definitive endoderm in response to the pathways downstream of Nodal and Wnt signalling. However, when these pathways are activated in naïve ESCs, they differentiate to a cell type resembling early primitive endoderm (PrE), the blastocyst-stage progenitor of the extra-embryonic endoderm. Here, we apply this context dependency to human ESCs, showing that activation of Nodal and Wnt signalling drives the differentiation of naïve pluripotent cells toward extra-embryonic PrE, or hypoblast, and these can be expanded as an *in vitro* model for naïve extra-embryonic endoderm (nEnd). Consistent with observations made in mouse, human PrE differentiation is dependent on FGF signalling *in vitro*, and we show that, by inhibiting FGF receptor signalling, we can simplify naïve pluripotent culture conditions, such that the inhibitor requirements closer resemble those used in mouse. The expandable nEnd cultures reported here represent stable extra-embryonic endoderm, or human hypoblast, cell lines.

This article has an associated 'The people behind the papers' interview.

KEY WORDS: Hypoblast, Human, Primitive endoderm, Naïve pluripotency, FGF signalling, Embryonic stem cells

INTRODUCTION

Embryonic stem cells (ESCs) are immortal, karyotypically normal cells derived from the mammalian pre-implantation embryo (reviewed by Smith, 2001). They are defined as pluripotent, as they can differentiate into all lineages of the future conceptus, and have been derived from a number of species, including humans (Thomson et al., 1998). At the stage of traditional ESC derivation, the pre-implantation embryo is in the process of generating three lineages as a result of two successive lineage segregation events. The first of these occurs as a result of morula compaction, leading to formation of the extra-embryonic trophoblast on the outside and the inner cell mass (ICM) on the inside. The blastocyst is then formed following cavitation, and ICM cells differentiate to either pluripotent epiblast (EPI) or the bipotent extra-embryonic primitive endoderm (PrE), also referred to as hypoblast in non-murine species.

ESCs are characterised by expression of a set of transcription factors (TFs) collectively referred to as the pluripotency network, including NANOG, OCT4 (also known as POU5F1) and SOX2, that are widely expressed in both the pre-segregation ICM and in the EPI (Dietrich and Hiiragi, 2007; reviewed by Chambers and Tomlinson, 2009; Morgani and Brickman, 2015; Martin Gonzalez et al., 2016). In addition to the pluripotency network, TFs specific to other lineages of the pre-implantation embryo include CDX2 in the trophoblast, and GATA6 and GATA4 in the nascent PrE (reviewed by Rossant, 2018). In mouse, the segregation of PrE and EPI from the ICM is regulated by FGF/ERK signalling, and the inhibition of this pathway can be used to support ESC self-renewal (Burdon et al., 1999; Chazaud et al., 2006; Yamanaka et al., 2010). This observation has been crucial for the development of defined conditions to derive and maintain ESCs in what has been referred to as the ground state of pluripotency, using a combination of inhibitors of MEK1 (MAP2K1) and MEK2 (MAP2K2) [both inhibited by PD0325901 (PD03)] and GSK-3 [inhibited by CHIR99021 (CHIR)], along with the cytokine LIF (2i/LIF) (Ying et al., 2008).

In contrast to mouse, conventional human ESCs (hESCs) differ in their culture requirements and instead exploit FGF/ERK activity to maintain pluripotency. An explanation for these differences has been that these *in vitro* cell types approximate different stages of embryonic development, and that the components required to support hESCs are instead required for pluripotency at later developmental stages in the mouse (Brons et al., 2007; Tesar et al., 2007). Mouse cell lines derived in human ESC conditions no longer bear resemblance to the peri-implantation embryo, but instead exhibit a gene expression signature and morphology similar to their human counterpart and the pre-gastrulation stages of embryonic development. These have therefore been termed EPI-derived stem cells (EpiSCs). The existence of two distinct cell types indicates that pluripotency exists across several developmental stages (reviewed by Nichols and Smith, 2012). Cells maintained in ground state conditions are considered naïve, whereas cells cultured in the presence of FGF represent a more advanced stage of development, in which cells are 'primed' for differentiation. A number of recent reports have identified culture conditions that capture a more naïve phase of human pluripotency, all of these conditions include 2i/LIF, but vary in the set of additional factors added to this media (Gafni et al., 2013; Takashima et al., 2014; Theunissen et al., 2014; Ware et al., 2014).

In mouse, immortal stem cell lines capable of indefinite self-renewal have been derived from the extra-embryonic, as well as the embryonic, lineages. Trophoblast stem (TS) cells can be derived from the polar trophectoderm and extra-embryonic endoderm (XEN) cells from the PrE (Tanaka et al., 1998; Kunath et al., 2005;

Novo Nordisk Foundation Center for Stem Cell Biology (DanStem), University of Copenhagen, Blegdamsvej 3B, DK-2200 Copenhagen N, Denmark.
*These authors contributed equally to this work

‡Author for correspondence (joshua.brickman@sund.ku.dk)

J.M.B., 0000-0003-1580-7491

Received 20 May 2019; Accepted 11 November 2019

Niakan et al., 2013). However, although derived from the blastocyst, these culture systems share characteristics of the later, more restricted extra-embryonic lineages, as TS cells resemble the extra-embryonic ectoderm and XEN cells resemble the parietal endoderm (Tanaka et al., 1998; Kunath et al., 2005). We recently demonstrated that in the presence of activin A (ActA) and CHIR, mouse naïve ESCs robustly differentiate to bipotent PrE that can be expanded as an *in vitro* naïve extra-embryonic endodermal progenitor (nEnd), capable of contributing to both the visceral and parietal endoderm (Anderson et al., 2017). Although evidence does exist for extra-embryonic endoderm differentiation in human models, none of these studies has produced expandable cell lines in the absence of genetic manipulation or abnormality. SOX7 overexpression in primed hESCs induced extra-embryonic endoderm progenitors that could maintain their phenotype through expansion (Séguin et al., 2008), and cell lines derived from human testicular teratomas have been reported to have either visceral or parietal characteristics (Pera et al., 1987). Furthermore, addition of retinoic acid to human embryonal carcinoma cells induced extra-embryonic endodermal gene expression (Roach et al., 1994). In hESCs, protein kinase C (PKC) stimulated differentiation to an endodermal cell type that undergoes epithelial-to-mesenchymal transition (EMT) and expresses markers of the extra-embryonic parietal endoderm (Feng et al., 2012). More recently, single-cell analysis of primed hESCs revealed a population of pluripotent founder cells that resemble extra-embryonic endoderm (Nakanishi et al., 2019). Taken together, although it appears to be possible to generate PrE cell types *in vitro*, no genetically stable self-renewing cell line or culture system has been reported for human PrE, or hypoblast.

In mammals, endoderm differentiation occurs in two waves, an early extra-embryonic wave giving rise to the PrE, and a later, gastrulation-stage differentiation resulting in the embryonic definitive endoderm (DE). Although these tissues share expression of several markers, they have traditionally been viewed as having different fates, with PrE contributing to the extra-embryonic visceral yolk sac and the majority of the embryonic organs coming from the DE (Lawson et al., 1987). This notion has recently been challenged by fate mapping and single-cell transcriptomics, asserting that mouse PrE can contribute to both the embryonic and extra-embryonic lineages (Kwon et al., 2008; Chan et al., 2019; Nowotschin et al., 2019; Peng et al., 2019; Pijuan-Sala et al., 2019). Moreover, although the induction of these two lineages were seen as modulated by distinct signalling pathways, we have previously shown that, in mouse, endoderm induction is context dependent. Differentiation is mediated by the same pathways acting on distinct developmental cell types, with signalling downstream of Wnt and Nodal inducing PrE from naïve ESCs, and gastrulation-stage DE from EpiSCs (Anderson et al., 2017). Together, these findings suggest that endoderm differentiation in mammals occurs based on common inductive mechanisms acting over a range of developmental stages, in which the stage determines the precise nature and fate of the endoderm induced.

Here, we ask whether context-dependent endoderm induction is evolutionarily conserved, by testing whether stimulation of Wnt and Nodal-related TGF-β signalling can give rise to distinct endodermal cell types depending on the starting population of pluripotent hESCs. We found that addition of ActA and CHIR to naïve and primed hESCs produced PrE and DE, respectively. As in the mouse, we also observed that FGF signalling is required for PrE differentiation *in vitro* and that we could exploit the activity of FGF signalling in PrE differentiation as a means to improve naïve hESC culture. Finally, we are able to expand PrE as nEnd, to produce stable human extra-embryonic endoderm cell lines resembling the *in vivo* hypoblast.

RESULTS

Naïve human pluripotent cells produce primitive endoderm in response to Wnt and nodal-related TGF-beta signalling

To investigate the context-dependent nature of the endoderm fate in the mouse, we previously used a fluorescent reporter mouse ESC line for the endodermal markers *Hhex* and *Hnf4a*, in which the order of induction of these two genes could distinguish between PrE and DE differentiation (Anderson et al., 2017). However, recent single-cell transcriptomic studies of the mouse and human embryo demonstrate that the role of HHEX in human PrE is not conserved and that *HHEX* is uniquely expressed in human DE, whereas *FOXA2* is expressed in both lineages (Fig. S1A) (D'Amour et al., 2005; Blakeley et al., 2015; Boroviak et al., 2018). To determine whether the same context dependence of embryonic versus extra-embryonic endoderm induction is also true in human, we generated a double fluorescent reporter for *HHEX* and FOXA2 in H9 hESCs (H9-HF). This was achieved by CRISPR/Cas9-mediated homologous recombination targeting *H2B-mCherry* to the first exon of *HHEX*, and replacing the endogenous stop codon of *FOXA2* with an in-frame fusion to a linker, followed by the *mVenus* coding sequence (Fig. S1B-F). To understand the relationship between differentiation and proliferation in PrE specification, we also generated a cell cycle *H2B-mCherry-F2A-mVenus-hGem(1-110)* (H9-G2M) reporter by random integration (Fig. S1G-I) in H9 hESCs. All reporter cell lines were generated using primed hESCs cultured in FGF/knockout serum replacement (KSR) medium.

To generate paired versions of primed and naïve cell lines, primed H9 wild-type and reporter cell lines were chemically reset (cR) following transient exposure to a histone deacetylase inhibitor (Guo et al., 2017). These were maintained as naïve hESCs in t2iLGö (naïve^{t2iLGö}) (Takashima et al., 2014) (Fig. S2A,B). Based on our analysis of published transcriptomic data of cR and *in vivo*-derived naïve hESCs, we identified CD53 as a putative cell-surface marker of human naïve pluripotency (Takashima et al., 2014; Guo et al., 2017). Whereas primed H9 expressed high levels of SSEA4, chemically reset H9 (cR-H9) and H9-HF (cR-H9-HF) cells expressed CD53, but not SSEA4 (Fig. S2C,D). This expression pattern was also observed in the embryo-derived human naïve ESCs (HNES1) (Guo et al., 2016) (Fig. S2C). We confirmed the identity of cR cell lines by morphology, immunofluorescence and qRT-PCR analysis of naïve-specific markers, such as KLF17, *DPPA3* and *TFCP2L1*, and the ubiquitous pluripotency markers SOX2 and NANOG (Fig. S2E-G).

We asked whether the primed and naïve H9-HF reporter ESCs could differentiate toward distinct endodermal subtypes in response to Wnt and Nodal signalling (Fig. 1A). In parallel, the cell lines were exposed to ActA, CHIR and LIF in RPMI 1640 GlutaMAX base medium supplemented with B27 minus insulin (RACL) (see Materials and Methods). Following 4-5 days of differentiation, primed H9-HF gave rise to a HHEX-mCherry/FOXA2-mVenus double-positive population, whereas up to 7 days of differentiation with cR-H9-HF did not, suggesting that only primed cells were able to generate DE (Fig. 1B,C). This was consistent with qRT-PCR analysis, which demonstrated the upregulation of various DE markers, such as *MIXL1*, *CXCR4* and *GSC* in RACL-treated primed H9-HF, but not in cR-H9-HF (Fig. 1D). In contrast, the figure also shows that RACL-treated cR-H9-HF cells expressed high levels of the early PrE marker *PDGFRA* (Artus et al., 2010), and the PrE marker *NID2* identified from single-cell sequencing data of the human blastocyst (Petropoulos et al., 2016). These data, and others

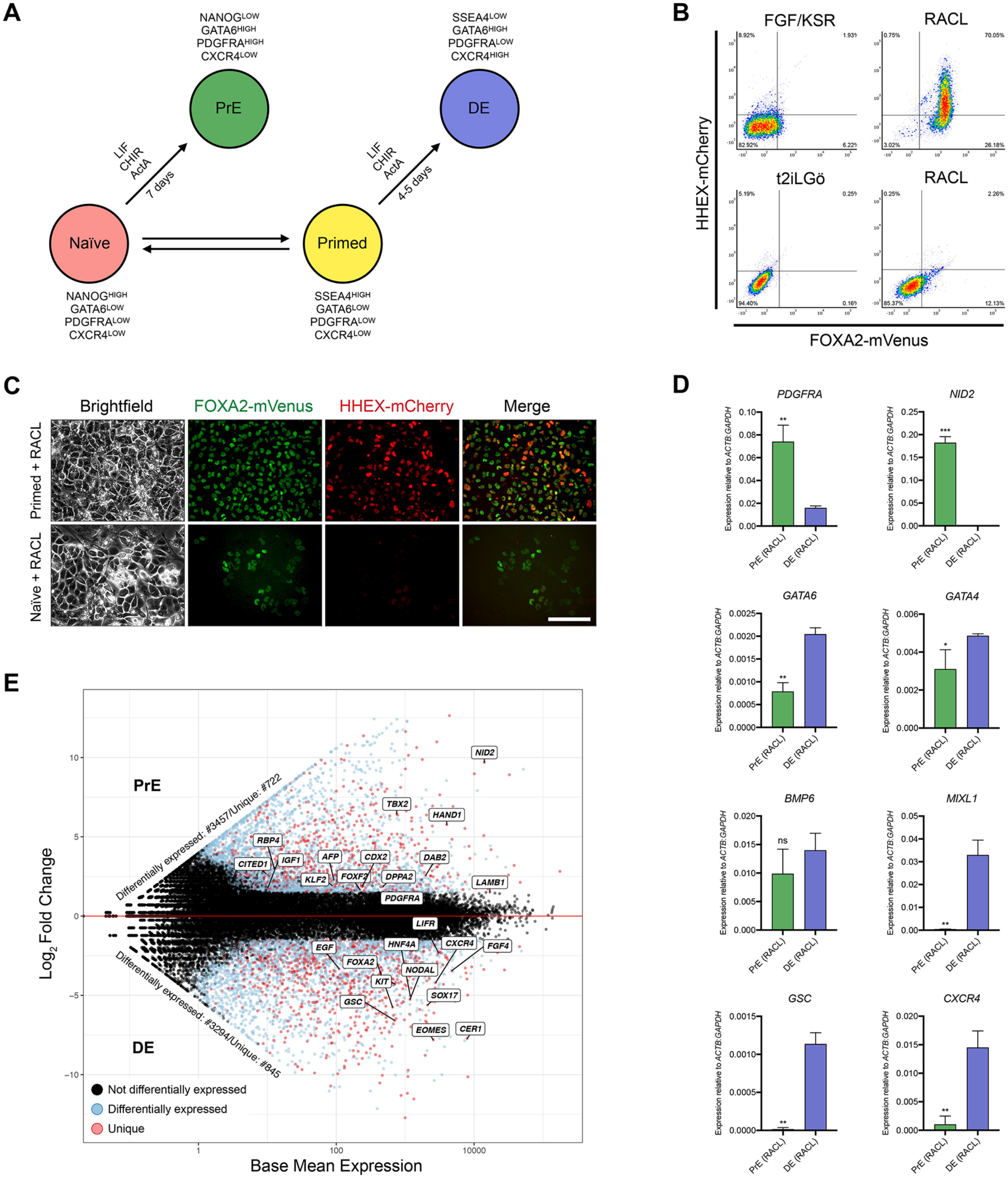


Fig. 1. Endoderm differentiation induced by Nodal and Wnt signalling is context dependent. (A) Schematic illustrating the approach to differentiation of primed hESCs to DE and naïve hESCs to PrE. (B) Flow cytometry density plots showing H9-HF reporter expression (top) in primed media (FGF/KSR) and RACL, as well as cR-H9-HF (bottom) in naïve media (t2iLGö) and RACL. For cR-H9-HF, voltage was compensated against autofluorescence of Gö6983. Bottom left quadrant indicates gating based on a negative control. (C) Primed H9-HF-derived DE in RACL showing high levels of FOXA2-Venus and HHEX-mCherry expression, and naïve cR-H9-HF-derived PrE showing FOXA2-Venus expression, but not HHEX-mCherry. Cells were imaged live by fluorescent microscopy. (D) qRT-PCR measuring expression of markers specific to naïve hESCs, PrE and DE in H9-HF-derived DE and cR-H9-HF-derived PrE. All were normalised to the expression of *ACTB* and *GAPDH*. *P<0.05, **P<0.01, ***P<0.001 (t-test). Error bars indicate ±s.d. ns, not significant. (E) MA-plot representing differential expression analysis of PrE compared with DE following differentiation in RACL, where blue dots indicate genes that are differentially expressed and red dots indicate genes that are uniquely expressed. ($\log_2$FC>1.5, P<0.05). Scale bar: 50 µm.

(Blakeley et al., 2015), also suggest BMP6 as a potential marker of human PrE. However, we observed *BMP6* mRNA to be expressed in both endodermal cultures, as were the endodermal TFs *GATA6* and *GATA4* (Fig. 1D). When the transcriptome of *in vitro*-derived human PrE was directly compared with DE, we observed considerable differences between these cell types, with 3457 genes upregulated and 3294 downregulated. As some of these differences can be attributed to developmental stage (i.e. primed or naïve), we filtered these datasets for genes that were also differentially expressed between the two hESC populations. From this we identified 722 and 845 genes uniquely up- and downregulated in PrE compared with DE, respectively, further suggesting that these are two distinct *in vitro* endodermal populations (Fig. 1E, Table S1A-D).

To probe the kinetics of endoderm differentiation from naïve hESCs, we assessed the timing of NANOG and GATA6 expression by immunofluorescence. Although only NANOG$^+$ cells were observed during the first few days of differentiation, a transient NANOG/GATA6 double-positive population emerged between days 3 and 4, and these cells eventually resolved their identity, with a GATA6 single-positive population arising from day 5 onwards. By day 7, all NANOG-expressing cells were lost and the culture was comprised entirely of GATA6$^+$ putative PrE colonies (Fig. 2A). Similar expression kinetics were observed by qRT-PCR (Fig. 2B). Pluripotency genes, such as *NANOG*, *DPPA3* and *OCT4*, were downregulated by day 4 of differentiation, and panendodermal genes *GATA6* and *GATA4* were upregulated from day 3. Specifically, we observed an initial upregulation of *GATA6* at day 2 followed by *GATA4* from day 3, mirroring the *in vivo* sequence of TF activation observed during mouse pre-implantation development (Yamanaka et al., 2010). We also observed an upregulation of the human PrE-specific marker *NID2* (Fig. 2B). Whole-transcriptome analysis revealed that a number of genes specifically upregulated in PrE compared with EPI *in vivo*, are upregulated in PrE derived from naïve hESCs (Yan et al., 2013; Blakeley et al., 2015) (Fig. S3A). The derivation of PrE from naïve hESCs was corroborated by similar experiments performed with two other naïve hESC lines, HNES1 and cR-Shef6 (Takashima et al., 2014; Guo et al., 2017) based on morphology (Fig. 2C), gene expression (Fig. 2D) and flow cytometry for PDGFRA (Fig. S3B).

Laminin-511 supports feeder-independent differentiation of PrE

A key difference between mouse and human naïve ESC cultures is the substrates on which they are propagated. Naïve hESCs are typically cultured on feeders, whereas mouse ESCs (mESCs) are cultured in feeder-free conditions, primarily on gelatine. Although naïve mESCs were originally cultured on feeders, the factor produced by these feeders that was sufficient to support pluripotency was identified as LIF (Smith et al., 1988). We therefore asked whether PrE differentiation could proceed in the absence of feeders.

Based on the previous observations that naïve hESCs could be stably maintained on a truncated laminin-511 fragment (LN511-E8) (Takashima et al., 2014; Guo et al., 2016), we attempted to culture these on the commercially available full-length laminin-511 (LN511, BioLamina). We confirmed that cR-H9 maintained their naïve phenotype by assessing morphology (Fig. S4A), CD53 expression (Fig. S4B) and their transcriptome (Fig. S4C, Table S2A-B), from which we confirmed individual naïve markers (Fig. S4D). As we found LIF to be essential for PrE specification in mouse (Morgani and Brickman, 2015; Anderson et al., 2017), we rationalised the continued use of the cytokine in human differentiation, and based on the transcriptome of the resulting cultures, PrE differentiation appeared to be feeder independent (Fig. S4E-G). Principal component analysis (PCA) of these datasets highlights the similarities in gene expression profiles of cells cultured and differentiated on LN511 or mouse embryonic fibroblasts (MEFs). PC1 separates naïve pluripotent cells from PrE (80%), whereas the small amount of variance in PC2 resolves any differences resulting from the different substrates (6%, accounting for a variation of <130 genes in both states) (Fig. S4E). Furthermore, cells differentiated to PrE on LN511 maintained similar levels of expression of PrE lineage markers to those differentiated on MEFs (Fig. S4G, Table S2C-D).

FGF signalling is required for PrE differentiation *in vitro*.

Although the role of FGF/ERK signalling in PrE specification is well established in the mouse, both *in vitro* and *in vivo*, its role in primates has remained unclear (Arman et al., 1998; Nichols and Smith, 2009; Yamanaka et al., 2010; Kuijk et al., 2012; Roode et al., 2012; Anderson et al., 2017). Based on culturing human blastocysts *ex vivo*, FGF signalling does not appear to be required for PrE differentiation. Treatment of embryos with inhibitors of FGF signalling [MEK1/2 inhibitor PD03 and FGFR1 inhibitor PD173074 (PD17)] does not affect the emergence of mutually exclusive NANOG- and GATA4-positive cells, with the hypoblast forming normally (Roode et al., 2012). However, as PD03 is included in most formulations of human naïve ground state conditions, coupled with its function in the mESCs to block PrE differentiation (Hamilton and Brickman, 2014), we reasoned that it may serve as an inhibitor of PrE differentiation *in vitro*.

We asked whether FGF signalling was required for specification of PrE from naïve hESCs *in vitro* in the presence of RACL. We found that both 1 μM PD03 and 1 μM PD17, when applied individually or combined, reproducibly blocked differentiation of cR-H9 to PrE (Fig. 3A-F). After 7 days with inhibitors, cells failed to upregulate GATA6 and GATA4 or downregulate NANOG (Fig. 3C,D). We observed that both PD03 and PD17 attenuated the upregulation of PDGFRA, and that although PD03 significantly upregulated CD53, both conditions with PD17 produced further increases in naïve surface marker expression (Fig. 3E). Based on qRT-PCR, inhibition of FGFR1 and MEK1/2 resulted in increased expression of *NANOG*, *OCT4* and *KLF17*, and blocked PrE gene expression (Fig. 3F). Although the inclusion of PD03 and PD17 to RACL resulted in a high level of cell death (Fig. 3B), the morphology of the remaining colonies appeared to be naïve-like (Fig. 3A).

FGFR1 inhibition supports naïve pluripotency in the absence of an atypical PKC inhibitor

Although the culture conditions for both mouse (2i/LIF) and human (t2iLGö) naïve ESCs are remarkably similar, with both sharing four key components (N2B27 supplemented with CHIR, PD03 and LIF), one notable difference is the addition of the broad-spectrum PKC inhibitor Gö6983 (Gö). Gö is reported to produce strong autofluorescence from bright red aggregates both *in vivo* and *in vitro*, making the use of fluorescent reporters challenging (Saiz et al., 2013; Takashima et al., 2014; Guo et al., 2017). We also observed this by fluorescent microscopy and flow cytometry analysis (Fig. S5A,B).

We attempted short-term culture of cR-H9 in the absence of Gö as a means to improve imaging, but found that these cells lost their naïve morphology, surface marker expression, and showed reduced expression of the naïve markers *NANOG*, *OCT4* and *KLF17* (Fig. 4A-C). As the knockdown phenotype of the atypical PKC (aPKC) isoform iota/lambda (aPKC$_{\iota/\lambda}$) (Takashima et al., 2014) demonstrated its role in supporting naïve pluripotency, we asked whether we could substitute Gö for another inhibitor targeting aPKC, including CRT0066854, PKC412 and ZIP. However, at all

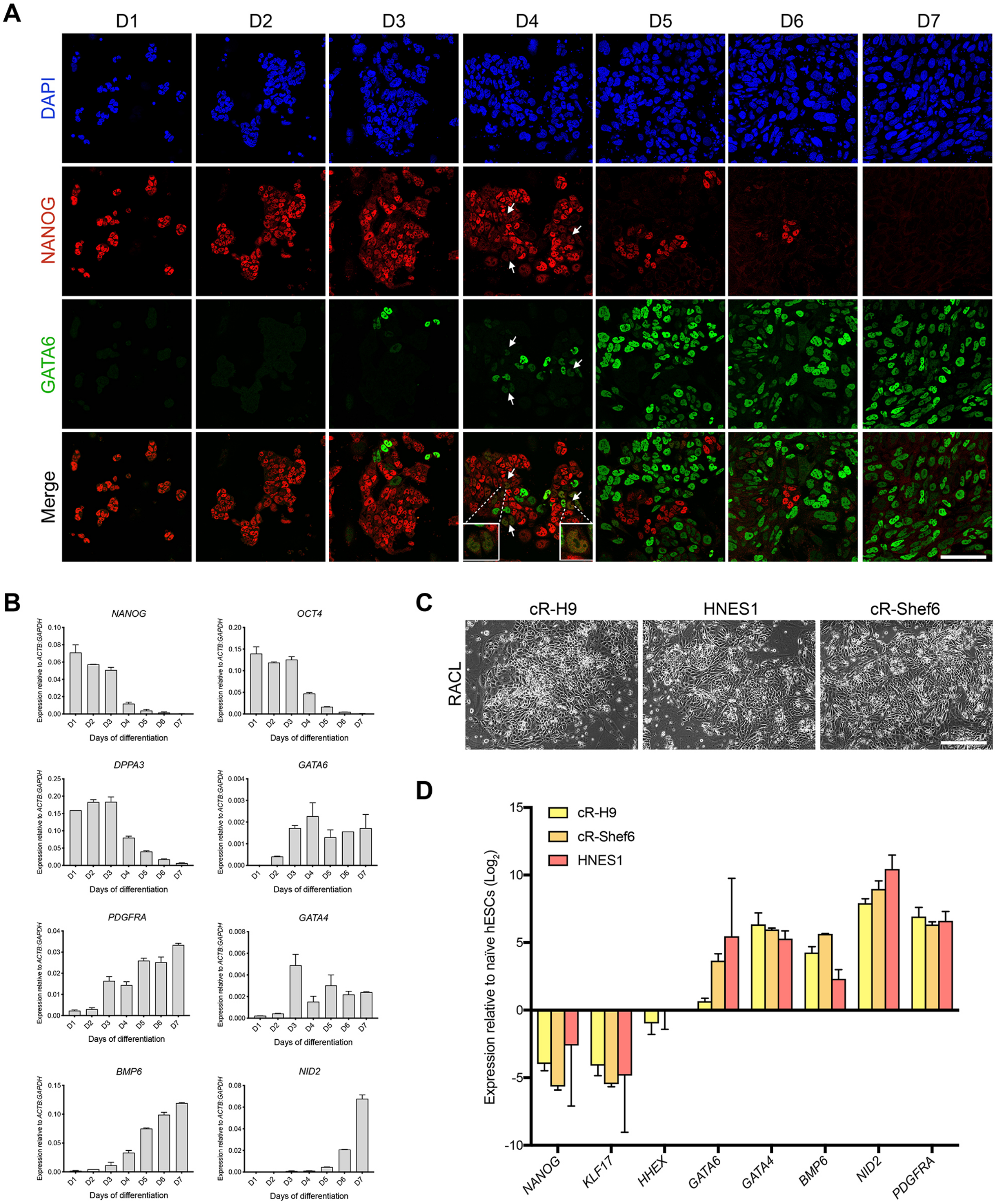

A
D1
D2
D3
D4
D5
D6
D7
DAPI
NANOG
GATA6
Merge
B
NANOG
OCT4
DPPA3
GATA6
PDGFRA
GATA4
BMP6
NID2
Expression relative to ACTB:GAPDH
Days of differentiation
D1
D2
D3
D4
D5
D6
D7
C
cR-H9
HNES1
cR-Shef6
RACL
D
cR-H9
cR-Shef6
HNES1
Expression relative to naïve hESCs (Log$_2$)
15
10
5
0
-5
-10
NANOG
KLF17
HHEX
GATA6
GATA4
BMP6
NID2
PDGFRA

Fig. 2. Differentiation to extra-embryonic PrE from naïve hESCs. (A) Immunostaining of cR-H9 PrE differentiation from day 1 (D1) to day 7 (D7), staining for NANOG and GATA6, including DAPI, with arrows indicating co-expression. Imaged by confocal microscopy. Insets (D4) show GATA6/NANOG double-positive cells. (B) Quantification of mRNA expression as determined by qRT-PCR of different pluripotency and PrE markers during a time course for the 7-day differentiation. All were normalised to the expression of *ACTB* and *GAPDH*. (C) Brightfield images of PrE differentiation across three cell lines at day 7. (D) qRT-PCR for naïve and PrE markers at day 7 of PrE differentiation relative to naïve hESCs across three cell lines. Error bars indicate ±s.d. Scale bars: 25 µm in A; 100 µm in C.

tested concentrations, hESCs lost their naïve morphology, similar to cells cultured without Gö, and some inhibitors caused cell death at high concentrations (Fig. S5C).

Based on our observation that 1 µM PD17 could block PrE differentiation *in vitro*, and that surviving cells in this condition retained a naïve-like morphology and high levels of CD53, we asked

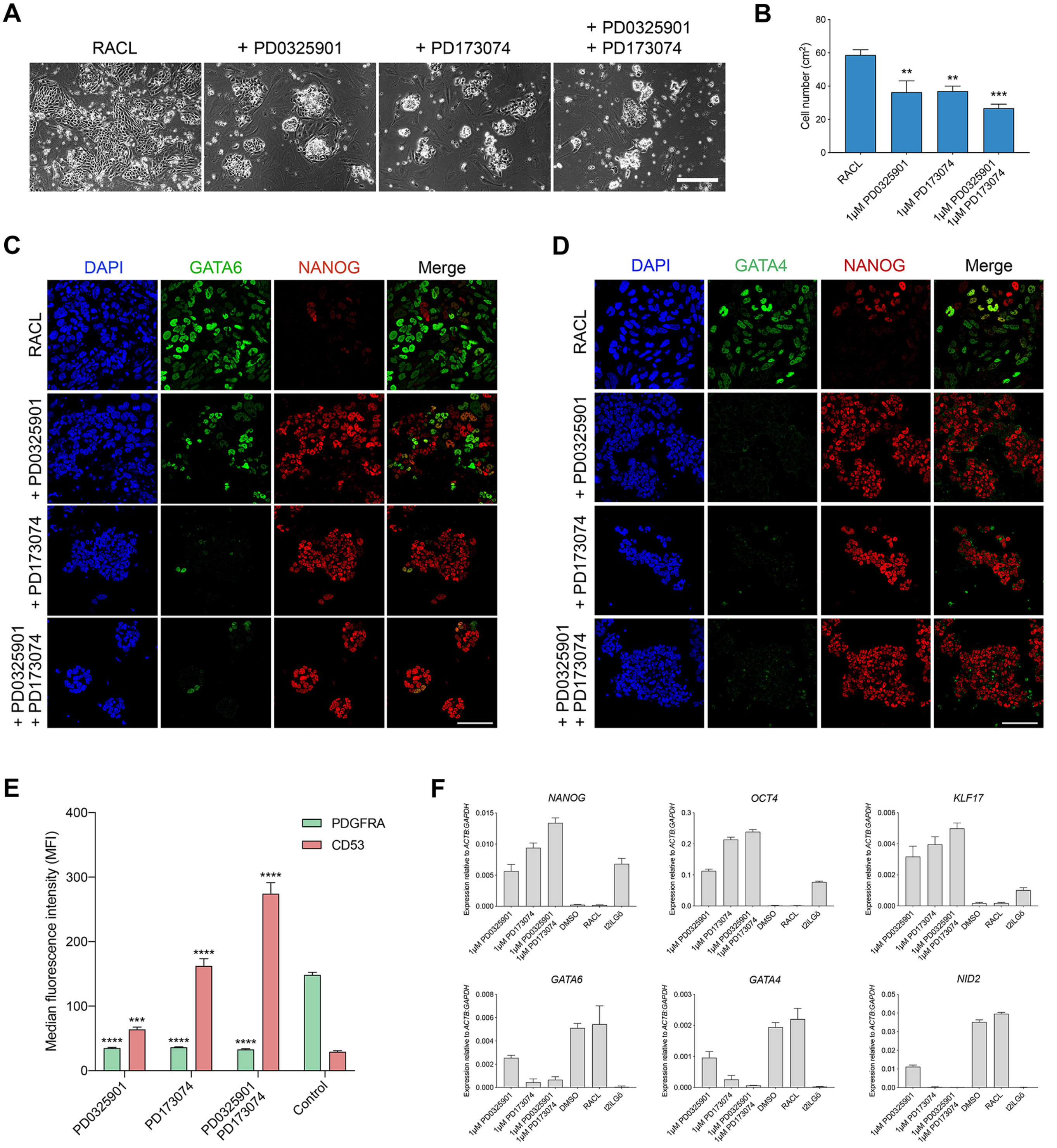


Fig. 3. PrE specification is dependent on FGF signalling *in vitro*. (A-F) Naïve hESCs were differentiated in the presence of the inhibitors PD0325901 (PD03, 1 µM) and/or PD173074 (PD17, 1 µM). Inhibitors were added on day 0 and maintained for 7 days. (A) Brightfield images of the three conditions and control differentiation at day 7. (B) Cell numbers at day 7 of differentiation. (C,D) Immunostaining of the three conditions and control differentiation for the indicated markers including DAPI, imaged by confocal microscopy. (E) Differential induction of PrE-specific PDGFRA and CD53 as determined by flow cytometry analysis, shown by median fluorescence intensity across the three conditions and control differentiation. (F) qRT-PCR for expression of pluripotency and PrE markers during differentiation in response to different FGF/ERK inhibitors, with DMSO as a control for the inhibitors, RACL as the control for differentiation and naïve hESCs in t2iLGö as the negative control. $^{**}P<0.01$, $^{***}P<0.001$, $^{****}P<0.0001$ (*t*-test). Error bars indicate ±s.d. Scale bars: 100 µm in A; 25 µm in C,D.

whether the FGFR1 inhibitor could replace Gö in naïve hESC culture. We cultured cR-H9 in N2B27 supplemented with 1 µM PD17, 1 µM CHIR and 10 ng/ml LIF (human-formulated 2iL, h2iL). Under these conditions, we found that hESCs could maintain their naïve morphology (Fig. 4A), CD53 expression (Fig. 4B), and differentiate to PrE (Fig. 4D). Importantly, cells cultured in h2iL (naïve$^{\mathrm{h2iL}}$) lost any Gö-derived autofluorescence within 3 passages (Fig. S5A,B).

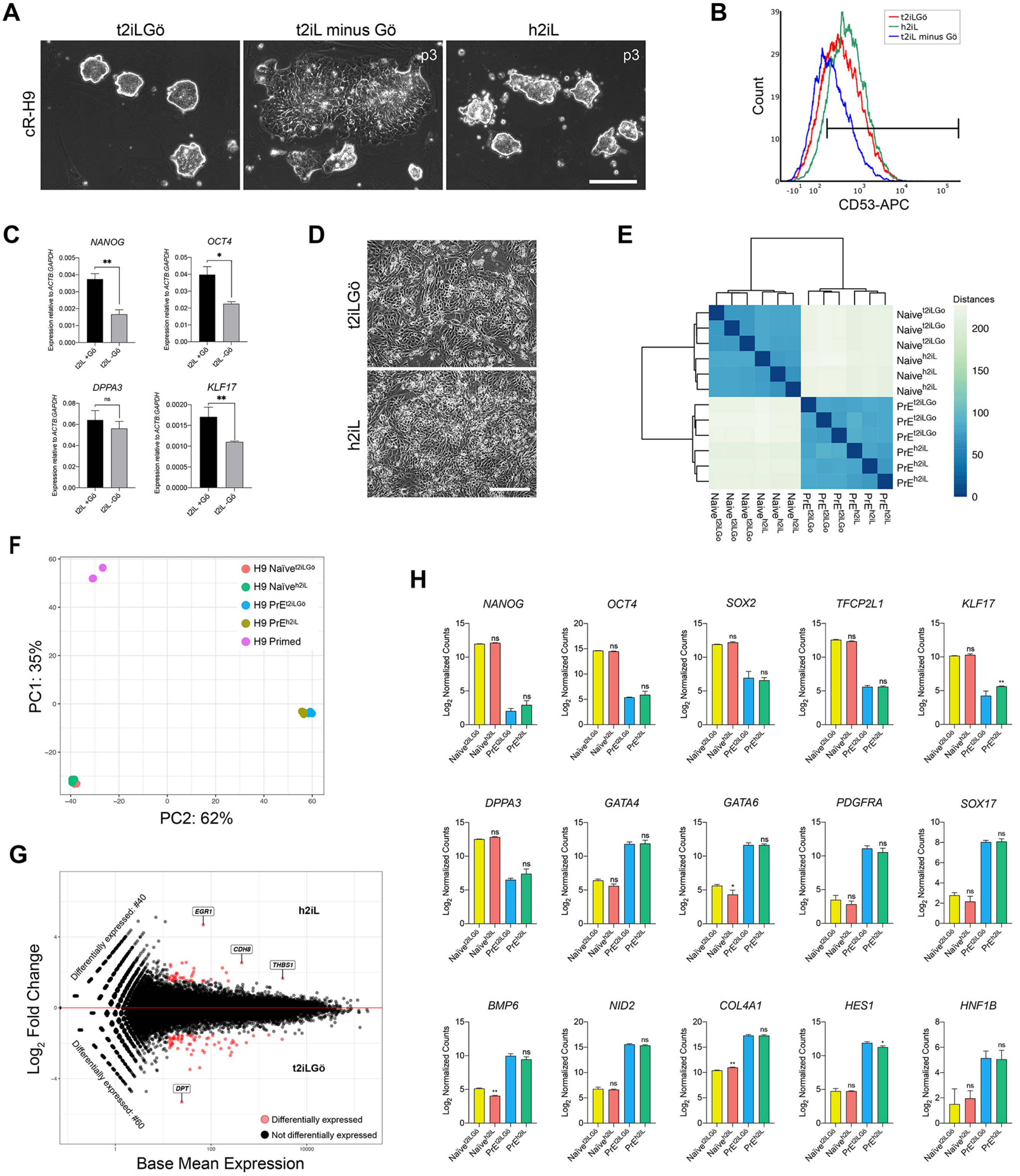


Fig. 4. FGFR1 inhibitor maintains naïve pluripotency in the absence of the aPKC inhibitor Gö6983. (A) Brightfield images showing cR-H9 cultured in naïve medium (t2iLGö), t2iL minus Gö (1 µM CHIR, 1 µM PD03, 10 ng/ml LIF) and h2iL (1 µM CHIR, 1 µM PD17, 10 ng/ml LIF). Cells were maintained in each condition for at least 3 passages. (B) Quantification of fluorescence distribution of CD53 across the three conditions as determined by flow cytometry analysis. Horizontal bar indicates gating based on the negative control with no antibody. (C) qRT-PCR showing expression of naïve pluripotency genes relative to *ACTB* and *GAPDH*. (D) Brightfield images showing cR-H9 cultured in either t2iLGö or h2iL at day 7 of differentiation in RACL. (E) Unbiased hierarchical clustering of whole-transcriptome RNA-seq datasets for naïve^{h2iL} and naïve^{t2iLGö}, as well as PrE^{h2iL} and PrE^{t2iLGö}. (F) PCA was performed based on log_2 normalised counts for gene expression (n=500 genes) of naïve conditions and PrE derived from these compared with primed hESCs in FGF/KSR. (G) MA-plot representing differential expression analysis of naïve^{h2iL} versus naïve^{t2iLGö} (>1.5 log_2 FC, P<0.05), red dots indicate genes that are differentially expressed. (H) Comparative expression of pluripotency and endodermal markers across all conditions described in E based on transcriptomic data. *P<0.05, **P<0.01 (Tukey's multiple comparison test). Error bars indicate ±s.d. ns, not significant. Scale bars: 50 µm in A; 100 µm in D.

To confirm that these h2iL cells were indeed naïve, we assessed their transcriptomes and compared them with cells cultured in t2iLGö, along with PrE derived from both culture conditions by RNA-seq. Based on unbiased hierarchical clustering, we observed a strong correlation between naïve$^{\mathrm{h2iL}}$ and naïve$^{\mathrm{t2iLG\ddot{o}}}$, and between the h2iL- and t2iLGö-derived PrE (PrE$^{\mathrm{h2iL}}$ and PrE$^{\mathrm{t2ilG\ddot{o}}}$, respectively) (Fig. 4E). When differentiated PrE gene expression was compared with both naïve states, these differentiation-specific changes were strongly correlated (Fig. S5D). PCA comparing the two naïve culture conditions with endoderm derived from these, in addition to primed pluripotent cells, suggests that the transcriptomes of naïve$^{\mathrm{h2iL}}$ and naïve$^{\mathrm{t2iLG\ddot{o}}}$, and PrE$^{\mathrm{h2iL}}$ and PrE$^{\mathrm{t2iLG\ddot{o}}}$ are strikingly similar (Fig. 4F). These observations were also supported by differential gene expression analysis between naïve$^{\mathrm{h2iL}}$ and naïve$^{\mathrm{t2iLG\ddot{o}}}$ that revealed 40 upregulated and 64 downregulated genes (Fig. 4G, Table S3A,B). Gene ontology (GO) analysis of this small set of differentially expressed genes revealed enrichment in specific terms associated with extracellular matrix components. Similar observations were made when comparing PrE$^{\mathrm{h2iL}}$ and PrE$^{\mathrm{t2iLG\ddot{o}}}$ (Fig. S5E; Table S3C,D). When inspecting normalised expression for a panel of pluripotency and extra-embryonic endoderm genes, we found no overall difference between the two naïve ESC conditions (Fig. 4H). When placing these cells back into primed culture, they readily converted to conventional hESCs expressing high levels of SSEA4 (Fig. S5F,G). As a relatively high concentration of PD17 (1 µM) was required (Fig. S5H), we considered a possible interaction with the related VEGF and EGF receptors (VEGFR and EGFR, respectively). We tested the ability of specific inhibitors of VEGFR (CBO-P11) (Deshayes et al., 2011) and EGFR (PD153035) (Bos et al., 1997) to mimic the activity of PD17 and found that neither was sufficient to maintain naïve pluripotency (Fig. S5I,J). Based on this observation, PD17 does not appear to be acting on lower affinity targets. Thus, together with LIF and GSK-3 inhibition, PD17 supports naïve hESC self-renewal, indicating that the underlying mechanisms governing naïve pluripotency may be better conserved than has previously been thought.

Human primitive endoderm can be expanded as an *in vitro* model for naïve extra-embryonic endoderm

In mouse, we recently identified conditions that support self-renewal in extra-embryonic PrE and that this was based on the activity of ActA, CHIR and LIF. As these cells can contribute to both parietal and visceral endoderm *in vivo*, we referred to these cultures as nEnd (Anderson et al., 2017). Here, following differentiation of the cR-H9-G2M reporter to PrE in RACL, we observed that these cultures contained proliferating cells, indicated by green fluorescing G2/M cells co-expressing GATA6 (Fig. S6A,B). However, the rate of cell death was significant and exceeded growth upon passaging, such that the culture expired in 3 passages (Fig. S6C,D). As mouse nEnd contains high levels of insulin-responsive AKT activation, including phosphorylations that are known responses to the PI3K and mTOR pathways (Anderson et al., 2017), we asked whether we could improve human nEnd survival by supplementing insulin to RACL. Here, we found that PrE cultures expanded in a dose-dependent manner to insulin (Fig. S6E). We recently reported that other components of the insulin-rich N2B27 medium also supported expansion of naïve mESCs in a similar cytokine cocktail to that used in PrE differentiation (Anderson et al., 2017). We therefore also assessed replacing the RACL basal media with N2B27 [supplemented with ActA, CHIR and LIF (NACL)] and found that these cells could be maintained for more than 13 passages (Fig. 5A). During PrE differentiation, flow cytometry analysis revealed a population expressing GATA6 that was heterogeneous with respect to PDGFRA. This resolves in expansion as an exclusive double-positive nEnd population (Fig. 5B). These consistently expressed PrE markers (Fig. 5C, Fig. S6F) and immunofluorescence demonstrated that these cultures are GATA6^{+}/NANOG^{-} (Fig. 5D). This is consistent with the observed uniform morphology of nEnd, in which ESCs do not re-emerge at any point during expansion. Following an initial passaging ratio of 1:2, nEnd cultures were routinely passaged 1:4 for 13 passages, which constitutes a 3×10^{7}-fold expansion. Therefore, as in mouse, human nEnd can be expanded as an *in vitro* model for extra-embryonic endoderm development. Our approach to establishing this culture system is summarised in Fig. 5E, with an overview of passaging techniques and media compositions given in Fig. S6G,H. Although FGF/ERK signalling is required for PrE differentiation *in vitro*, it is not required for self-renewal of nEnd (Fig. S6I,J). This suggests that the role of FGF/ERK inhibition in suppressing differentiation is not based on a selective mechanism.

During early mammalian development, PrE and EPI precursors within the ICM undergo cell sorting. The PrE is the first of the two cell types to become polarised, forming an epithelialised sheet that produces basement membrane (BM) components, facilitating subsequent embryonic development (Koutsourakis et al., 1999; Smyth et al., 1999; Gerbe et al., 2008; Schrode et al., 2014). To determine the functional capacity of nEnd, we assessed BM production. Here, we found that nEnd produces BM factors normally expressed in both human and marmoset hypoblast *in vivo* (Blakeley et al., 2015; Boroviak et al., 2018), including FN1, LAMA1 and COL4A1, all of which are notably absent in naïve hESCs (Fig. 5F,G). In support of this observation, we also found that expression of mRNA encoding a range of additional hypoblast-associated BM components was upregulated in nEnd relative to naïve hESCs (Fig. S6K).

In mouse, XEN cells can be converted to a visceral endoderm (VE)-like cell type through culture on laminin- or gelatine-coated substrates in the presence of BMP4 (Artus et al., 2012; Paca et al., 2012). We previously reported similar observations when applying these conditions to mouse nEnd (Anderson et al., 2017). We therefore assessed the ability of human nEnd to generate VE by culturing these cells on a range of substrates (Fig. 5H). nEnd cultured on LN511 in N2B27 supplemented with 50 ng/ml BMP4 resulted in the downregulation of *GATA6* and *PDGFRA*, upregulation of VE markers *AFP* and *HNF4A*, and suppression of basal levels of the parietal marker *PLAT*. We observed an increase in a second parietal marker, *SPARC*, but this has also been reported in the VE *in vivo* (Mason et al., 1986).

Transcriptomic states of naïve epiblast and endoderm

As functional tests of PrE identity in human are ethically restricted, we sought to validate our cell culture model based on *in vivo* gene expression from the human blastocyst. We utilised a panel of EPI- and PrE-specific markers defined from single-cell transcriptomic analysis of the human blastocyst (Table S4A) (Stirparo et al., 2018). Here, we included the transcriptomes of naïve and primed hESCs, alongside putative PrE. We found that genes specifically associated with EPI in the human blastocyst also exhibit a strong correlation with naïve hESCs, whereas those associated with endoderm show the highest level of correlation with PrE and nEnd (Fig. 6A).

To understand the relative contribution of these EPI and PrE-specific markers in separating ESCs from different endoderm subtypes (i.e. PrE versus DE), we performed PCA on all samples using only these predetermined lineage markers (Fig. 6B). This

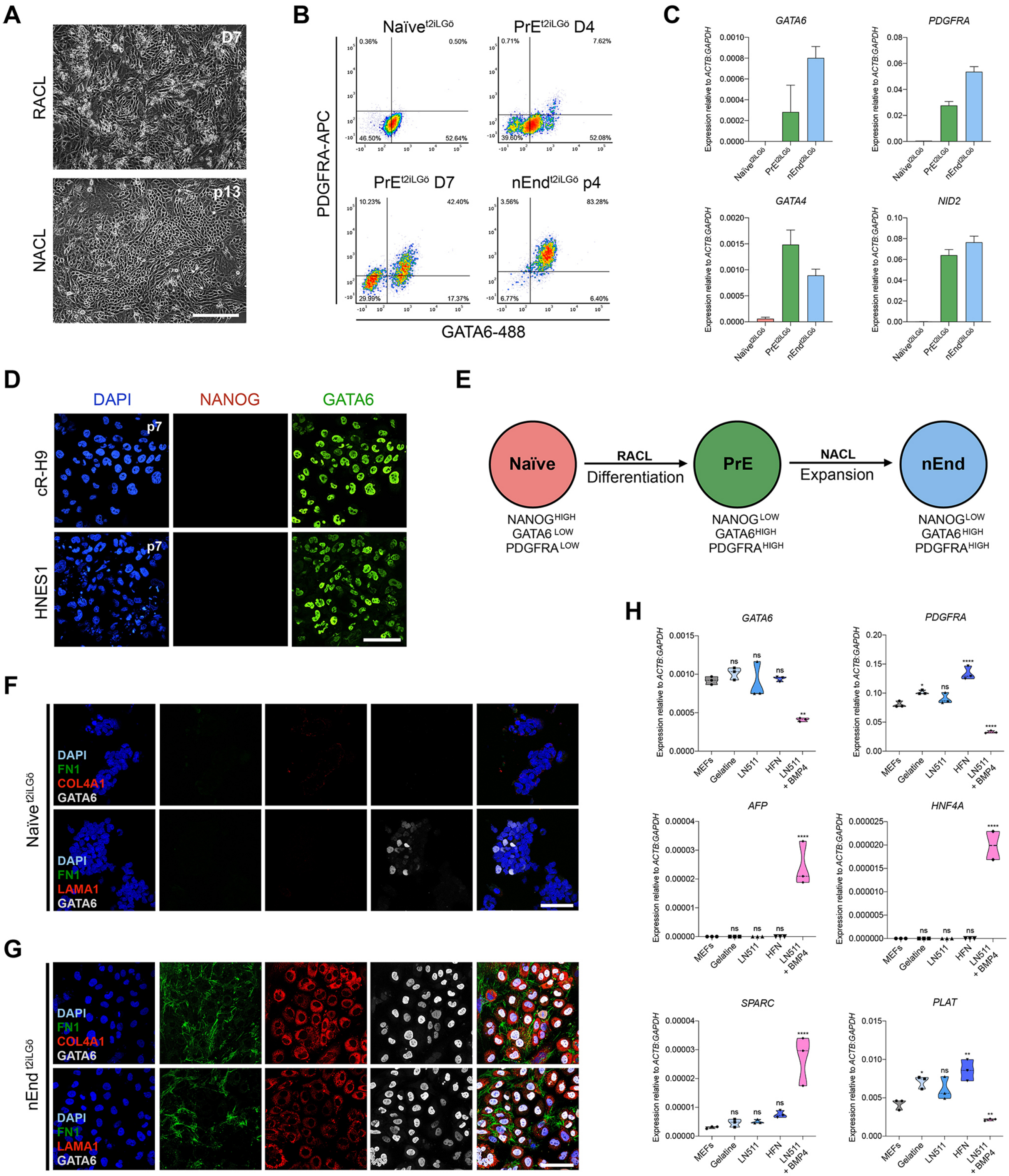


Fig. 5. Expansion of PrE as an *in vitro* model for extra-embryonic naïve endoderm. (A) Brightfield images of cR-H9 naïve$^{\text{t2iLGö}}$ differentiated to PrE in RACL (day 7) and passaged in NACL medium (passage 13). (B) Expression of the cell-surface marker PDGFRA and intracellular staining for GATA6 in naïve$^{\text{t2iLGö}}$ hESCs, during PrE differentiation (day 4 and day 7), and in nEnd (passage 4). Flow cytometry density plots show quadrants based on gating of the negative control and is representative of five differentiations. (C) qRT-PCR showing relative expression of PrE markers in cR-H9 at day 7 of PrE differentiation and nEnd at passage 10 in NACL medium, compared with naïve$^{\text{t2iLGö}}$. (D) Confocal images of immunostaining of nEnd in NACL at passage 7, cultures derived from cR-H9 and HNES1 ESCs, stained for NANOG and GATA6, including DAPI. (E) Schematic of differentiation to PrE and subsequent expansion as nEnd. (F,G) Immunostaining showing BM production in naïve$^{\text{t2iLGö}}$ hESCs (F) compared with nEnd$^{\text{t2iLGö}}$ (G) for the indicated markers, including GATA6 and DAPI. (H) Violin plot of qRT-PCR showing relative expression of primitive, visceral and parietal endoderm markers in nEnd$^{\text{t2iLGö}}$ cultured on different substrates [MEFs, gelatine, LN511 and human fibronectin (HFN)] and in conditions promoting visceral identity (LN511+50 ng/ml BMP4). The dotted lines indicate the median value of the triplicates contained in the violin plot. *$P<0.05$, **$P<0.01$, ****$P<0.0001$ (*t*-test). Error bars indicate ±s.d. ns, not significant. Scale bars: 100 µm in A; 25 µm in D,F,G.

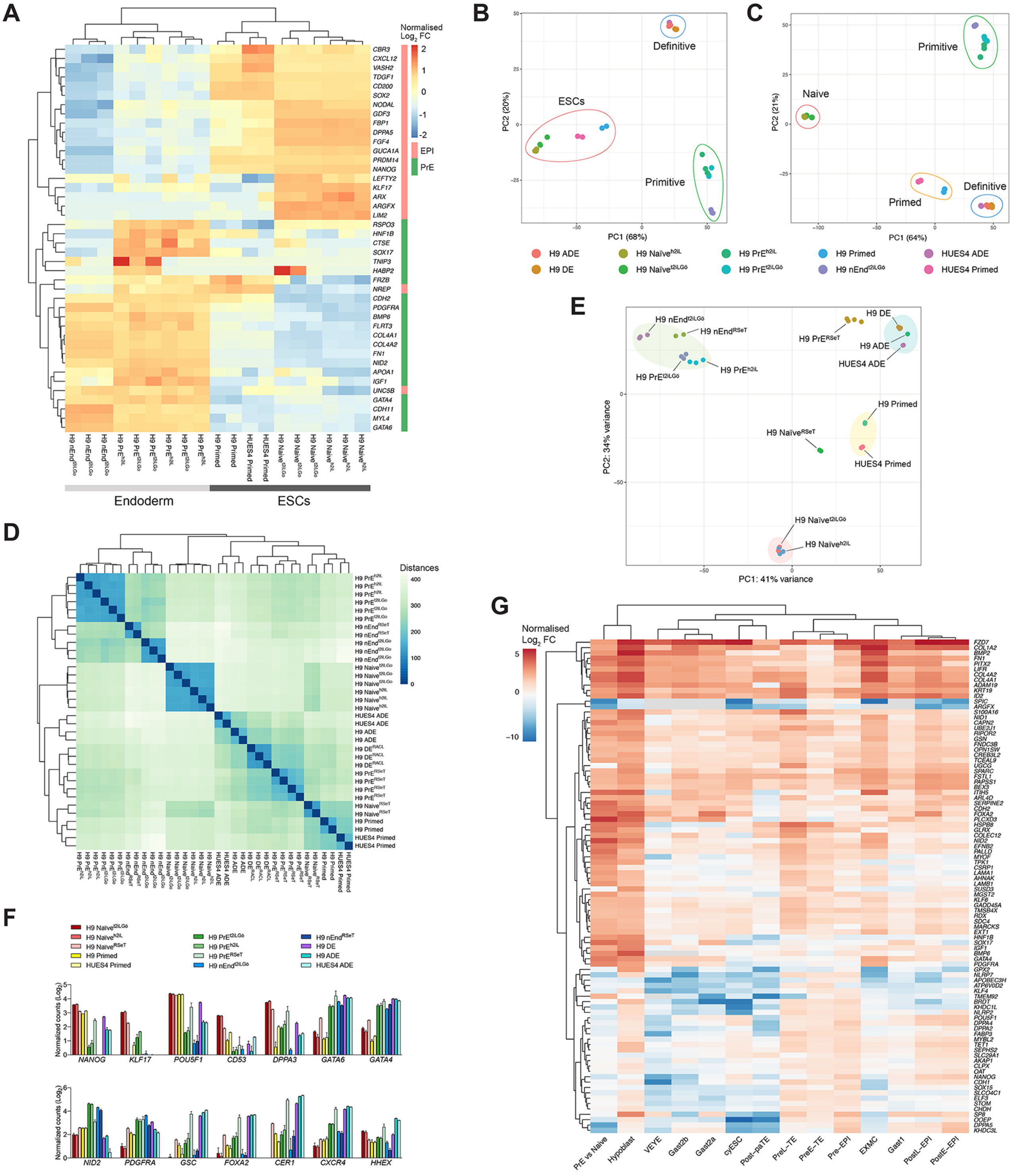


Fig. 6. Transcriptomic states of naïve epiblast and endoderm. (A) Clustered expression data from ESC and endodermal types for a panel of EPI and PrE markers (n=42 genes) specific to the human blastocyst (Stirparo et al., 2018). (B,C) PCA of transcriptomic data for all samples clustered against a panel of EPI- and PrE markers (n=42 genes) identified in the human blastocyst (B) (Stirparo et al., 2018), and naïve and primed markers (n=24 genes) identified *in vitro* (C) by single-cell RNA-seq (Messmer et al., 2019). Indicated clusters correspond to specific cell types and stages of development or differentiation. (D) Unbiased hierarchical clustering of whole-transcriptome datasets of various types of ESCs and endoderm, including ESCs and PrE derived from the RSeT naïve culture system, showing sample-to-sample distances. Sequencing was performed on biological duplicates or triplicates. (E) PCA of the samples in D, performed based on log$_2$ normalised counts for gene expression of all samples (n=1000 genes). (F) Comparative expression of pluripotency and endodermal markers across all samples. Error bars indicate ±s.d. (G) Comparative analysis of the fold changes (log$_2$ FC>1.5, P<0.05) of common genes (n=92 genes) between cynomolgus monkey visceral endoderm/yolk-sac endoderm (VEYE)/hypoblast versus pre-implantation EPI (Pre-EPI)/ICM (Nakamura et al., 2016, 2017) and human PrE/nEnd versus naïve hESCs. Columns and rows are grouped by hierarchical clustering. cyESC, cynomolgus ESCs; EXMC, extra-embryonic mesenchyme; Gast1, gastrulating cells 1; Gast2a, gastrulating cells 2a; Gast2b, gastrulating cells 2b; PreE-TE, pre-implantation early TE; PreL-TE, pre-implantation late TE; PostE-EPI, post-implantation early EPI; PostL-EPI, post-implantation late EPI; Post-PaTE, post-implantation parietal TE.

analysis included *in vitro*-differentiated PrE, nEnd and DE, as well as DE sorted for prospective foregut identity (based on CXCR4 and KIT expression), referred to as anterior definitive endoderm (ADE) (Morrison et al., 2008; Cheng et al., 2012). We found that these markers were sufficient to distinguish not only ESCs from endoderm, as represented by PC1, but also *in vitro*-derived PrE and DE in PC2. When including a panel of genes specific to naïve or primed hESCs (Table S4B) (Messmer et al., 2019), in addition to the previous set of markers from the human blastocyst, we observed a further separation of different ESC types. Primed hESCs more closely identified with DE cell types, as indicated by PC1 (Fig. 6C).

To further define the relationship between the developmental stage in pluripotent culture to that derived upon differentiation, we focused on the capacity of another naïve human state to differentiate into endoderm. It has been suggested that naïve hESCs generated using the RSeT system (Gafni et al., 2013) represent a later stage of EPI development, albeit more naïve than conventional primed hESCs (Stirparo et al., 2018). As the early post-implantation EPI was recently shown to generate VE (Nowotschin et al., 2019), we predicted that RSeT cells would differ from primed cultures in their capacity to generate PrE, but might also generate DE. We therefore adapted H9 hESCs to RSeT conditions (naïve$^{\text{RSeT}}$) (Fig. S7A) and assessed endoderm differentiation. After 7 days in RACL, naïve$^{\text{RSeT}}$ cells differentiated to a population that exhibited co-expression of the PrE marker PDGFRA and the DE marker CXCR4 (Fig. S7B). In addition, these cells downregulated NANOG and OCT4, and homogenously expressed GATA6 (Fig. S7C,D). Based on normalised expression of a panel of markers and flow cytometry, these cells exhibit a combination of primitive and definitive characteristics (Fig. S7B,E).

Although naïve$^{\text{RSeT}}$ cells appeared to generate a compound endodermal population, we considered whether a PrE cell type could be expanded, and how this expansion might impact on endodermal phenotypes. We placed naïve$^{\text{RSeT}}$-derived PrE (PrE$^{\text{RSeT}}$) in NACL and these cultures expanded efficiently as putative nEnd, similar to nEnd$^{\text{t2iLGö}}$ (Fig. S7F). Attempts to expand primed hESC-derived DE in NACL have thus far been unsuccessful, with the differentiated endoderm in these cultures being outcompeted by overgrowth of ESCs (Fig. S7G). To determine the nature of nEnd$^{\text{RSeT}}$, we performed whole-transcriptome analysis of these cultures and combined them with the transcriptomes described previously for unbiased hierarchical clustering of the different naïve, primed and differentiated states (Fig. 6D). Here, it is apparent that nEnd$^{\text{RSeT}}$ clusters with nEnd$^{\text{t2iLGö}}$, suggesting that both naïve states can generate extra-embryonic endoderm upon expansion. This analysis also supports the notion that naïve$^{\text{RSeT}}$ hESCs either generate an intermediate endodermal cell type that initially clusters with DE/ADE, or a mixture of differentiated cell types that average out to this position (Fig. 6D). To distinguish between these possibilities, we differentiated naïve$^{\text{t2iLGö}}$, naïve$^{\text{RSeT}}$ and primed ESCs in the presence of RACL. We assessed each ESC and endodermal cell type for pluripotency, and primitive and definitive endoderm markers using quantitative immunofluorescence (Fig. S7H,I). Endoderm derived from primed cultures contains a large number of brachyury (TBXT)- and FOXA2-positive cells, which are notably reduced in PrE$^{\text{t2iLGö}}$ cultures. In contrast, PrE$^{\text{RSeT}}$ cells contain populations of both positive and negative cells for these DE markers. This suggests that differentiation of naïve$^{\text{RSeT}}$ cultures produces a mix of primitive and definitive endoderm that resolves to PrE when placed in NACL, as only the primitive population expands.

The similarities of these different cell states were also observed by PCA (Fig. 6E). Here, PC1 appears to resolve primed and naïve cell types. This would suggest that the t2iLGö/h2iL and RSeT naïve states constitute different steps in a potential developmental progression, with naïve$^{\text{RSeT}}$ hESCs representing an intermediate in a continuum from a potential pre-implantation t2iLGö/h2iL state to conventional primed pluripotent cells. This observation is also supported by the expression of established markers, with naïve$^{\text{RSeT}}$ expressing high levels of *NANOG* and *DPPA3* (Fig. 6F). Consistent with the single-cell analysis shown in Fig. S7I, which indicates a heterogeneous endoderm population, the whole-transcriptome analysis shows expression of both PrE and DE markers, including *PDGFRA* and *NID2*, as well as *GSC*, *CER1* and *HHEX* (Fig. 6F). Upon expansion in NACL, this heterogeneity is resolved with the PrE/nEnd population apparently selected for, and with the gene expression signature of nEnd$^{\text{RSeT}}$ closely resembling nEnd$^{\text{t2iLGö}}$ (Fig. 6D-F).

To define sets of markers unique to these populations of endoderm, we assessed the top 70 genes varying across all endodermal cultures and defined clusters of genes in which expression correlated (Fig. S8A). From these, we observed genes known to be upregulated during gastrulation, such as *GSC*, *CER1* and *LEFTY1*, clustering to DE, ADE and RSeT-derived endoderm. The potential of *NID2* as a specific marker for PrE is reinforced here, with *WNT2*, *TBX4* and *COL3A1* also appearing as unique markers for *in vitro*-differentiated or expanding nEnd.

As a means to verify the hypoblast nature of our *in vitro*-derived PrE, we compared our bulk transcriptomic data with that derived from single-cell sequencing of cynomolgus monkey embryos (Nakamura et al., 2016, 2017). We performed differential expression analysis of genes that define the transition between naïve pluripotency and PrE/nEnd, as well as naïve ESCs and DE, and equivalent stages of primate development (Fig. 6G, Fig. S8B-E). Based on common genes present in these comparisons, it appears that *in vitro*-derived PrE/nEnd closer resembles primate hypoblast, rather than later post-implantation stages of endoderm development. In addition, *in vitro*-derived DE more strongly correlates with a cluster of post-implantation gastrulating cells, Gast2a/b, than with the hypoblast. We applied the same comparison to published human single-cell datasets (Yan et al., 2013; Blakeley et al., 2015) and similarly found that our PrE/nEnd, but not DE, cultures correlate to the changes in gene expression observed in extra-embryonic hypoblast (Fig. S8F-H). Unfortunately, as it is not possible to obtain material for early human *in vivo* DE and additional extra-embryonic endoderm material, the comparisons that we are able to perform are inherently limited.

DISCUSSION

Using a reporter cell line marking endodermal types, we have demonstrated that the context-dependent response of pluripotent cells to Wnt and Nodal signalling previously demonstrated in mouse is conserved in human, and that these pathways could induce naïve hESCs to generate PrE and support its expansion. Based on this model for PrE differentiation, we provide a partial resolution to a controversy, showing that FGF signalling is essential for PrE specification *in vitro*, suggesting that the canonical role of this pathway defined in mouse is conserved. Moreover, consistent with our observations in mouse that FGF/ERK inhibition in ground state conditions supports self-renewal via a block of PrE differentiation (Hamilton and Brickman, 2014), we found that inhibiting this pathway at the level of the FGFR was sufficient to support naïve pluripotency alongside LIF and CHIR.

The overwhelming similarities at the transcriptional level between cells in t2iLGö and h2iL suggests that signalling between human and other vertebrate species, including mouse, is more conserved than has been previously thought. However, why is inhibition of MEK1/2 alone insufficient to support human naïve pluripotency? In PrE differentiation *in vitro*, we observed that the FGFR inhibitor PD17 was a more effective antagonist of endoderm differentiation than targeted MEK1/2 inhibition. This suggests the presence of a second pathway downstream of the FGFR, required for PrE differentiation, and that this pathway could exploit aPKC. Previous studies have shown that the activities of aPKC are necessary for mitogen signalling via the MAPK/ERK cascade (Berra et al., 1993, 1995). It has also been suggested that aPKC$_{\iota/\lambda}$ is recruited to modulate FGF signalling through the docking protein FRS2 (Lim et al., 1999). In this context, aPKC$_{\iota/\lambda}$ acts downstream of Ras, but upstream of ERK1/2 (Bjørkøy et al., 1997). As aPKC$_{\iota/\lambda}$ has been implicated in PrE specification in mouse development (Saiz et al., 2013), there may be a fundamental aPKC$_{\iota/\lambda}$/FGF-dependent signalling loop involved in PrE induction and segregation from the EPI. In mouse, this loop appears to be MEK1/2-dependent, whereas in human this signalling loop has diverged slightly, so that MEK1/2 and aPKC$_{\iota/\lambda}$ represent parallel branch points. Thus, in human, simultaneous inhibition of aPKC$_{\iota/\lambda}$ and MEK1/2 or a block to the FGFR is required to block differentiation, support self-renewal and sustain naïve pluripotency.

A number of naïve ESC states have been reported and extensively characterised in human. Here, we have exploited the parallel differences seen in the transcriptomes from ESCs and ESC-derived endoderm that we initially described in mouse, to demonstrate that a similar trajectory is observed in human and that multiple naïve states lie at different positions along this continuum. Thus, naïve$^{\mathrm{RSeT}}$ cells lie between naïve$^{\mathrm{t2iLG\ddot{o}}}$ and naïve$^{\mathrm{h2iL}}$ and conventional primed hESCs, suggesting that these cultures could represent a particularly early state of the post-implantation EPI. This intermediate naïve state appears to be unique in its capacity to generate both embryonic and extra-embryonic endoderm, with the PrE component exhibiting properties of nEnd when placed into expansion. Moreover, as the EPI *in vivo* generates VE, and the later EPI generates DE (Nowotschin et al., 2019), naïve$^{\mathrm{RSeT}}$ may represent a model with differentiation competence for both lineages. In addition, unlike in mouse where only low levels of endodermal transcription is detected in naïve ESC culture (Canham et al., 2010), we and others have observed significant GATA6 protein expression in different human cultures (Chan et al., 2013; Guo et al., 2016). Specifically, in both steady state self-renewal and during early endoderm differentiation, GATA6/NANOG double-positive cells arise, suggesting that cR naïve ESC culture may be closer in identity to the unsegregated ICM. Taken together, these observations suggest that, in human, it appears possible to trap a broad range of putative pluripotent states in culture, capturing stages ranging from unsegregated ICM, through implantation, and eventually to later post-implantation, pre-gastrulation stages in primed culture.

Common to the pluripotent states captured in human ESC culture is the capacity to respond to the evolutionarily conserved Wnt and Nodal pathways (reviewed by Grapin-Botton and Constam, 2007). These pathways induce endoderm differentiation in both naïve and primed hESCs, but in a context-dependent manner. In the naïve ground state, these cells are driven towards the developmentally proximal PrE, whereas this signalling in primed cells mediates differentiation towards DE. The conservation of this level of developmental context in mammals serves to provide better coherence to the fundamental role of these pathways in endoderm induction observed across a range of vertebrate species. Thus, both embryonic and extra-embryonic endoderm lineages are produced as a result of the same conserved signalling cascades.

Extra-embryonic stem cell lines, including TS and XEN cells, have been derived in rodent species (Tanaka et al., 1998; Kunath et al., 2005; Wamaitha et al., 2015). Whereas XEN cells resemble the later parietal endoderm, we recently reported that when applied to mESCs, RACL induced the more developmentally plastic naïve PrE. These nEnd cultures did not show expression of markers exclusive to either parietal or visceral endoderm (Anderson et al., 2017), an observation repeated here in the human model. We find that these human hypoblast cell lines possess the same naïve character as mouse nEnd and therefore should be considered as human nEnd, rather than XEN cells. Although these cell lines are derived *in vitro*, and their exclusive extra-embryonic character is difficult to verify functionally, they represent an expandable culture system that approximates human extra-embryonic endoderm. As it was recently demonstrated that VE contributes to the mouse embryonic gut (Kwon et al., 2008; Chan et al., 2019; Nowotschin et al., 2019; Peng et al., 2019; Pijuan-Sala et al., 2019), the distinction between embryonic and extra-embryonic endoderm has become blurred. As such, human nEnd represents a culture system with potential for probing the differentiation competence of the hypoblast for human organ culture. Given the focus on extra-embryonic endoderm as a primordial state for the reprogramming of somatic cells to pluripotency (Parenti et al., 2016; Zhao et al., 2018), the existence of a closely related naïve endodermal culture system could serve as an important stepping stone in understanding mechanisms regulating the final stages of induced pluripotency.

MATERIALS AND METHODS

Cell lines

H9 (WA09, WiCell) and HUES4-170-3 (Ameri et al., 2017) hESCs were used for primed culture and subsequent chemical resetting. The cR-Shef6 cell line was a generous gift from the Smith lab (Takashima et al., 2014) and the embryo-derived HNES1 cell line was a generous gift from the Nichols lab (Guo et al., 2016). All lines were routinely screened for mycoplasma, and all were negative.

Conventional primed hESC culture

Primed hESCs were maintained on tissue culture plates pre-coated with 0.1% gelatine and 25×10^3 per cm^2 irradiated MEF feeder cells in KSR (Thermo Fisher Scientific) medium supplemented with 10 ng/ml FGF2 (R&D Systems). Cells were passaged every 3-5 days with Accutase (00455556, Thermo Fisher Scientific), re-plated with 10 µM ROCKi (Y-27632, Stemcell Technologies) for 24 h and maintained in 20% O$_2$, 5% CO$_2$, 37°C.

Naïve cells culture

Chemically reset naïve human ESC lines were generated and cultured using the protocols as specified from Austin Smith's lab (Guo et al., 2017) or by RSeT™ Medium (Stemcell Technologies), a commercial medium based on NHSM formula (Gafni et al., 2013).

Naïve hESCs were maintained in N2B27 supplemented with either t2iLGö [1 µM CHIR (1386, Axon), 1 µM PD03 (PZ0162, Sigma-Aldrich), an activity equivalent of 10 ng/ml LIF (made in house) and 2 µM Gö (2285, Tocris)] or h2iL [1 µM CHIR, 1 µM PD17 (3044, Tocris) and 10 ng/ml LIF]. Additional aPKC inhibitors tested include CRT0066854 (2656, Axon), PKC412 (2992, Tocris) and ZIP (2549, Tocris) at varying concentrations. Cells were seeded on either plates pre-coated with 0.2% gelatine and 15×10^3 per cm^2 MEFs or on 5 µg/ml LN511 (BioLamina) for feeder-free culture. All naïve cells were passaged by dissociation every 3-5 days with Accutase, re-plated with 10 µM ROCKi for 24 h and maintained under hypoxic conditions (5% O$_2$, 5% CO$_2$, 37°C).

Generation of hESC reporter lines

HHEX-mCherry and *FOXA2-mVenus* reporter cells were generated in parental H9 hESCs (H9-HF) by replacing the endogenous start codon of the *HHEX* gene with an *H2B-mCherry-LoxP-SV40-neoR-LoxP* cassette, and by replacing the endogenous stop codon of the *FOXA2* gene with a *GS-mVenus-LoxP-SV40-neoR-LoxP* cassette, respectively (Fig. S1B,D). We first introduced the *HHEX-mCherry* reporter into

H9 hESC via CRISPR nickase-mediated homologous recombination (with gRNA1: 5′-TTATCCGCGCTCCGCGCTCC-3′ and gRNA2: 5′-CGCGGATAAATGTAGCGCCG-3′). Then, the *FOXA2-mVenus* reporter was introduced to the targeted *HHEX-mCherry* reporter hESC via CRISPR HF nuclease-mediated homologous recombination (with gRNA: 5′-GAAGCCGTCGTCTTCTTAAG-3′). All cell lines were routinely checked for appropriate antibiotic resistance and the presence of the correct reporters. All lines generated were characterised by PCR and Southern blot (Fig. S1C,E).

For karyotyping, fibroblasts and H9-HF hESCs were treated for 45 min with KaryoMAX Colcemid (Life Technologies), harvested in freshly prepared fixative consisting of 25% acetic acid and 75% methanol and sent for G-band karyotyping (Cell Guidance Systems).

The G2M proliferation tracing hESC reporter was made by stably introducing an *H2B-mCherry-F2A-mVenus-hGem(1-110)* cassette through random integration (Fig. S1G). The reporter was confirmed by flow cytometry analysis, as shown by all cells marked by H2B-mCherry, with a subset of proliferating cells expressing mVenus-hGem(1-110) (Fig. S1H).

Differentiation to PrE and DE

To differentiate to PrE, naïve or primed hESCs were plated at 5×10^4 per cm^2 onto feeders in RACL [RPMI 1640 medium with GlutaMAX (61870044, Gibco) with B27 minus insulin (A1895601, Gibco) supplemented with 100 ng/ml ActA, 3 µM CHIR and 10 ng/ml LIF] for 7-8 days (PrE) and 4-5 days (DE). The medium was changed every other day until cells reached confluency, at which point medium was changed every day.

Endoderm expansion

When PrE cells reached confluency, they were passaged by dissociation using Accutase, detached from the plate using a cell scraper and collected in clusters using a stripette (Fig. S6G). These were re-plated on feeders at a ratio of 1:2 to 1:4 in NACL (N2B27 supplemented with 100 ng/ml ActA, 3 µM CHIR and 10 ng/ml LIF) every 4-7 days, supplemented with 10 µM ROCKi for the first 24 h. nEnd cultures were frozen 1:1 in 60% NACL supplemented with ROCKi, 30% foetal bovine serum (FBS; Gibco) and 10% dimethyl sulfoxide (DMSO).

Immunostaining and imaging

Cells were fixed with 4% paraformaldehyde for 10 min at 37°C, blocked and permeabilised with 2% donkey serum, 0.3% Triton X-100 and 0.1% bovine serum albumin in PBS for 1 h at room temperature. Primary antibodies (Table S5) were incubated with 3% FBS in PBS overnight at 4°C and secondary antibodies were added for 1 h at room temperature. Brightfield imaging of cells was performed using a Nikon microscope. Fluorescent imaging was carried out using either a Leica AF6000 fluorescent microscope or a Leica SP8 confocal microscope.

Flow cytometry

Cells were dissociated with Accutase and resuspended as single cells with 3% FBS in PBS. Antibodies (Table S5) were incubated for 45 min to 1 h at 4°C and cells were stained with DAPI to exclude dead cells. Cells were analysed using an LSRFortessa (BD Biosciences) and data analysed with the FACSDiva and FCS Express 6 software (BD Biosciences).

RNA preparation and quantitative RT-PCR

Total RNA was extracted using the RNeasy Mini Kit (Qiagen) and cDNA was synthesised using SuperScript III reverse transcriptase (Invitrogen, 11904018) and Random Hexamer primers (Invitrogen, N8080127). RT-qPCR was performed using the LightCycler 480 Instrument II (Roche Life Science) using the primers and probes listed in Table S6. The housekeeping genes *ACTB* and *GAPDH* were used to normalise expression.

Microarray

We mixed 100 ng of total RNA with RNA standards (Agilent; 5188-5282) and labelled with the LowInput QuickAmp Labeling Kit One-Color (Agilent; 5190-2305) according to the manufacturer's instructions. For each sample, 600 ng of labelled RNA was hybridised to a Human Gene Expression 8×60K v2 (Agilent; G4851B). RNA quantification and assessment of integrity was performed using the total RNA nanochip (Agilent; 5067-1511) measured on an Agilent Bioanalyzer. Microarrays were hybridised for 17 h, washed according to the manufacturer's instructions and scanned on a Surescan G2600D scanner using Agilent Scan Control 9.1.7.1 software with default settings (GE 8×60k microarrays). The resulting single-channel TIFF images were processed using feature extraction software (Agilent). All microarray expression data were analysed using ExAtlas (Sharov et al., 2015).

Transcriptome sequencing

RNA integrity was assessed using a Fragment Analyzer (Advanced Analytical Technologies) and ribosomal RNA was removed from the samples using the NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs). Sequencing libraries were prepared from 250 ng of purified total mRNA using NEBNext Ultra II RNA Library Prep Kit from Illumina (New England Biolabs) according to the manufacturer's instructions. Single-read RNA-seq was performed using the NextSeq 500 (Illumina) with High Output Kit (Illumina, FC-404-2005). Fastq files were aligned to the hg38/GRCh38 genome using STAR v2.5.3a (Dobin et al., 2013). Transcript expression levels were estimated with the –quantMode GeneCounts option and GRCh38p10.v27 annotations. FastQC v0.11.7 was used for QC metrics, and multiqc v1.7 (Ewels et al., 2016) for reporting. Data analysis was then performed with R/Bioconductor (Gentleman et al., 2004). Normalisation was performed with DESeq2 (Love et al., 2014). All RNA-seq data were analysed using the Bioconductor package DESeq2 (Love et al., 2014). Human (GSE36552, GSE66507) and cynomolgus monkey (GSE74767) single-cell RNA-sequencing datasets were analysed using the R package 'Seurat' (Stuart et al., 2019). Differentially expressed genes for each cell stage were calculated using the 'FindMarkers' function against Hypoblast versus EPI (human) or Hypoblast versus ICM/Pre-EPI (monkey) stages. Log2 fold changes of shared differentially expressed genes between single-cell RNA-seq stages and bulk RNA-seq samples were compared using Pearson correlations.

Statistics and reproducibility

Statistical tests were performed using GraphPad Prism 8. Significance was defined as: not significant>0.05; *$P<0.05$, **$P<0.01$, ***$P<0.001$, ****$P<0.0001$. Unless otherwise stated, error bars indicate ±s.d., P values were determined by standard *t*-test and all experiments are representative of three experiments or were reproduced with three biological replicates ($n=3$).

Acknowledgements

We thank Austin Smith and Jennifer Nichols (University of Cambridge) for the cR-Shef6 and HNES1 cell lines; Henrik Semb for the HUES4-170-3 cell line; and Helen Neil, Magali Michaut and the DanStem Genomics Platform for technical expertise, support and the use of instruments. Data processing and analysis were performed using Computerome Supercomputer at DTU (www.computerome.dk). We thank the DanStem Flow Cytometry Platform, the DanStem Imaging Platform and the DanStem Stem Cell Culture Platform for training, technical expertise, support and the use of instruments. We also thank Anne Grapin-Botton, Henrik Semb and members of the Brickman lab for critical comments on this manuscript. The Novo Nordisk Foundation Center for Stem Cell Biology is supported by a Novo Nordisk Foundation grant number NNF17CC0027852.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: M.L.-A., Y.F.W., J.M.B.; Methodology: M.L.-A., Y.F.W., K.G.V.A., J.M.B.; Software: M.L.-A., J.A.R.H., R.S.M.; Validation: M.L.-A., Y.F.W.; Formal analysis: M.L.-A., Y.F.W., J.A.R.H., R.S.M.; Investigation: M.L.-A., Y.F.W.; Data curation: M.L.-A., J.A.R.H.; Writing - original draft: M.L.-A., J.M.B.; Writing - review & editing: M.L.-A., Y.F.W., J.A.R.H., R.S.M., J.M.B.; Visualization: M.L.-A.; Supervision: Y.F.W., J.M.B.; Project administration: J.M.B.; Funding acquisition: R.S.M., J.M.B.

Funding

This work was funded by HumEn under the European Union Seventh Framework Programme FP7/2007-2013 (HEALTH-F4-2013-602889) and by the Danmarks Frie Forskningsfond (DFF-6110-00009). R.S.M. is supported by a fellowship from the Lundbeckfonden (R303-2018-2939).

Data availability

Full microarray expression and Illumina sequencing data have been deposited in GEO under the accession numbers GSE138012 and GSE138013.

Supplementary information

Supplementary information available online at http://dev.biologists.org/lookup/doi/10.1242/dev.180620.supplemental

References

Ameri, J., Borup, R., Prawiro, C., Ramond, C., Schachter, K. A., Scharfmann, R. and Semb, H. (2017). Efficient generation of glucose-responsive beta cells from isolated gp2 + human pancreatic progenitors. *Cell Rep.* **19**, 36-49. doi:10.1016/j.celrep.2017.03.032

Anderson, K. G. V., Hamilton, W. B., Roske, F. V., Azad, A., Knudsen, T. E., Canham, M. A., Forrester, L. M. and Brickman, J. M. (2017). Insulin fine-tunes self-renewal pathways governing naive pluripotency and extra-embryonic endoderm. *Nat. Cell Biol.* **19**, 1164-1177. doi:10.1038/ncb3617

Arman, E., Haffner-Krausz, R., Chen, Y., Heath, J. K. and Lonai, P. (1998). Targeted disruption of fibroblast growth factor (FGF) receptor 2 suggests a role for FGF signaling in pregastrulation mammalian development. *Proc. Natl. Acad. Sci. USA* **95**, 5082-5087. doi:10.1073/pnas.95.9.5082

Artus, J., Panthier, J.-J. and Hadjantonakis, A. K. (2010). A role for PDGF signaling in expansion of the extra-embryonic endoderm lineage of the mouse blastocyst. *Development* **137**, 3361-3372. doi:10.1242/dev.050864

Artus, J., Douvaras, P., Piliszek, A., Isern, J., Baron, M. H. and Hadjantonakis, A. K. (2012). BMP4 signaling directs primitive endoderm-derived XEN cells to an extraembryonic visceral endoderm identity. *Dev. Biol.* **361**, 245-262. doi:10.1016/j.ydbio.2011.10.015

Berra, E., Diaz-Meco, M. T., Dominguez, I., Municio, M. M., Sanz, L., Lozano, J., Chapkin, R. S. and Moscat, J. (1993). Protein kinase C ζ isoform is critical for mitogenic signal transduction. *Cell* **74**, 555-563. doi:10.1016/0092-8674(93)80056-K

Berra, E., Díaz-Meco, M. T., Lozano, J., Frutos, S., Municio, M. M., Sánchez, P., Sanz, L. and Moscat, J. (1995). Evidence for a role of MEK and MAPK during signal transduction by protein kinase C zeta. *EMBO J.* **14**, 6157-6163. doi:10.1002/j.1460-2075.1995.tb00306.x

Bjorkoy, G., Perander, M., Overvatn, A. and Johansen, T. (1997). Reversion of Ras- and phosphatidylcholine-hydrolyzing phospholipase C-mediated transformation of NIH 3T3 cells by a dominant interfering mutant of protein kinase C lambda is accompanied by the loss of constitutive nuclear mitogen-activated protein kinase/extracellular signal-regulated kinase activity. *J. Biol. Chem.* **272**, 11557-11565. doi:10.1074/jbc.272.17.11557

Blakeley, P., Fogarty, N. M. E., del Valle, I., Wamaitha, S. E., Hu, T. X., Elder, K., Snell, P., Christie, L., Robson, P. and Niakan, K. K. (2015). Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. *Development* **142**, 3151-3165. doi:10.1242/dev.123547

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**, dev167833. doi:10.1242/dev.167833

Bos, M., Mendelsohn, J., Kim, Y. M., Albanell, J., Fry, D. W. and Baselga, J. (1997). PD153035, a tyrosine kinase inhibitor, prevents epidermal growth factor receptor activation and inhibits growth of cancer cells in a receptor number-dependent manner. *Clin. Cancer Res.* **3**, 2099-2106.

Brons, I. G. M., Smithers, L. E., Trotter, M. W. B., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S. M., Howlett, S. K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R. A. et al. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**, 191-195. doi:10.1038/nature05950

Burdon, T., Stracey, C., Chambers, I., Nichols, J. and Smith, A. (1999). Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells. *Dev. Biol.* **210**, 30-43. doi:10.1006/dbio.1999.9265

Canham, M. A., Sharov, A. A., Ko, M. S. H. and Brickman, J. M. (2010). Functional heterogeneity of embryonic stem cells revealed through translational amplification of an early endodermal transcript. *PLoS Biol.* **8**, e1000379. doi:10.1371/journal.pbio.1000379

Chambers, I. and Tomlinson, S. R. (2009). The transcriptional foundation of pluripotency. *Development* **136**, 2311-2322. doi:10.1242/dev.024398

Chan, Y.-S., Göke, J., Ng, J.-H., Lu, X., Gonzales, K. A. U., Tan, C.-P., Tng, W.-Q., Hong, Z.-Z., Lim, Y.-S. and Ng, H.-H. (2013). Induction of a human pluripotent state with distinct regulatory circuitry that resembles preimplantation epiblast. *Cell Stem Cell* **13**, 663-675. doi:10.1016/j.stem.2013.11.015

Chan, M. M., Smith, Z. D., Grosswendt, S., Kretzmer, H., Norman, T. M., Adamson, B., Jost, M., Quinn, J. J., Yang, D., Jones, M. G. et al. (2019). Molecular recording of mammalian embryogenesis. *Nature* **570**, 77-82. doi:10.1038/s41586-019-1184-5

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. doi:10.1016/j.devcel.2006.02.020

Cheng, X., Ying, L., Lu, L., Galvão, A. M., Mills, J. A., Lin, H. C., Kotton, D. N., Shen, S. S., Nostro, M. C., Choi, J. K. et al. (2012). Self-renewing endodermal progenitor lines generated from human pluripotent stem cells. *Cell Stem Cell* **10**, 371-384. doi:10.1016/j.stem.2012.02.024

D'Amour, K. A., Agulnick, A. D., Eliazer, S., Kelly, O. G., Kroon, E. and Baetge, E. E. (2005). Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nat. Biotechnol.* **23**, 1534-1541. doi:10.1038/nbt1163

Deshayes, S., Maurizot, V., Clochard, M.-C., Baudin, C., Berthelot, T., Esnouf, S., Lairez, D., Moenner, M. and Déléris, G. (2011). "Click" conjugation of peptide on the surface of polymeric nanoparticles for targeting tumor angiogenesis. *Pharm. Res.* **28**, 1631-1642. doi:10.1007/s11095-011-0398-5

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. doi:10.1093/bioinformatics/bts635

Dietrich, J.-E. and Hiiragi, T. (2007). Stochastic patterning in the mouse pre-implantation embryo. *Development* **134**, 4219-4231. doi:10.1242/dev.003798

Ewels, P., Magnusson, M., Lundin, S. and Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047-3048. doi:10.1093/bioinformatics/btw354

Feng, X., Zhang, J., Smuga-Otto, K., Tian, S., Yu, J., Stewart, R. and Thomson, J. A. (2012). Protein kinase C mediated extraembryonic endoderm differentiation of human embryonic stem cells. *Stem Cells* **30**, 461-470. doi:10.1002/stem.1018

Gafni, O., Weinberger, L., Mansour, A. A. F., Manor, Y. S., Chomsky, E., Ben-Yosef, D., Kalma, Y., Viukov, S., Maza, I., Zviran, A. et al. (2013). Derivation of novel human ground state naive pluripotent stem cells. *Nature* **504**, 282-286. doi:10.1038/nature12745

Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J. et al (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5**, R80.

Gerbe, F., Cox, B., Rossant, J. and Chazaud, C. (2008). Dynamic expression of Lrp2 pathway members reveals progressive epithelial differentiation of primitive endoderm in mouse blastocyst. *Dev. Biol.* **313**, 594-602. doi:10.1016/j.ydbio.2007.10.048

Grapin-Botton, A. and Constam, D. (2007). Evolution of the mechanisms and molecular control of endoderm formation. *Mech. Dev.* **124**, 253-278. doi:10.1016/j.mod.2007.01.001

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. doi:10.1016/j.stemcr.2016.02.005

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. doi:10.1242/dev.146811

Hamilton, W. B. and Brickman, J. M. (2014). Erk signaling suppresses embryonic stem cell self-renewal to specify endoderm. *Cell Rep.* **9**, 2056-2070. doi:10.1016/j.celrep.2014.11.032

Koutsourakis, M., Langevold, A., Patient, R., Beddington, R. and Grosveld, F. (1999). 'The transcription factor GATA6 is essential for early extraembryonic development'. *Development* **126**, 723-732.

Kuijk, E. W., van Tol, L. T. A., Van de Velde, H., Wubbolts, R., Welling, M., Geijsen, N. and Roelen, B. A. J. (2012). The roles of FGF and MAP kinase signaling in the segregation of the epiblast and hypoblast cell lineages in bovine and human embryos. *Development* **139**, 871-882. doi:10.1242/dev.071688

Kunath, T., Arnaud, D., Uy, G. D., Okamoto, I., Chureau, C., Yamanaka, Y., Heard, E., Gardner, R. L., Avner, P. and Rossant, J. (2005). Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts. *Development* **132**, 1649-1661. doi:10.1242/dev.01715

Kwon, G. S., Viotti, M. and Hadjantonakis, A.-K. (2008). The endoderm of the mouse embryo arises by dynamic widespread intercalation of embryonic and extraembryonic lineages. *Dev. Cell* **15**, 509-520. doi:10.1016/j.devcel.2008.07.017

Lawson, K. A., Pedersen, R. A., Van, S. and Geer, D. (1987). Cell fate, morphogenetic movement and population kinetics of embryonic endoderm at the time of germ layer formation in the mouse. *Development* **101**, 627-652.

Lim, Y. P., Low, B. C., Lim, J., Wong, E. S. M. and Guy, G. R. (1999). Association of atypical protein kinase C isotypes with the docker protein FRS2 in fibroblast growth factor signaling. *J. Biol. Chem.* **274**, 19025-19034. doi:10.1074/jbc.274.27.19025

Love, M. I., Huber, W. and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550. doi:10.1186/s13059-014-0550-8

Martin Gonzalez, J., Morgani, S. M., Bone, R. A., Bonderup, K., Abelchian, S., Brakebusch, C. and Brickman, J. M. (2016). Embryonic stem cell culture conditions support distinct states associated with different developmental stages and potency. *Stem Cell Reports* **7**, 177-191. doi:10.1016/j.stemcr.2016.07.009

Mason, I. J., Murphy, D., Münke, M., Francke, U., Elliott, R. W. and Hogan, B. L. (1986). Developmental and transformation-sensitive expression of the Sparc gene on mouse chromosome 11. *EMBO J.* **5**, 1831-1837. doi:10.1002/j.1460-2075.1986.tb04434.x

Messmer, T., von Meyenn, F., Savino, A., Santos, F., Mohammed, H., Lun, L., Marioni, J. C. and Reik, W. (2019). Transcriptional heterogeneity in naive and primed human pluripotent stem cells at single-cell resolution. *Cell Rep.* **26**, 815-824. doi:10.1016/j.celrep.2018.12.099

Morgani, S. M. and Brickman, J. M. (2015). LIF supports primitive endoderm expansion during pre-implantation development. *Development* **142**, 3488-3499. doi:10.1242/dev.125021

Morrison, G. M., Oikonomopoulou, I., Migueles, R. P., Soneji, S., Livigni, A., Enver, T. and Brickman, J. M. (2008). Anterior definitive endoderm from ESCs reveals a role for FGF signaling. *Cell Stem Cell* **3**, 402-415. doi:10.1016/j.stem.2008.07.021

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. doi:10.1038/nature19096

Nakamura, T., Yabuta, Y., Okamoto, I., Sasaki, K., Iwatani, C., Tsuchiya, H. and Saitou, M. (2017). Single-cell transcriptome of early embryos and cultured embryonic stem cells of cynomolgus monkeys. *Sci. Data* **4**, 170067. doi:10.1038/sdata.2017.67

Nakanishi, M., Mitchell, R. R., Benoit, Y. D., Orlando, L., Reid, J. C., Shimada, K., Davidson, K. C., Shapovalova, Z., Collins, T. J., Nagy, A. et al. (2019). Human pluripotency is initiated and preserved by a unique subset of founder cells. *Cell* **177**, 910-924.e22. doi:10.1016/j.cell.2019.03.013

Niakan, K. K., Schrode, N., Cho, L. T. Y. and Hadjantonakis, A.-K. (2013). Derivation of extraembryonic endoderm stem (XEN) cells from mouse embryos and embryonic stem cells. *Nat. Protoc.* **8**, 1028-1041. doi:10.1038/nprot.2013.049

Nichols, J. and Smith, A. (2009). Naive and primed pluripotent states. *Cell Stem Cell* **4**, 487-492. doi:10.1016/j.stem.2009.05.015

Nichols, J. and Smith, A. (2012). Pluripotency in the embryo and in culture. *Cold Spring Harb. Perspect. Biol.* **4**, a008128-a008128. doi:10.1101/cshperspect.a008128

Nowotschin, S., Hadjantonakis, A. K. and Campbell, K. (2019). The endoderm: a divergent cell lineage with many commonalities. *Development* **146**, dev150920. doi:10.1242/dev.150920

Paca, A., Séguin, C. A., Clements, M., Ryczko, M., Rossant, J., Rodriguez, T. A. and Kunath, T. (2012). BMP signaling induces visceral endoderm differentiation of XEN cells and parietal endoderm. *Dev. Biol.* **361**, 90-102. doi:10.1016/j.ydbio.2011.10.013

Parenti, A., Halbisen, M. A., Wang, K., Latham, K. and Ralston, A. (2016). OSKM induce extraembryonic endoderm stem cells in parallel to induced pluripotent stem cells. *Stem Cell Reports* **6**, 447-455. doi:10.1016/j.stemcr.2016.02.003

Peng, G., Suo, S., Cui, G., Yu, F., Wang, R., Chen, J., Chen, S., Liu, Z., Chen, G., Qian, Y. et al. (2019). 'Molecular architecture of lineage allocation and tissue organization in early mouse embryo'. *Nature* **572**, 528-532. doi:10.1038/s41586-019-1469-8

Pera, M. F., Lafita, M. J. B. and Mills, J. (1987). Cultured stem-cells from human testicular teratomas: The nature of human embryonal carcinoma, and its comparison with two types of yolk-sac carcinoma. *Int. J. Cancer* **40**, 334-343. doi:10.1002/ijc.2910400309

Petropoulos, S., Edsgärd, 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. doi:10.1016/j.cell.2016.03.023

Pijuan-Sala, B., Griffiths, J. A., Guibentif, C., Hiscock, T. W., Jawaid, W., Calero-Nieto, F. J., Mulas, C., Ibarra-Soria, X., Tyser, R. C. V., Ho, D. L. L. et al. (2019). A single-cell molecular map of mouse gastrulation and early organogenesis. *Nature* **566**, 490-495. doi:10.1038/s41586-019-0933-9

Roach, S., Schmid, W. and Pera, M. F. (1994). Hepatocytic transcription factor expression in human embryonal carcinoma and yolk sac carcinoma cell lines: expression of HNF-3α in models of early endodermal cell differentiation. *Exp. Cell Res.* **215**, 189-198. doi:10.1006/excr.1994.1331

Roode, M., Blair, K., Snell, P., Elder, K., Marchant, S., Smith, A. and Nichols, J. (2012). Human hypoblast formation is not dependent on FGF signalling. *Dev. Biol.* **361**, 358-363. doi:10.1016/j.ydbio.2011.10.030

Rossant, J. (2018). Genetic control of early cell lineages in the mammalian embryo. *Annu. Rev. Genet.* **52**, 185-201. doi:10.1146/annurev-genet-120116-024544

Saiz, N., Grabarek, J. B., Sabherwal, N., Papalopulu, N. and Plusa, B. (2013). Atypical protein kinase C couples cell sorting with primitive endoderm maturation in the mouse blastocyst. *Development* **140**, 4311-4322. doi:10.1242/dev.093922

Schrode, N., Saiz, N., Di Talia, S. and Hadjantonakis, A.-K. (2014). GATA6 levels modulate primitive endoderm cell fate choice and timing in the mouse blastocyst. *Dev. Cell.* **29**, 454-467. doi:10.1016/j.devcel.2014.04.011

Séguin, C. A., Draper, J. S., Nagy, A. and Rossant, J. (2008). Establishment of endoderm progenitors by SOX transcription factor expression in human embryonic stem cells. *Cell Stem Cell* **3**, 182-195. doi:10.1016/j.stem.2008.06.018

Sharov, A. A., Schlessinger, D. and Ko, M. S. H. (2015). ExAtlas: an interactive online tool for meta-analysis of gene expression data. *J. Bioinform. Comput. Biol.* **13**, 1550019. doi:10.1142/S0219720015500195

Smith, A. G. (2001). Embryo-derived stem cells: of mice and men. *Annu. Rev. Cell Dev. Biol.* **17**, 435-462. doi:10.1146/annurev.cellbio.17.1.435

Smith, A. G., Heath, J. K., Donaldson, D. D., Wong, G. G., Moreau, J., Stahl, M. and Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature* **336**, 688-690. doi:10.1038/336688a0

Smyth, N., Vatansever, H. S., Murray, P., Meyer, M., Frie, C., Paulsson, M. and Edgar, D. (1999). Absence of basement membranes after targeting the *LAMC1* gene results in embryonic lethality due to failure of endoderm differentiation. *J. Cell Biol.* **144**, 151-160. doi:10.1083/jcb.144.1.151

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**, dev169672. doi:10.1242/dev.169672

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., Hao, Y., Stoeckius, M., Smibert, P. and Satija, R. (2019). Comprehensive integration of single-cell data. *Cell* **177**, 1888-1902. doi:10.1016/j.cell.2019.05.031

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* **162**, 452-453. doi:10.1016/j.cell.2015.06.052

Tanaka, S., Kunath, T., Hadjantonakis, A. K., Nagy, A. and Rossant, J. (1998). Promotion of trophoblast stem cell proliferation by FGF4. *Science* **282**, 2072-2075. doi:10.1126/science.282.5396.2072

Tesar, P. J., Chenoweth, J. G., Brook, F. A., Davies, T. J., Evans, E. P., Mack, D. L., Gardner, R. L. and McKay, R. D. G. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448**, 196-199. doi:10.1038/nature05972

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. doi:10.1016/j.stem.2014.07.002

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. doi:10.1126/science.282.5391.1145

Wamaitha, S. E., del Valle, I., Cho, L. T. Y., Wei, Y., Fogarty, N. M. E., Blakeley, P., Sherwood, R. I., Ji, H. and Niakan, K. K. (2015). Gata6 potently initiates reprograming of pluripotent and differentiated cells to extraembryonic endoderm stem cells. *Genes Dev.* **29**, 1239-1255. doi:10.1101/gad.257071.114

Ware, C. B., Nelson, A. M., Mecham, B., Hesson, J., Zhou, W., Jonlin, E. C., Jimenez-Caliani, A. J., Deng, X., Cavanaugh, C., Cook, S. et al. (2014). Derivation of naive human embryonic stem cells. *Proc. Natl. Acad. Sci. USA* **111**, 4484-4489. doi:10.1073/pnas.1319738111

Yamanaka, Y., Lanner, F. and Rossant, J. (2010). FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. *Development* **137**, 715-724. doi:10.1242/dev.043471

Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, R., Liu, P., Lian, Y., Zheng, X., Yan, J. et al. (2013). Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat. Struct. Mol. Biol.* **20**, 1131-1139. doi:10.1038/nsmb.2660

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. doi:10.1038/nature06968

Zhao, T., Fu, Y., Zhu, J., Liu, Y., Zhang, Q., Yi, Z., Chen, S., Jiao, Z., Xu, X., Xu, J. et al. (2018). Single-cell RNA-Seq reveals dynamic early embryonic-like programs during chemical reprogramming. *Cell Stem Cell* **23**, 31-45.e7. doi:10.1016/j.stem.2018.05.025

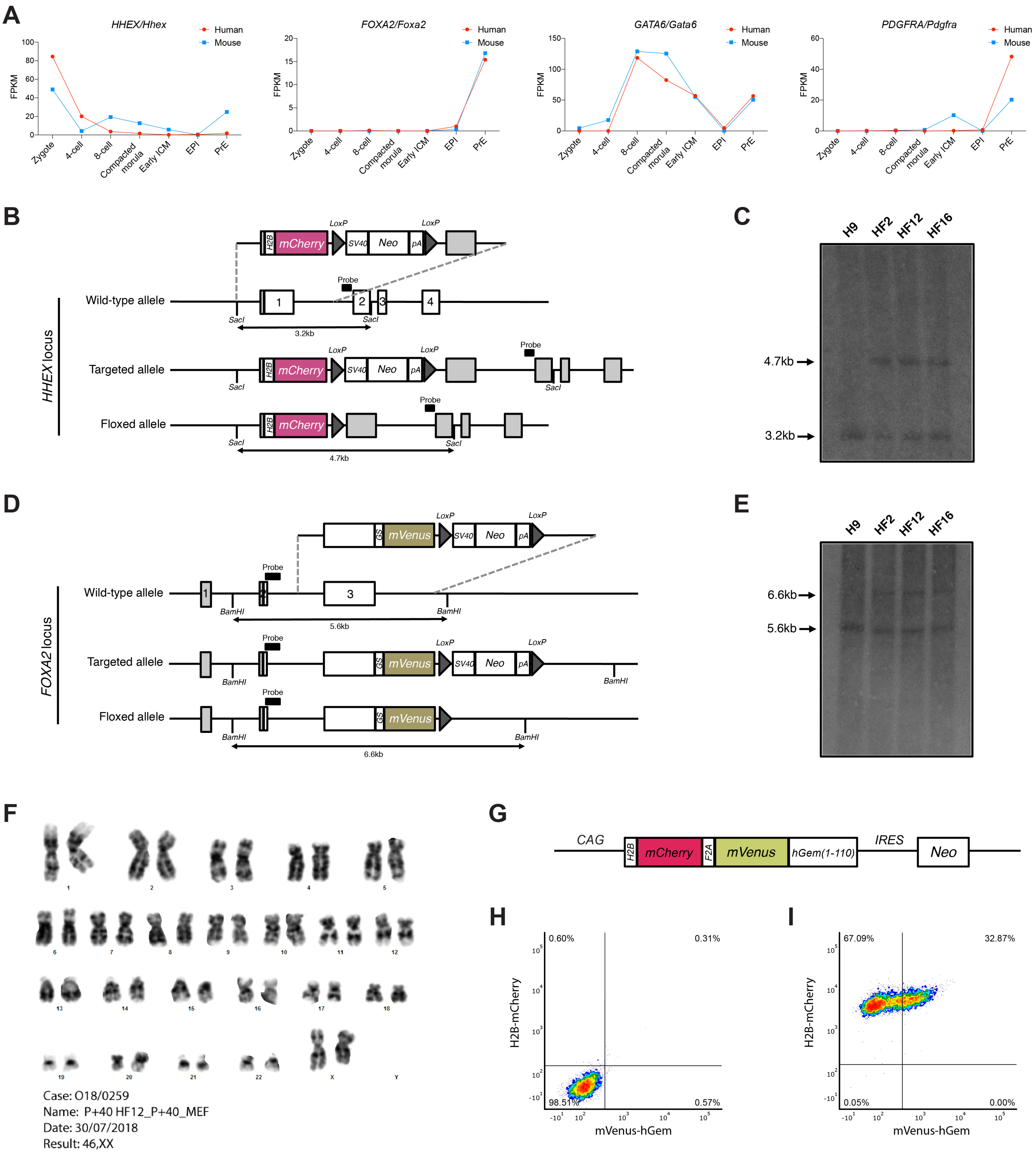

A
HHEX/Hhex
FOXA2/Foxa2
GATA6/Gata6
PDGFRA/Pdgfra
Human
Mouse
FPKM
Zygote
4-cell
8-cell
Compacted morula
Early ICM
EPI
PrE
B
HHEX locus
Wild-type allele
Targeted allele
Floxed allele
H2B
mCherry
SV40
Neo
pA
LoxP
Probe
SacI
3.2kb
4.7kb
C
H9
HF2
HF12
HF16
4.7kb
3.2kb
D
FOXA2 locus
GS
mVenus
BamHI
5.6kb
6.6kb
E
6.6kb
5.6kb
F
Case: O18/0259
Name: P+40 HF12_P+40_MEF
Date: 30/07/2018
Result: 46,XX
G
CAG
F2A
hGem(1-110)
IRES
H
I
H2B-mCherry
mVenus-hGem
0.60%
0.31%
98.51%
0.57%
67.09%
32.87%
0.05%
0.00%

Figure S1: Description of reporter cell lines. (A) Expression of endodermal markers in the human blastocyst across early developmental stages from single-cell sequencing data **(Stirparo et al., 2018)**. (B, D) Schematics depicting the *HHEX-mCherry* and *FOXA2*-mVenus (H9-HF) reporter generated by CRISPR/Cas9-mediated homologous recombination. (C, E) Southern blot of the H9-HF reporter cell line showing the wild type gene sizes and the labelled targeted fragments. (F) HF12 was karyotyped, where the chromosome count was 46. (G) Schematic depicting the *H2B-mCherry-F2A-mVenus-hGem(1-110)* (H9-G2M) reporter driven by the *CAG* promoter generated by random integration. (H) Representative flow cytometry density plot showing gating strategy for wild type H9 hESCs for H2B-mCherry and mVenus-hGem expression. (I) Flow cytometry analysis confirming the H9-G2M reporter in primed media (KSR/FGF), as shown by all cells marked by H2B-mCherry, with a subset of proliferating cells expressing mVenus-hGem.

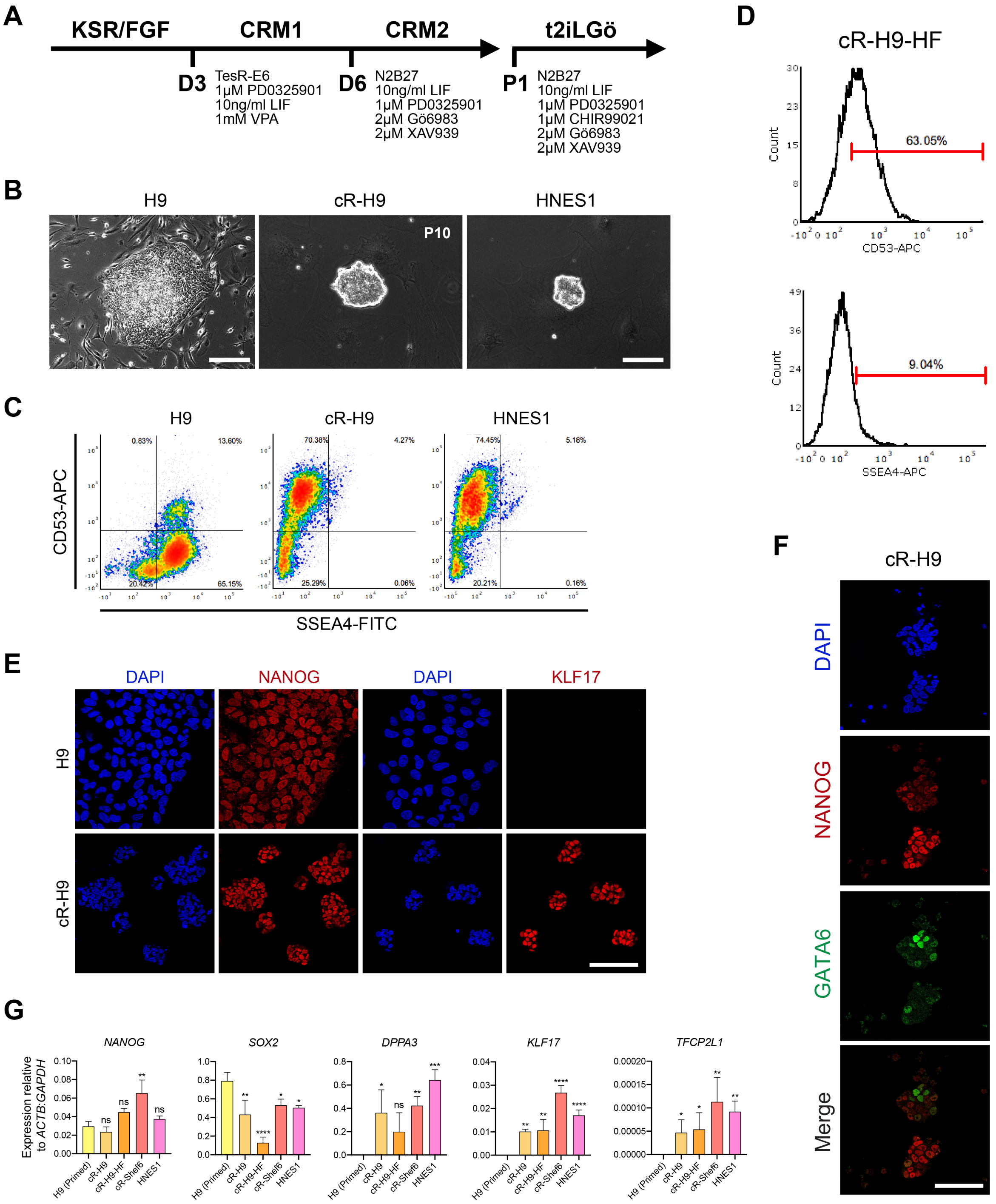

A
KSR/FGF
CRM1
CRM2
t2iLGö
D3
TesR-E6
1µM PD0325901
10ng/ml LIF
1mM VPA
D6
N2B27
10ng/ml LIF
1µM PD0325901
2µM Gö6983
2µM XAV939
P1
N2B27
10ng/ml LIF
1µM PD0325901
1µM CHIR99021
2µM Gö6983
2µM XAV939
B
H9
cR-H9
P10
HNES1
C
H9
cR-H9
HNES1
CD53-APC
SSEA4-FITC
0.83%
13.60%
20.42%
65.15%
70.38%
4.27%
25.29%
0.06%
74.45%
5.18%
20.21%
0.16%
D
cR-H9-HF
Count
63.05%
CD53-APC
9.04%
SSEA4-APC
E
DAPI
NANOG
DAPI
KLF17
H9
cR-H9
F
cR-H9
DAPI
NANOG
GATA6
Merge
G
Expression relative to ACTB:GAPDH
NANOG
SOX2
DPPA3
KLF17
TFCP2L1
H9 (Primed)
cR-H9
cR-H9-HF
cR-Shef6
HNES1
ns
**
*
***
****

Figure S2: Chemical resetting of primed hESCs to the naïve state. (A) Schematic of the chemical resetting protocol including media compositions. (B) Brightfield images of primed H9 compared to cR-H9 and the embryo-derived HNES1. (C) Flow cytometry analysis of primed H9, cR-H9 and HNES1 with staining for primed-specific SSEA4 and naïve-specific CD53. Lower left quadrant indicates gating based on the negative control. (D) As cR-H9-HF could not be stained for SSEA4-FITC, a SSEA4-APC antibody was used instead. Histograms show fluorescence distribution of SSEA4-APC and CD53-APC expression in cR-H9-HF. cR-H9-G2M could not be analysed by flow cytometry due to problems with compensation. Horizontal bar indicates gating against the negative control with no antibody. (E) Immunostaining of primed H9 and cR-H9 for the pluripotency-associated marker NANOG and the naïve-specific marker KLF17, including DAPI, and imaged by confocal microscopy. (F) Immunostaining of cR-H9 for the indicated markers, including DAPI, and imaged by confocal microscopy. (G) qRT-PCR for pluripotency and naïve marker expression across various cell lines relative to *ACTB* and *GAPDH*. Scale bars: 100μm and 50μm in B (left to right); 25μm in E, F.

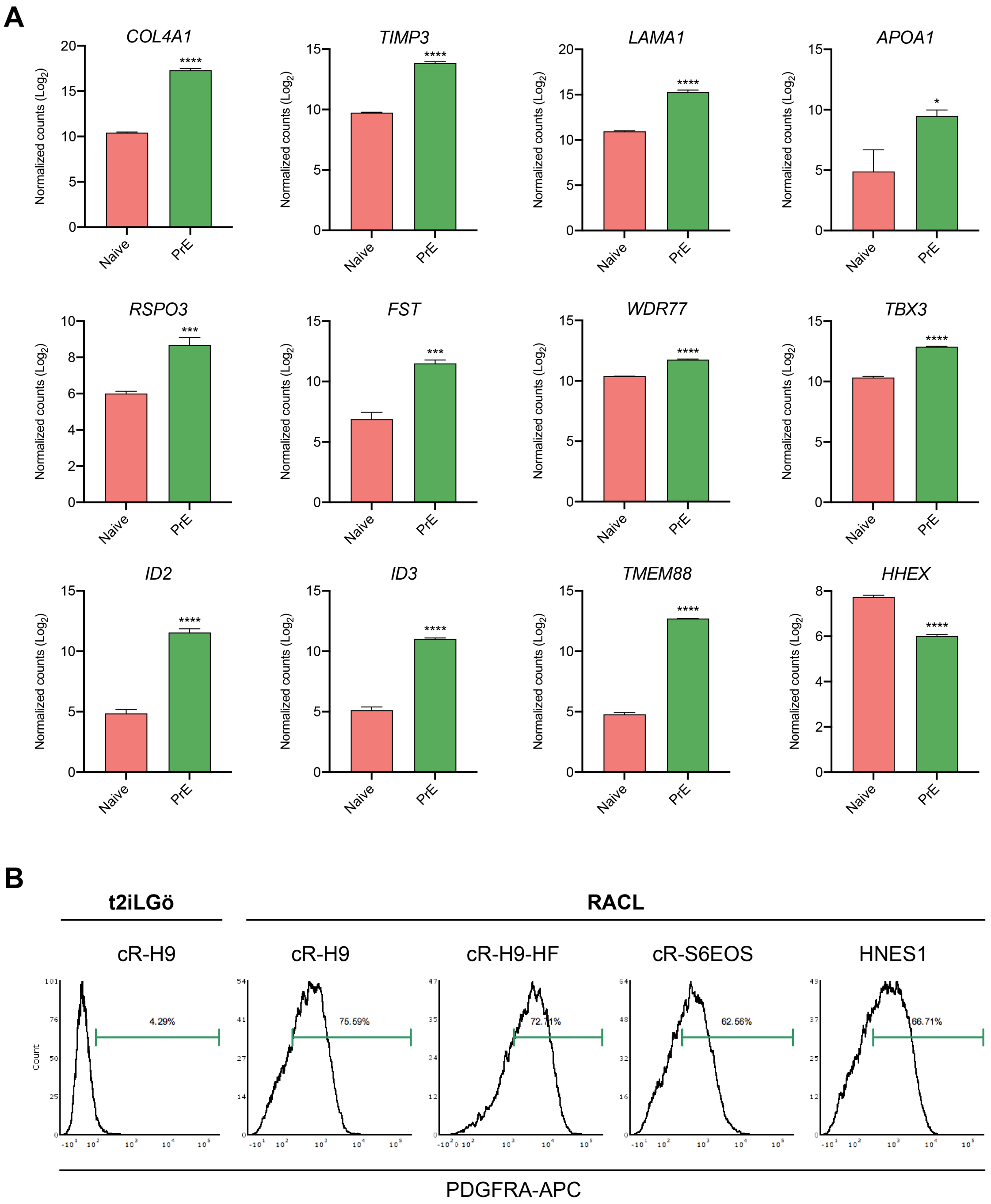


Figure S3: Human blastocyst endoderm marker expression in *in* vitro cell types. (A) Expression of indicated PrE markers **(Yan et al., 2013; Blakeley et al., 2015)** in RNA-sequencing data from naïve hESCs in t2iLGö and PrE. (B) Flow cytometry histograms for PDGFRA expression after PrE differentiation across different cell lines. Horizontal bar indicates gating based on the negative control with no antibody. Representative of 5 differentiations.

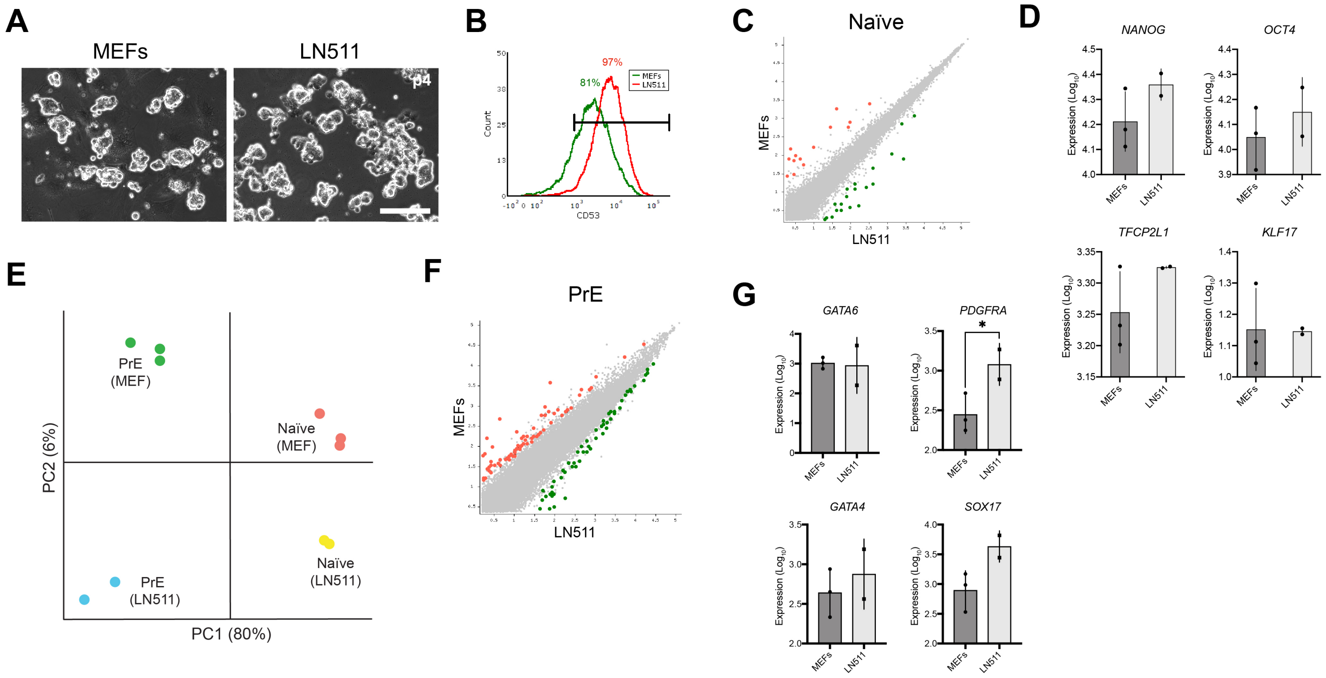


Figure S4: LN511 supports naïve hESCs and PrE differentiation in a feeder-free system. (A) Brightfield images of cR-H9 cultured on MEFs and LN511 for 4 passages in t2iLGö. (B) Histograms representing fluorescence for CD53 expression in cR-H9 maintained on either MEFs or LN511. Horizontal bar indicates gating against the negative control with no antibody. (C, F) Pairwise comparison of gene expression of microarray data for (C) naïve hESCs cultured on LN511 vs MEFs and (F) PrE differentiated on LN511 vs MEFs (fold change (FC) ≤ 1.5, FDR threshold ≤ 0.05). (D, G) Total expression of (D) pluripotency markers in naïve hESCs maintained on MEFs or LN511, and (G) PrE-specific markers in PrE differentiated on MEFs or LN511 based on microarray datasets. *n*=3 biological replicates for naïve cR-H9 on MEFs, *n*=2 from cR-H9 and *n*=1 from HNES1 differentiated to PrE on MEFs; *n*=2 biological replicates for cR-H9 cultured and differentiated on LN511 (E) PCA of microarray datasets from cR-H9 maintained on either LN511 or MEFs and differentiated to PrE on each substrate, as well as HNES1 differentiated to PrE on MEFs. *n* values as above. Scale bars: 100µm in A.

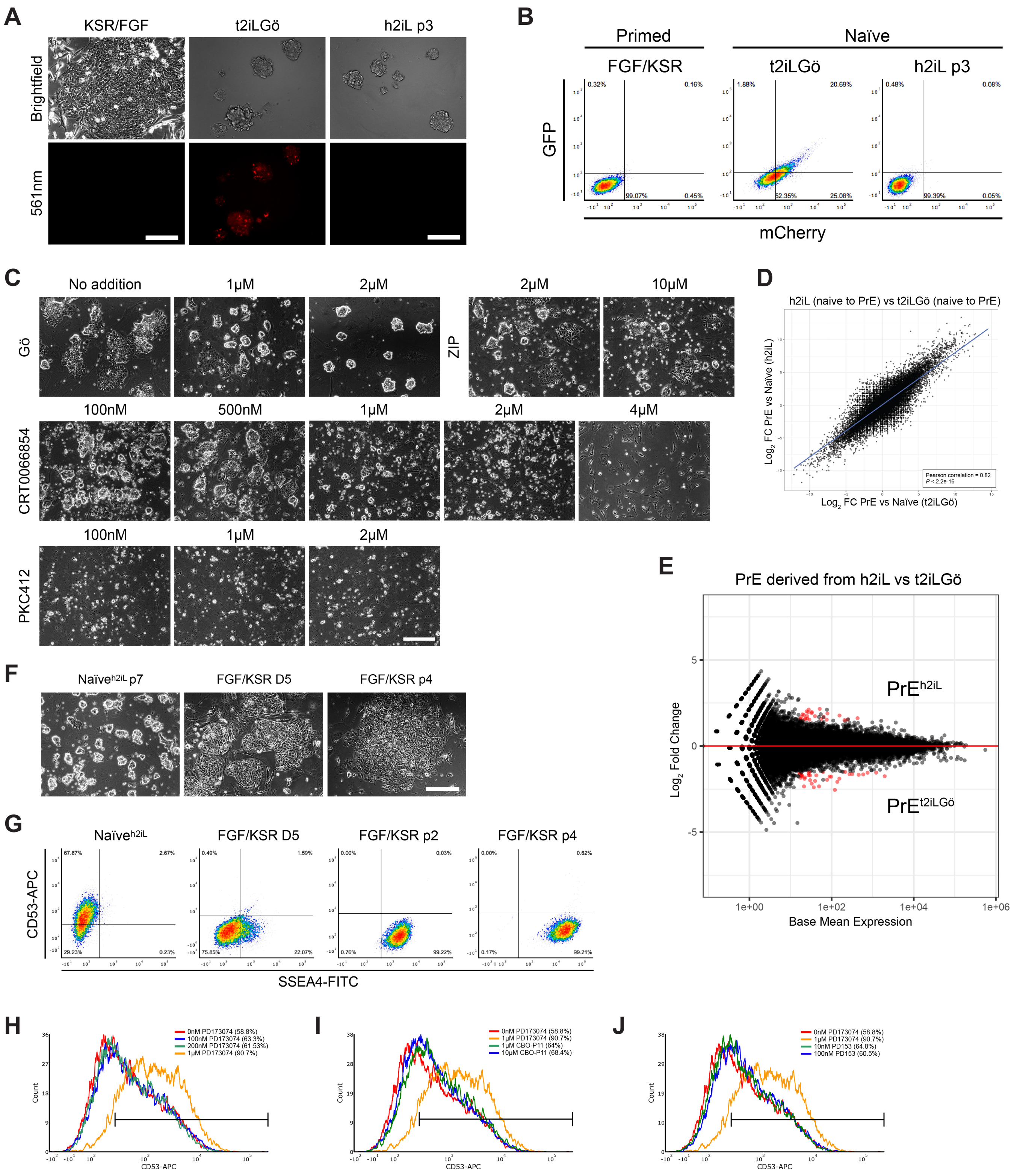
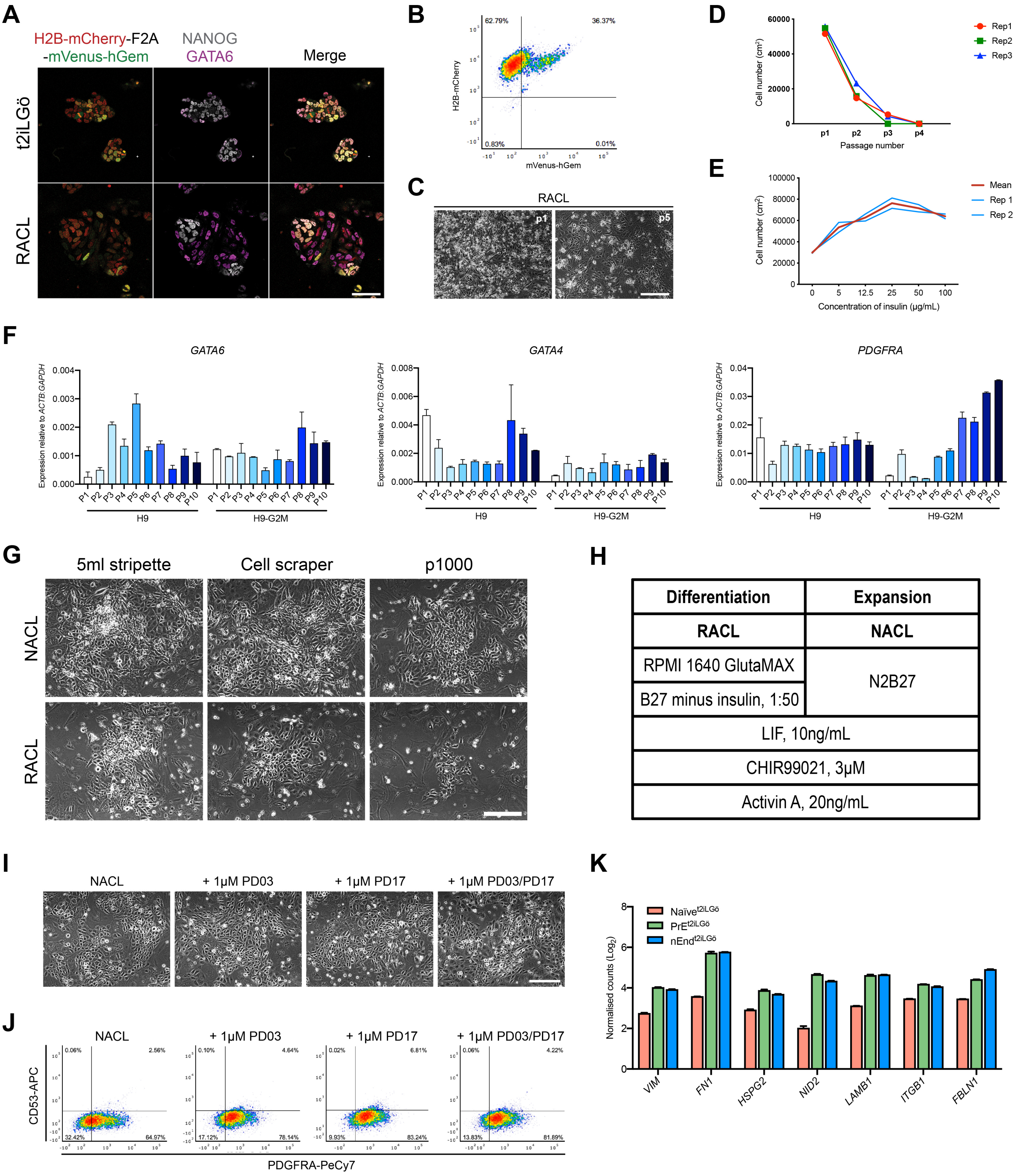

A
KSR/FGF
t2iLGö
h2iL p3
Brightfield
561nm
B
Primed
Naïve
FGF/KSR
t2iLGö
h2iL p3
GFP
mCherry
0.32%
0.16%
99.07%
0.45%
1.88%
20.69%
52.35%
25.08%
0.48%
0.08%
99.39%
0.05%
C
No addition
1μM
2μM
2μM
10μM
Gö
ZIP
100nM
500nM
1μM
2μM
4μM
CRT0066854
100nM
1μM
2μM
PKC412
D
h2iL (naive to PrE) vs t2iLGö (naive to PrE)
Log2 FC PrE vs Naive (h2iL)
Log2 FC PrE vs Naïve (t2iLGö)
Pearson correlation = 0.82
P < 2.2e-16
E
PrE derived from h2iL vs t2iLGö
Log2 Fold Change
Base Mean Expression
PrEh2iL
PrEt2iLGö
5
0
-5
1e+00
1e+02
1e+04
1e+06
F
Naïveh2iL p7
FGF/KSR D5
FGF/KSR p4
G
Naïveh2iL
FGF/KSR D5
FGF/KSR p2
FGF/KSR p4
CD53-APC
SSEA4-FITC
67.87%
2.67%
29.23%
0.23%
0.49%
1.59%
75.85%
22.07%
0.00%
0.03%
0.76%
99.22%
0.00%
0.62%
0.17%
99.21%
H
0nM PD173074 (58.8%)
100nM PD173074 (63.3%)
200nM PD173074 (61.53%)
1μM PD173074 (90.7%)
Count
CD53-APC
I
0nM PD173074 (58.8%)
1μM PD173074 (90.7%)
1μM CBO-P11 (64%)
10μM CBO-P11 (68.4%)
Count
CD53-APC
J
0nM PD173074 (58.8%)
1μM PD173074 (90.7%)
10nM PD153 (64.8%)
100nM PD153 (60.5%)
Count
CD53-APC

Figure S5: Assessment of signalling inhibitors in naïve pluripotent and differentiating culture. (A) Live primed H9 in KSR/FGF, cR-H9 in t2iLGö and cR-H9 in h2iL imaged by fluorescent microscopy. Bright red aggregates can be seen in cR-H9 cultured in t2iLGö. (B) Representative density plots of primed H9 ESCs in KSR/FGF, cR-H9 in t2iLGö and h2iL, showing autofluorescence produced by Gö in the t2iLGö condition. Lower left quadrant indicates gating based on the negative control. (C) Brightfield images of cR-H9 cultured in t2iL supplemented with different aPKC inhibitors at the indicated concentrations. (D) Pearson correlation dot plot showing $\log_2$ FC in global gene expression between naïve and PrE samples in t2iLGö and h2iL. (E) MA-plot representing differential expression analysis of PrE differentiated from naïve hESCs maintained in h2iL vs t2iLGö (1.5 $\log_2$ FC, $p<0.05$), where red dots indicate genes that are differentially expressed ($n=25$ genes in PrE^{h2iL} and $n=35$ genes in $\mathrm{PrE}^{t2iLGö}$). (F-G) Conversion of $\mathrm{naïve}^{h2iL}$ cells back to primed pluripotency. cR-H9 were cultured in h2iL for 7 passages and then placed back into primed culture (FGF/KSR) for 4 passages. (F) Brightfield images of cR-H9 in initial naïve h2iL culture and 4 passages after transfer to primed medium. (G) Flow cytometry density plots of the naïve-to-primed transition for CD53 and SSEA4 expression. Quadrants based on gating against the negative control with no antibody. (H-I) Flow cytometry histograms for naïve hESCs cultured in N2B27 supplemented with 10ng/mL LIF, 1µM CHIR, as well as various concentrations of (H) PD17, (I) the VEGFR inhibitor CBO-P11, and (J) the EGFR inhibitor PD153053. Horizontal bar indicates gating against the negative control with no antibody. Scale bars: 100µm and 50µm in A (left to right); 50µm in C; 100µm in F.



Differentiation	Expansion
RACL	NACL
RPMI 1640 GlutaMAX	N2B27
B27 minus insulin, 1:50	
LIF, 10ng/mL	
CHIR99021, 3µM	
Activin A, 20ng/mL	

Figure S6: Expansion of naïve extra-embryonic endoderm is supported by insulin. (A) cR-H9-G2M naïve^{h2iL} hESCs and subsequent differentiation to PrE imaged by confocal microscopy. H2B-mCherry expression marks all cells, while mVenus-hGem specifies cells in G2/M. (B) Dot plot based on flow cytometry analysis of H2B-mCherry and mVenus-hGem expression in H9-G2M passaged nEnd. (C) Brightfield images of naïve cR-H9 differentiated to PrE in RACL and passaged in RACL medium. (D) Cell numbers of cR-H9 differentiated and passaged in RACL, where cells were plated at a 1:1 ratio at each passage. (E) Cell numbers of cR-H9 PrE cultured in titrated concentrations of insulin (*n*=2 biological replicates, with the mean shown in red). (F) qRT-PCR showing relative expression of PrE markers during expansion of cR-H9 and cR-H9-G2M nEnd in NACL. Error bars indicate ± s.d. of technical replicates (*n*=1 biological replicate for each cell line shown). (G) Brightfield images of cR-H9 PrE expansion in RACL or NACL, using three different passaging techniques. Representative of 3 different cell lines. (H) Overview of culture media composition in PrE differentiation and expansion. (I-J) nEnd expansion does not require FGF/ERK signalling. HNES1 nEnd^{t2iLGö} (passage 8) was passaged to NACL supplemented with 1μM PD032501, 1μM PD173074 or both for 5 days. (I) Brightfield microscopy showing nEnd morphology, and (J) expression of cell surface markers PDGFRA and CD53 in all three conditions by flow cytometry analysis. (K) Comparative expression of BM components in naïve^{t2iLGö} cells compared to PrE and nEnd derived from these, based on RNA sequencing data. Scale bars: 25μm in A; 100μm in C, G, and I.

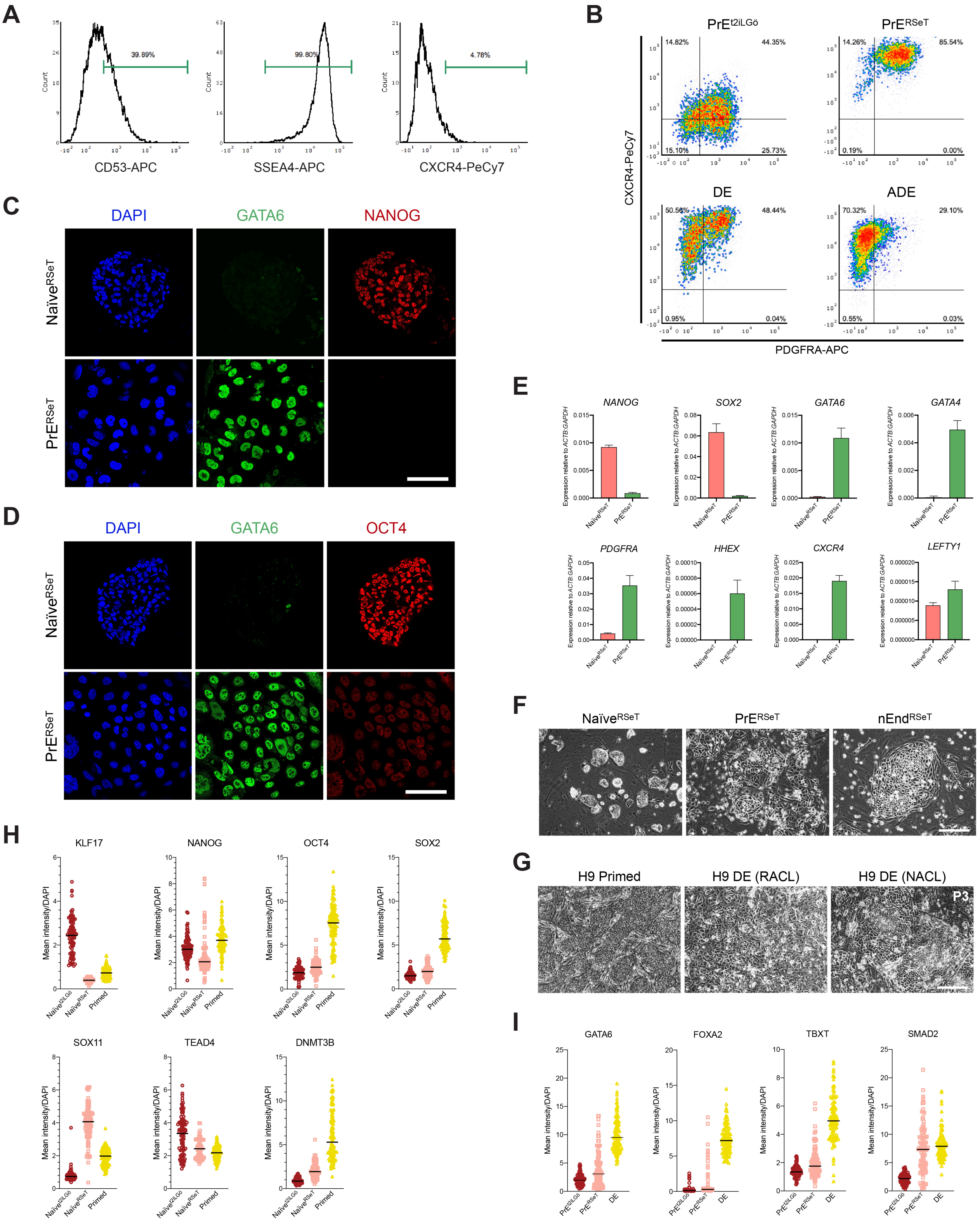

A
39.89%
99.80%
4.78%
Count
CD53-APC
SSEA4-APC
CXCR4-PeCy7
B
PrE^{t2iLGö}
14.82%
44.35%
15.10%
25.73%
PrE^{RSeT}
14.26%
85.54%
0.19%
0.00%
DE
50.56%
48.44%
0.95%
0.04%
ADE
70.32%
29.10%
0.55%
0.03%
CXCR4-PeCy7
PDGFRA-APC
C
DAPI
GATA6
NANOG
Naïve^{RSeT}
PrE^{RSeT}
D
DAPI
GATA6
OCT4
Naïve^{RSeT}
PrE^{RSeT}
E
NANOG
SOX2
GATA6
GATA4
PDGFRA
HHEX
CXCR4
LEFTY1
Expression relative to ACTB:GAPDH
Naïve^{RSeT}
PrE^{RSeT}
F
Naïve^{RSeT}
PrE^{RSeT}
nEnd^{RSeT}
G
H9 Primed
H9 DE (RACL)
H9 DE (NACL)
P3
H
KLF17
NANOG
OCT4
SOX2
SOX11
TEAD4
DNMT3B
Mean intensity/DAPI
Naïve^{t2iLGö}
Naïve^{RSeT}
Primed
I
GATA6
FOXA2
TBXT
SMAD2
Mean intensity/DAPI
PrE^{t2iLGö}
PrE^{RSeT}
DE

Figure S7: Derivation and differentiation of endoderm from RSeT naïve ESCs. (A) Histograms depicting expression of indicated surface markers in ESCs maintained in RSeT medium. Horizontal bar indicates gating against the negative control with no antibody. (B) Density plot for flow cytometry analysis of naïve^{RSeT}-derived PrE for CXCR4 and PDGFRA compared to PrE from naïve^{t2iLGö} cells, as well as DE and ADE from primed hESCs. (C-D) Immunostaining of naïve^{RSeT} and PrE derived from these for (D) NANOG and GATA6 and (E) OCT4 and GATA6, with inclusion of DAPI. (E) qRT-PCR showing relative expression of indicated pluripotency and endoderm markers in naïve^{RSeT} hESCs and subsequent endoderm. (F) Brightfield images of naïve^{RSeT} ESCs differentiated to endoderm in RACL for 7 days and expansion in NACL. (G) Brightfield images of primed H9 hESCs differentiated to DE in RACL for 5 days and attempt at expansion in NACL. (H-I) Single-cell analysis of (H) ESCs (naïve^{t2iLGö}, naïve^{RSeT}, primed) and (I) endoderm differentiation from these ESCs (PrE^{t2iLGö}, PrE^{RSeT}, DE) by quantitative immunofluorescence for indicated markers normalised to DAPI nuclear stain (*n*=100 cells). Scale bars: 50µm for C, D; 100µm for F, G.

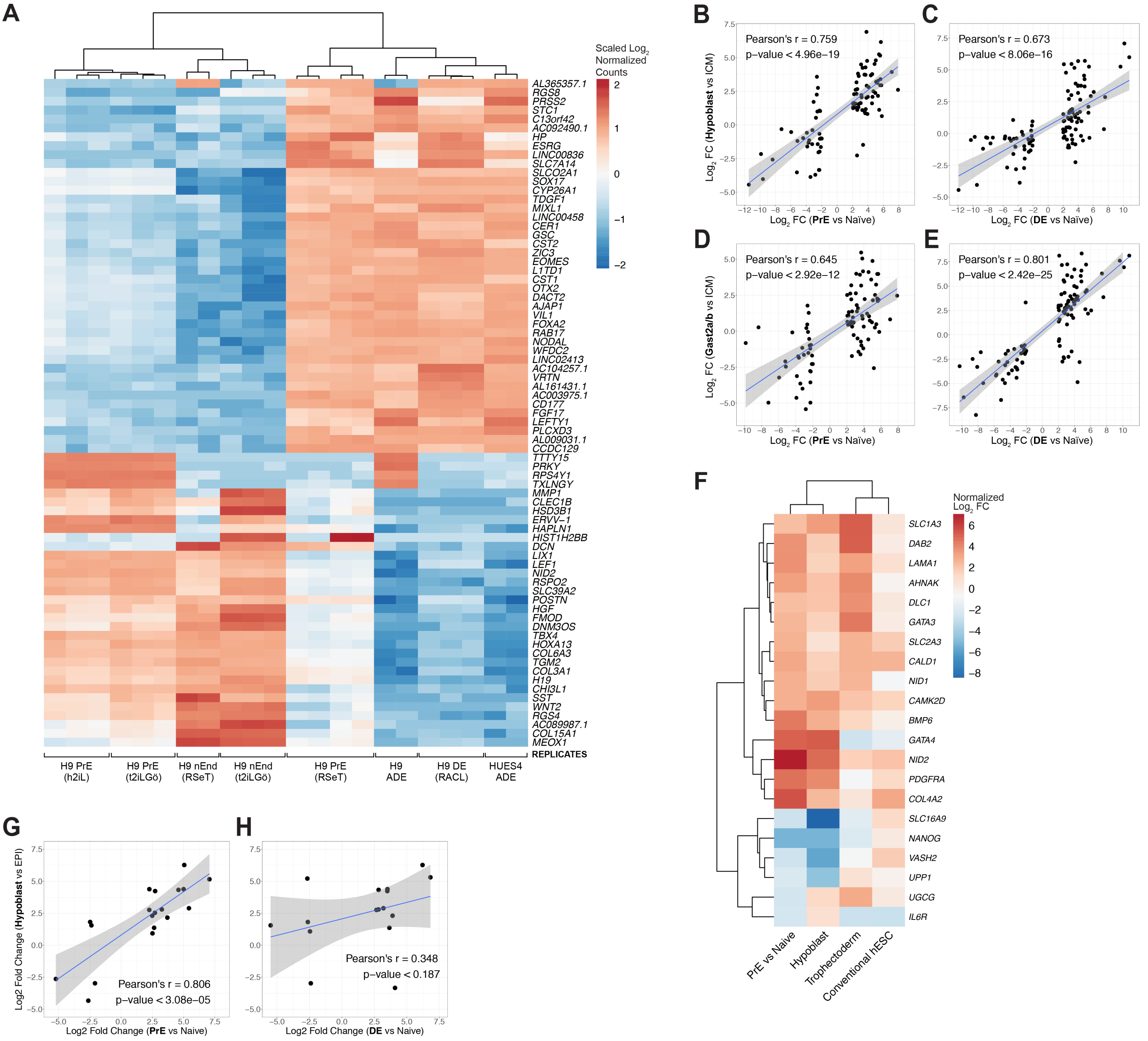


Figure S8: Analysis of *in vitro* transcriptomes and comparison to single-cell *in vivo* datasets derived from human and cynomolgus monkey.

(A) Representation of top 70 most variable genes across all endodermal cell types based on whole-transcriptome analysis of *invitro* cultures. (B-E) Pearson correlation of $\log_2$ FC of common genes between cynomolgus monkey **(Nakamura et al., 2016)** *invivo* and human *invitro* cell types for (B) hypoblast vs ICM and PrE/nEnd vs naïve ESCs (*n*=92 genes), (C) hypoblast vs ICM and DE vs naïve ESCs (*n*=110 genes), (D) Gast2a/b vs ICM and PrE/nEnd vs naïve ESCs (*n*=93 genes), and (E) Gast2a/b vs ICM and DE vs naïve ESCs (*n*=108 genes). (F) Comparative analysis of $\log_2$ FC of common genes (*n*=19) between human early cell types vs EPI **(Yan et al., 2013; Blakeley et al., 2015)** and human *invitro* PrE/nEnd vs naïve cells. (G-H) Pearson correlation of $\log_2$ FC of common genes between human early cell types vs EPI **(Yan et al., 2013; Blakeley et al., 2015)** and human *invitro*; (G) PrE/nEnd vs naïve ESCs (*n*=19 genes), and (H) DE vs naïve ESCs (*n*=16 genes).

Table S1. A: Differentially upregulated genes in PrE compared to DE; B: Differentially downregulated genes in PrE compared to DE; C: Uniquely differentially upregulated genes in PrE compared to DE; D: Uniquely differentially downregulated genes in PrE compared to DE

Click here to Download Table S1

Table S2. A: Overexpressed genes in naive hESCs cultured on LN511 vs MEFs; B: Underexpressed genes in naive hESCs cultured on LN511 vs MEFs; C: Overexpressed genes in PrE differentiated on LN511 vs MEFs; D: Underexpressed genes in PrE differentiated on LN511 vs MEFs

Click here to Download Table S2

Table S3. A: Overexpressed genes in naive hESCs cultured in h2iL vs t2iLGö; B: Underexpressed genes in naive hESCs cultured in h2iL vs t2iLGö; C: Overexpressed genes in PrE differentiated from h2iL vs t2iLGö; D: Underexpressed genes in PrE differentiated from h2iL vs t2iLGö

Click here to Download Table S3

Table S4. A: EPI and PrE markers (Stirparo *et al*., 2018); B: Naive and primed markers (Messner *et al*., 2019)

Click here to Download Table S4

Table S5: Antibodies

Antibody	Vendor	Catalogue number	Concentration
CD53-APC	Miltenyi Biotec	130101795	1:50
CXCR4-PeCy7	BD Bioscience	560669	1:100
DAPI	Thermo Fisher	D1306	1:5000
PDGFRa-APC	BD Bioscience	562798	1:100
SSEA4-647	Molecular Probes	SSEA421	1:1000
SSEA4-FITC	BioLegend	330409	1:100
GATA4	Santa Cruz	sc1237	1:1000
GATA6	R&D	AF1700	1:500
KLF17	Sigma-Aldrich	hpa024629	1:1000
NANOG	Abcam	ab21624	1:1000

Table S6: qRT-PCR primers and probe pairs

Gene	Forward Primer	Reverse Primer	Probe
ACTB	ccaaccgccgagaagatga	ccagcggcgtacagggatag	64
BMP6	tttccaagacctgggatgg	gcattcatgtgtgcgttga	12
CXCR4	cctgcctggtattgtcatcc	gatggggatgattgtggtct	49
DPPA3	gggaaatcgaagatgagtgg	aggctccttgtttgttggtc	18
GAPDH	agccacatcgctcagacac	gcccaatacgaccaaatcc	60
GATA4	ggaagcccaagaacctgaat	gttgctggagttgctggaa	17
GATA6	gcgggctctacagcaagat	tggcacaggacaatccaag	30
GSC	cctccgcgaggagaaagt	cgttctccgactcctctgat	29
HHEX	gcggacggtgaacgacta	ggccgcctttcctttttat	50
HNF4A	acaatcgtcaagcccctct	ccagcggcttgctagataac	9
KLF17	ctcctgctgctggtccttag	cagttgccacgtccagtg	64
LEFTY1	aaagaggttcagccagagctt	caccagcaggtgtgtgct	72
MIXL1	ggtaccccgacatccactt	gcctgttctggaaccatacct	32
NANOG	gggaaaaagccagaagtcg	ctttggggacaagctgga	52
NID2	cctgcagctacctgctacaa	gtgtcaggcttgaggtggag	2
OCT4	gcttcaagaacatgtgtaagctg	cacgagggtttctgctttg	69
PDGFRA	tgcctgacattgaccctgt	ccgtctcaatggcactctc	63
SOX2	ttaaaagttctagtggtacggtaggag	ttcgtcgcttggagactagc	4
TFCP2L1	cctggtccaccacacctatt	atggtcatctttggcctcac	2