Capacitation of human naïve pluripotent stem cells for multi-lineage differentiation

Maria Rostovskaya*, Giuliano G. Stirparo and Austin Smith*

Wellcome-MRC Cambridge Stem Cell Institute
Cambridge CB2 1QR,
United Kingdom

* Authors for correspondence (mr631@cam.ac.uk; austin.smith@cscr.cam.ac.uk)

Running title: Naïve to competent transition

SUMMARY STATEMENT

Human naïve pluripotent stem cells acquire competence for multilineage differentiation following a developmental trajectory comparable to primate embryonic epiblast that is facilitated by inhibition of the Wnt pathway

ABSTRACT

Human naïve pluripotent stem cells (PSC) share features with pre-implantation epiblast. They thus provide an unmatched opportunity for characterising the developmental programme of pluripotency in *Homo sapiens*. Here we confirm that naïve PSC do not respond directly to germ layer induction, but must first acquire competence. Capacitation for multi-lineage differentiation occurs without exogenous growth factor stimulation and is facilitated by inhibition of Wnt signalling. Whole transcriptome profiling during this formative transition highlights dynamic changes in gene expression, affecting many cellular properties including metabolism and epithelial features. Notably, naïve pluripotency factors are exchanged for post-implantation factors, but competent cells remain devoid of lineage-specific transcription. The gradual pace of transition for human naïve PSC is consistent with the timespan of primate development from blastocyst to gastrulation. Transcriptome trajectory during in vitro capacitation of human naïve cells tracks the progression of epiblast during embryogenesis in *Macaca fascicularis*, but shows greater divergence from mouse development. Thus the formative transition of naïve PSC in a simple culture system may recapitulate essential and specific features of pluripotency dynamics during an inaccessible period of human embryogenesis.

INTRODUCTION

Pluripotency refers to a flexible potential for individual cells to give rise to all lineages of the embryo. This property is a foundational feature in amniote embryogenesis (Sheng, 2015). Pluripotency extends from initial emergence of the epiblast a few days after fertilisation until lineage commitment during gastrulation. The period varies from around 4 days in mouse and other rodents to 8-10 days or longer in primates, including *Homo sapiens,* and in many other mammals. Over this time, pluripotent cells change in character from the initial naïve condition to a lineage-primed state poised for commitment (Morgani et al., 2017; Nakamura et al., 2016; Nichols and Smith, 2009; Smith, 2017). This dynamic transition is manifest at the cellular level by epithelialisation and molecularly by reconfiguration of transcriptome, epigenome and metabolism (Bedzhov and Zernicka-Goetz, 2014; Buecker et al., 2014; Kalkan et al., 2017; Mohammed et al., 2017; Zhou et al., 2012).

Cultures of pluripotent stem cells (PSC) can be derived from embryos (Evans and Kaufman, 1981; Martin, 1981; Thomson et al., 1998) or generated by molecular reprogramming (Takahashi and Yamanaka, 2006). Different PSC phenotypes present a spectrum of pluripotent states (Enver et al., 2009; Hackett et al., 2017; Hackett and Surani, 2014; Hough et al., 2014; Nichols and Smith, 2009), some of which show correspondence with stages of in vivo progression of epiblast, while others may be artefactual products of the culture environment (Gokhale et al., 2015; Smith, 2013). In the mouse, canonical embryonic stem (ES) cells cultured in defined conditions are considered to be counterparts of the naïve epiblast from which they are derived (Boroviak et al., 2014; Brook and Gardner, 1997). Mouse post-implantation epiblast-derived stem cells (EpiSC) (Brons et al., 2007; Tesar, 2005) in contrast resemble gastrulating epiblast of the anterior primitive streak (Kojima et al., 2014; Tsakiridis et al., 2014), and are accordingly classified as primed (Nichols and Smith, 2009). Human and other primate PSC, as conventionally established and propagated, are overtly different from mouse ES cells and are transcriptionally distinct from pre-implantation epiblast (Nakamura et al., 2016; Rossant, 2015; Rossant and Tam, 2017; Yan et al., 2013). They display post-implantation features (Nakamura et al., 2016) although positioning on the developmental axis is uncertain, both because of variation between cell lines and culture conditions and because there is no human reference available for early post-implantation embryogenesis. Recently, culture conditions have been devised that sustain human PSC with many of the expected properties of naïve pluripotency (Takashima et al., 2014; Theunissen et al., 2016; Theunissen et al., 2014). Naïve cells can be generated by resetting conventional PSC (Guo et al., 2017), by somatic cell reprogramming (Kilens et al., 2018; Liu et al., 2017), or by derivation directly from dissociated human inner cell mass cells (Guo et al., 2016). They exhibit transcriptome correlation with pre-implantation epiblast (Nakamura et al., 2016; Stirparo et al., 2018) and show protein expression of naïve epiblast-specific transcription factors such as KLF4, KLF17 and TFCP2L1 (Guo et al., 2016; Takashima et al., 2014).

Human naïve PSC provide an opportunity for simulation of the developmental programme of human pluripotency prior to gastrulation. They may thereby open a window into events during the second week of gestation that cannot be characterised or even observed in human embryos in utero. This is a period of major change that appears critical for establishing differentiation competence (Rossant and Tam, 2017; Sheng, 2015; Smith, 2017). Notably, mouse naïve ES cells do not differentiate directly into germ cell or somatic lineages, but first transit through the

peri- and early post-implantation phase of epiblast development (Hayashi et al., 2011; Kalkan et al., 2017; Mulas et al., 2017). During this formative transition, naïve cells are proposed to gain competence for lineage induction through a process of capacitation (Kalkan and Smith, 2014; Smith, 2017). In mouse ES cells, capacitation occurs within 24-48 hours (Hayashi et al., 2011; Mulas et al., 2017), reflecting the rapid rate of peri-implantation development in rodents (Acampora et al., 2016). For primate naïve PSC, the process may be expected to extend over several days, in line with slower embryogenesis (Nakamura et al., 2016; Smith, 2017). However, current methods for capacitating naïve PSC require prolonged culture for more than 20 days to achieve robust multilineage differentiation (Guo et al., 2017). The developmental relevance of this protracted conversion is further questioned by poor viability, cellular heterogeneity, and rather low efficiency. Here we set out to determine conditions under which naïve PSC may recapitulate in utero progression to late epiblast, fully competent for germ layer induction.

RESULTS

Naïve hPSC do not respond immediately to somatic lineage induction

Throughout this study we compared the conventional hESC line H9EOS with reset naïve derivative cR-H9EOS (Guo et al., 2017) and with the embryo derived naïve line HNES1 (Guo et al., 2016). We first tested multilineage differentiation via embryoid body formation in non-instructive serum-free conditions, a context permissive for the three primary germ layers. PSC were aggregated in suspension in N2B27 medium for up to 14 days. Conventional cells developed into typical embryoid body structures with downregulation of pluripotency markers *NANOG* and *POU5F1* (OCT4) and robust expression of markers of the three germ layers (Fig. 1A, B). Naïve cells formed fewer and smaller aggregates, with extensive cell death. They retained substantial expression of *POU5F1* and *NANOG* while *SOX17* and *VIM* differentiation markers were modestly upregulated, but markers for neuroectoderm, *SOX1*, *PAX6*, and *MAP2,* were not detected. The neural markers were also not expressed in outgrowths from naïve PSC aggregates (Fig. S1A). Thus, neural induction from naïve cells is not observed in embryoid body conditions, whilst other differentiation markers are lowly expressed.

Embryoid body differentiation is dependent on efficiency of cell aggregation and cell-cell interactions, parameters that are difficult to standardise. We therefore investigated the response of naïve cells to directed differentiation in adherent culture using protocols proven for conventional PSC. For these experiments naïve PSC were exchanged directly from self-renewal medium containing inhibitors of MEK/ERK and aPKC (Guo et al., 2017; Takashima et al., 2014) to the respective lineage induction media.

For neuroectoderm induction we employed dual SMAD inhibition (Chambers et al., 2009). By 10 days, quantification of SOX1 and PAX6 immunostaining by flow cytometry showed that H9EOS cultures comprise 90% neural lineage cells. In contrast, cR-H9EOS or HNES1 cultures contained fewer than 0.4% cells stained for either marker (Fig. 1C). Failure of direct neural induction is consistent with our previous observations on cR-Shef6 (Guo et al., 2017) and is further substantiated by absence of mRNA expression for *SOX1*, *PAX6*, *BRN2*, *FOXG1* (Fig. 1D). Definitive endoderm induction (Loh et al., 2014) applied to conventional hPSC such as H9EOS or Shef6 generally results in around 90% CXCR4+ SOX17+ cells detected by flow

cytometry on day 3. In contrast, naïve PSC cultures remained negative for both markers (Fig. 1E), again consistent with previous observations (Guo et al., 2017). Naïve PSC also failed to upregulate mRNA for *SOX17*, *CER1*, *HHEX*, *LHX1* and *FZD8* (Fig. 1F). During paraxial mesoderm differentiation (Chal et al., 2016), conventional hPSC expand during the 6-day protocol (Fig. 1G), undergo epithelial-to-mesenchymal transition (EMT), upregulate markers characteristic for paraxial mesoderm and EMT (*TBX6*, *MSGN1*, *HES7*, *CDH2*, *SNAI1*, *ZEB1*, *ZEB2*, *VIM*), and downregulate epithelial *CDH1* (Fig. 1H). In contrast, naïve PSC showed high levels of cell death and the few remaining cells did not adopt mesenchymal morphology, lacked EMT markers, retained expression of *CDH1*, and showed no or little upregulation of PM markers (Fig. 1G, H).

We further assessed the fate of naïve PSC exposed to differentiation conditions, either via embryoid body formation (Fig. S1B) or by monolayer induction of neuroectoderm or definitive endoderm (Fig. S1C). Naïve and general pluripotency markers (*KLF4*, *KLF17*, *TFCP2L1*, *OCT4*, *NANOG*) were downregulated in most cases, although still detectable. Genes characteristic of post-implantation epiblast such as *TCF15*, *SOX11* and *HES1* (Boroviak et al., 2015; Nakamura et al., 2016) were generally upregulated, although to variable levels. These observations indicate that upon withdrawal from self-renewing conditions a proportion of naïve PSC may progress towards a post-implantation formative epiblast identity irrespective of environment.

These findings confirm and extend previous indications (Guo et al., 2017; Liu et al., 2017) that human naïve PSC lack competence to respond productively to inductive cues for lineage specification.

Naïve hPSC begin transition following withdrawal from self-renewal culture

We have previously shown that naïve hPSC are able to differentiate into somatic lineages following a period of adaptation to culture in conventional hPSC media, such as mTESR, FGF/KSR or E8 (Guo et al., 2017; Takashima et al., 2014). Therefore, we surmise that naïve hPSC can be capacitated for somatic lineage induction, a process we have termed formative transition (Smith, 2017). However, this transition is accompanied by significant cell death and considerable cellular heterogeneity in the above conditions, and moreover takes longer than 20 days before stabilisation in a conventional PSC-like state (Guo et al., 2017). We therefore sought to achieve capacitation with improved consistency and efficiency, and over a developmentally more relevant time scale.

We first compared capacitation in E8, which contains TGFβ and FGF2 (Chen et al., 2011), with transition in N2B27 without added growth factors. In both conditions, a proportion of cells acquired and maintained flattened epithelioid morphology similar to conventional hPSC during the first 7 days. However, we observed that cell survival was much improved in N2B27. In E8 cells could be propagated further, albeit with considerable cell death, and eventually stabilise in a conventional PSC-like state. In contrast, in N2B27 the population became increasingly heterogeneous after day 7, and by day 12 all cells appeared differentiated and ceased proliferation (Fig. 2A).

We tested ability to re-form naïve colonies after periods in E8 or N2B27. Cells were re-plated at clonal density in naïve PSC culture conditions each day after switching to E8 or N2B27, and alkaline phosphatase-positive colonies were scored 5-7 days later. In both media, the proportion of cells capable of colony formation dropped

markedly during the first 3 days (Fig. 2B). We monitored expression of pluripotency genes by RT-qPCR (Fig. 2C). In both media naïve-specific transcription factors *KLF4* and *TFCP2L1* were downregulated over 3 days while *KLF17* reduced more gradually. *OCT4* and *NANOG* declined less and then stabilised from day 7 in E8 but continued to fall in N2B27. *SOX2* was relatively stable in E8 but downregulated progressively in N2B27.

We assayed monolayer differentiation into neuroectoderm and definitive endoderm after 7 days of treatment, when cells in N2B27 or E8 appear similar. Both populations produced SOX1-positive neuroectodermal cells and CXCR4/SOX17 double positive endoderm (Fig. 2D). However, the efficiencies of differentiation from multiple independent experiments were variable and generally low, with maximum values around 27% for neuroectoderm and 50% for definitive endoderm (Fig. 2E). In contrast we have previously shown that longer term culture (>20 days) in E8 results in efficiencies of 80-90% for both lineages, comparable to conventional PSC (Guo et al., 2017).

These results indicate that naïve cells can reach somatic lineage competence over 7 days, with or without provision of exogenous FGF and TGFβ, but these conditions are only partially effective.

Inhibition of WNT signalling facilitates capacitation

At day 7 we noted a significant fraction of *TBXT* (*Brachyury*) expressing cells that were exclusive to NANOG-positive cells (Fig. 2F, G). *TBXT* is a known target of canonical Wnt signalling (Yamaguchi et al., 1999). We also detected expression of Wnt ligands and pathway components (Fig. S2). We therefore examined whether endogenous WNT activity might disrupt or divert the transition process. We tested capacitation in N2B27 supplemented with the tankyrase inhibitor XAV939 (2μm), which blocks canonical Wnt signalling (Huang et al., 2009). In contrast to cultures in N2B27 alone or E8, we observed relatively uniform establishment of epithelial morphology similar to conventional hPSC throughout the culture in the presence of XAV939 (Fig. 3A). Cells expanded continuously throughout this conversion (Fig. 3B). Moreover, unlike in N2B27 only, proliferation was sustained for at least 20 days. Thereafter, the cultures could still be maintained, but became heterogeneous. We found that addition of FGF2 and activin A to XAV939 (hereafter XAF) after day 10 allowed for continued propagation with minimal overt differentiation for at least 50 days (Fig. S3). XAF is similar to medium used for culturing mouse EpiSC as more homogeneous populations (Sumi et al., 2013; Tsakiridis et al., 2014).

Colony assay showed that the ability to self-renew in naïve conditions was greatly diminished after 2 days in N2B27 plus XAV and almost eliminated by day 3, similar to cultures in N2B27 only or E8 (Fig. 3C). Likewise, naïve pluripotency factor (*KLF4*, *TFCP2L1*) transcripts were reduced to a very low or undetectable level over this period, while *KLF17* declined more slowly and was extinguished only by 10 days of transition. In distinction to cultures in N2B27 alone, however, *NANOG* and *POU5F1* expression stabilised from day 7 at similar levels to those observed in conventional hPSC (Fig. 3D). We also noted that *DPPA3* was downregulated from day 6 onwards. These observations were substantiated by immunostaining (Fig. 3E, F).

We tested directed differentiation after 10 days in XAV939 when cells display a pluripotency marker profile similar to conventional PSC. Upon dual SMAD inhibition, we observed robust upregulation of PAX6 and SOX1, as shown by immunostaining

and quantified by flow cytometry (Fig. 4A, B). The efficiency was comparable to conventional hPSC. RT-qPCR confirmed expression of these markers, along with *BRN2*, *MAP2*, *FOXG1* (Fig. 4C). These cells could be further differentiated to post-mitotic neurons, validated by immunostaining for TUBB3 (β-III-tubulin), MAP2, NEUN, and RT-qPCR for *MAP2*, *NEUN*, *NCAM1* and *ASCL1* (Fig. S4A, B). Definitive endoderm induction was also highly efficient, assayed by immunostaining for SOX17 and FOXA2 (Fig. 4D), and by flow cytometry which quantified co-expression of CXCR4 and SOX17 in 83.0% of cells (Fig. 4E). Importantly the primitive endoderm marker PDGFRa (Blakeley et al., 2015; Plusa et al., 2008; Stirparo et al., 2018; Yan et al., 2013) was not induced (Fig. S4C). RT-qPCR showed expression of definitive endoderm markers at levels similar to induction from conventional hPSC (Fig. 4F). Differentiation could be continued to PDX1-expressing foregut progenitors, with efficiencies >80% (Fig. S4D). Finally, in response to paraxial mesoderm induction, lineage markers and EMT genes were upregulated while *CDH1* was downregulated. TBX6 and CDX2 protein expression was confirmed by immunostaining (Fig. 4G). Paraxial mesoderm identity was substantiated by further differentiation to myotubes with expression of transcripts for sarcomeric proteins TTN and DMD (Fig. S4E). Immunostaining for sarcomeric myosin showed striated myofibers (Fig. S4F) and spontaneous contractions confirmed functional sarcomere assembly (Movie 1).

We repeated capacitation with XAV939 on multiple naïve PSC, including: embryo-derived HNES1, HNES5c1 and HNES5c5; reset ESC, cR-H9EOS and cR-S6EOS; reset iPSC, cR-LQT1. Lineage competence was consistently achieved across multiple independent experiments (Fig. S5, for selected cell lines) (Table S1).

XAV939 is a potent tankyrase (TNKS1 and 2) inhibitor (Huang et al., 2009), which stabilizes the axin-GSK3β complex resulting in degradation of β-catenin, but potentially can affect other cellular pathways (Lehtio et al., 2013). In order to confirm that capacitation is facilitated specifically by inhibition of the Wnt pathway, we tested an alternative mode of inhibition. IWP2 and C59 act on porcupine (PORCN) to prevent Wnt processing and secretion (Chen et al., 2009; Proffitt et al., 2013). Naïve hPSC cultured in N2B27 with C59, or IWP2, also produced cells with flat epithelioid morphology (Fig. S6A) that displayed similar multi-lineage competence as cells cultured with XAV939 (Fig. S6B-F).

PSC capacitated in XAV939 for 10 days could be further propagated for multiple passages in either XAF or E8 without signs of growth arrest or differentiation. In either media cells maintained their abilities to produce derivatives of three germ layers with efficiencies similar to conventional PSC, even >33d after initiation of transition (Fig. S7)

From these results we conclude that Wnt inhibition facilitates formative transition of human naïve cells. By 10 days, cells appear fully capacitated to produce neurectoderm, endoderm and mesoderm lineage precursors that can undergo further differentiation into tissue progenitors and post-mitotic cell types.

Global gene expression profiling during capacitation of naïve hPSC

To characterise gene expression dynamics during capacitation, we performed whole transcriptome RNA sequencing at days 0, 1, 2, 3, 7 and 10 (Fig. S8A). In addition, on day 10 cells were split into two conditions for continued maintenance, E8 or XAF medium, and passaged until 22-28 days in total (indicated as d22+). Conventional

H9EOS hESC in E8 were used as a reference. We prepared biological triplicate samples for all cell lines and conditions.

Pearson correlation analysis based on all expressed genes divided samples into two major populations; early, day 0-3, and late, day 7 onwards (Fig. S8B). HNES1 and cR-H9EOS samples were highly correlated over the entire time course, indicating a shared trajectory. We determined the number of variable genes for all pairwise comparisons, which confirmed that differential expression increased markedly between early and late samples (Fig. S8C).

We further examined expression dynamics after 10 days, following transfer to either E8 or XAF. Pearson correlation analysis of samples on day 10, day 22+, and control cells, reveals overall similarity with correlation coefficients greater than 0.95 in all comparisons (Fig. S8D). PCA shows two major groups correlating to the culture conditions; E8 medium or XAV939-containing media (Fig. S8E). Of note, cR-H9EOS cells capacitated and then maintained in E8 medium are most similar to conventional H9EOS hESC cultured in E8. Thus, following capacitation naïve cells can regain their original conventional PSC state in terms of global gene expression if they are expanded in comparable culture conditions.

We examined the dynamics of gene expression during the time course. PCA for all variable genes confirmed that HNES1 and cR-H9EOS follow similar trajectories (Fig. 5A). Biological triplicates showed high consistency. We then generated a list comprising all genes differentially expressed between any two time points of the time course and in at least one cell line (padj<0.01; fold change >2). Soft clustering analysis distinguished 5 major clusters, identified by minimizing the total variation within each cluster (Fig. S9A): 1 – “early down”, 2 – “late down”, 3 – “up-n-dn, 4 – “early up”, 5 – “late up”, (Fig. 5B, C). We compared the clusters between the two cell lines, considering genes of the same and similar clusters. Similar clusters are those in which genes changed expression in the same direction, but with different dynamics: cluster 1 and 2 (downregulated); clusters 4 and 5 (upregulated). For the two cell lines, 61.7% of genes belonged to the same clusters, and 83.9% were in the same and similar clusters (Fig. 5D, E). PCA showed proximity between the same clusters, indicating that not only the dynamics, but also the levels of gene expression are similar between the two cell lines (Fig. S9B).

We inspected representation of KEGG pathways related to the clusters (Fig. S9C). We noted that genes related to oxidative phosphorylation are mostly downregulated (clusters 1 and 2) while those associated with glycolysis show no prevailing direction of change (Fig. S10A, Table S6). We used TMRE staining to assay mitochondrial membrane potential and observed reduced activity in capacitated compared with naïve PSC (Fig. S10B-D). This is in accord with reduction of mitochondrial respiration observed during formative transition in mouse (Fiorenzano et al., 2016; Kalkan et al., 2017; Zhou et al., 2012). Genes related to focal adhesion, cell-cell adhesion, ECM-receptor interactions and adherent junctions were predominant in the upregulated clusters (4 and 5). We inspected expression of a panel of genes associated with tight junctions, cell polarity, integrins and cadherins and observed that many are upregulated, some continuously but others only after day 3 or even later (Fig. S11). An exception is *Cdh1* (E-cadherin), which is downregulated but still expressed. These data highlight gain of epithelial features as a major aspect of

formative transition in vitro, consistent with development of a laminar epithelial epiblast disk in the early post-implantation embryo (Sheng, 2015).

We inspected expression of ligands and receptors of the TGF-β family and noted that both nodal and the convertase furin are down-regulated at the end of transition, while TGFBR1 (ALK5), TGFBR2 and ACVR2 receptors gain expression (Fig. S12). These dynamics are consistent with the requirement for exogenous activin for robust expansion after capacitation.

Dynamics of selected genes were then assessed (Fig. 5E). Genes characteristic for naïve epiblast, such as *KLF4*, *TFCP2L1*, *DNMT3L*, *FGF4*, and *KLF17* (Stirparo et al., 2018; Takashima et al., 2014), were downregulated to very low levels. Broader expressed pluripotency factors *NANOG* and *POU5F1* were partially downregulated while *SOX2* expression was relatively constant. These results are consistent with RT-qPCR and immunofluorescence analyses. We noted that DNMT3B was upregulated early whereas DNMT3A expression was maintained at a similar level (Fig. S12). Markers characteristic for early post-implantation epiblast (Boroviak et al., 2015; Nakamura et al., 2016) were among the upregulated genes, including *SOX11*, *FZD7*, *CDH2*, and *SALL2*. We also evaluated genes that differ in expression between mouse and primate post-implantation epiblast. Among genes reported as specific for mouse early post-implantation stages (Boroviak et al., 2015), some were not upregulated (*FGF5*, *POU3F1*, *NODAL*), whereas others showed mild upregulation (*SOX3* and *SOX4*). On the other hand, markers distinctive for primate post-implantation epiblast were upregulated, among them *TCF7L1* (encoding TCF3) and *FGF2*. Of note, most early lineage markers, such as *SOX1*, *PAX6*, *CDX2*, *GATA4*, *GATA6*, *SOX17*, *FOXA2* showed very low or undetectable expression, even at the end of the time course. Only *TBXT* and *MIXL1* were found at low but detectable levels. Furthermore, this profile was maintained after extended expansion in E8 or XAF (Fig. S13). We validated dynamics of gene expression during formative transition by RT-qPCR analysis of independent experiments using HNES1, HNES5c2 and cR-H9EOS cell lines (Fig. S14). This also revealed that while cells at day 7 and 10 of capacitation share many features, some genes such as *KLF17*, *TCF7L1* and *CDH2* show ongoing expression changes between these timepoints, indicating that up to 10 days are required to achieve a stable profile.

These analyses reveal a global and dynamic reconfiguration of gene expression during formative transition, with implications for metabolism, transcriptional regulation and cell biological properties.

Comparison of gene expression dynamics between *in vitro* capacitation and *in utero* progression of primate epiblast

To assess concordance between *in vitro* capacitation and formative transition *in vivo*, we compared our transcriptome results to the published single cell RNAseq data for early embryogenesis in the mouse (Boroviak et al., 2018; Mohammed et al., 2017) and the cynomolgus monkey, *Macaca fascicularis* (Nakamura et al., 2016). From the mouse embryo dataset, we used cells of pre-implantation (day 4.5, EPI), early (day 5.5, post-E) and late (day 6.5, post-L) post-implantation epiblast. From the macaque embryo, we selected cells assigned as pre-implantation (EPI), early post-implantation (post-E) and late post-implantation (post-L) epiblast in the original study. For both species, we selected variable genes during epiblast progression (padj<0.01, fold change >2) and identified 5 clusters of genes with distinct dynamic behaviour by soft clustering analysis (Fig. 6A).

We compared dynamics of gene expression in the clusters of the embryonic epiblast progression with the *in vitro* time course. We noted that clusters 1, 2, 4 and 5 showed comparable behaviour between mouse and macaque embryo and hPSC, and could be regarded as equivalent. Cluster 3 genes behaved slightly differently. In mouse and macaque embryos, these genes were upregulated from EPI to post-E, then downregulated in post-L, but maintained at higher level than in EPI. In contrast, most genes of cluster 3 of the *in vitro* time course showed similar expression at the start and end points. Therefore, we considered two groups of similar clusters: downregulated (embryo 1 and 2, and hPSC 1 and 2); upregulated (embryo 3, 4 and 5, and hPSC 4 and 5). Examination of overlap between the clusters of dynamic genes during the hPSC transition and mouse embryo revealed that a significant proportion of genes belonged to the same (28.9%) or similar clusters (66.3%). Furthermore, comparison of the hPSC transition with macaque embryo showed a greater overlap, 39.0% of variable genes were in the same clusters and 83.1% were in similar clusters (Fig. 6B).

Among genes with similar dynamics during capacitation of hPSC and in both mouse and macaque embryos are: *KLF4* and *SPIC*, downregulated; *GBX2*, up and down; *SOX11, ETV4 and ETV5,* upregulated (Fig. 6C). Genes dynamically expressed in hPSC and macaque that were not detected in the mouse and have not been characterised in the context of pluripotency progression include *DLL3* and *GLI2*. Overall, therefore, *in vitro* capacitation of hPSC shares gene expression features with mouse but more closely resembles development of primate epiblast.

We further focused on the comparison of the formative transition of hPSC *in vitro* and macaque embryo epiblast *in utero*. We performed PCA with hPSC during transition time course and the embryo cells from ICM, EPI, post-E and post-L stages. The two datasets are separated in PC1, which can be attributed to differences in methodology (single cell vs bulk RNAseq, alternative sequencing chemistries), species and environment. PC2 and PC3 reflect gradual progression in the embryo from ICM to EPI, then to post-E and finally to post-L (Fig. 6D). Naïve hPSC are closest to the EPI, consistent with previous analyses (Nakamura et al., 2016; Stirparo et al., 2018), and further during *in vitro* capacitation hPSC align with the progression of embryonic epiblast. We validated these results by comparing samples during capacitation to the embryo stages EPI, post-E, post-L (Fig. 6E, F). Gene expression for the embryo samples was calculated as average values of single cells. Using quadratic programming (Gong and Szustakowski, 2013), we measured relative similarities of hPSC samples to EPI, post-E and post-L. Similarity to EPI decreased continuously during capacitation. Fraction of similarity to post-E increased during the early steps of the time course (d1-3) and then reduced on day 7-10. Conversely, similarity to post-L remained relatively low during day 0-3, and increased on days 7 and 10. Therefore, the trajectory of *in vitro* capacitation follows to the progression of embryonic epiblast of non-primate embryo *in utero*.

DISCUSSION

Our results demonstrate that human naïve PSC are not equipped to enter directly into lineage specification but must first undergo a formative transition. During this capacitation process, PSC down-regulate naïve pluripotency transcription factors, rewire metabolism and signalling pathways, develop epithelial character, and become fully competent for differentiation to embryonic lineages. At the end of this transition cells exhibit dependence on exogenous FGF and activin/TGFβ for

continued expansion. Capacitation takes up to 10 days and follows a trajectory and timeline reflective of the progression of primate epiblast from ICM to gastrulation. These findings are consistent with the postulate of naïve pluripotency as a *tabula rasa*, in which potential for multi-lineage differentiation is created, but not actuated (Nichols and Smith, 2009; Smith, 2017).

Classically, pluripotency is considered as the capacity of single cells to form all embryonic lineages with no predetermination. Highlighting capacitation as a pre-requisite for multi-lineage differentiation accords with a more refined concept, in which pluripotency refers to potential rather than actual capacity. Naïve pluripotent cells can ultimately give rise to all cell types of the body, but only via an obligate transition during which competence is instated. Until recently this distinction has been obscured by the compression of events in rapidly advancing mouse ES cell differentiation. Indeed, mouse ES cells have often been presented as responding directly to lineage cues. However, closer inspection has revealed that in reality naïve cells first lose ES cell identity and transit to a population, termed epiblast-like (EpiLC) or formative, that is enabled for lineage induction (Buecker et al., 2014; Hayashi et al., 2011; Hoffman et al., 2013; Kalkan et al., 2017; Mulas et al., 2017). When human naïve PSC are exposed directly to inductive environments, they show a heterogeneous response, including cell death, partial transition to post-implantation epiblast, and differentiation into miscellaneous phenotypes. Acquisition of competence is therefore a shared requirement for mouse and human naïve PSC.

For mouse ES cells capacitation can proceed in medium without growth factors or inhibitors other than insulin. In human, however, this process is unreliable due to disruption by endogenous Wnt activity. Interestingly, and in contrast to mouse ES cells, human naïve PSC self-renewal may also be perturbed by Wnt signalling (Guo et al., 2011; Theunissen et al., 2016; Zimmerlin et al., 2016). Indeed, for most experiments in the present study we used human naïve PSC that had been maintained in the presence of XAV939 along with MEK and aPKC inhibitors (Bredenkamp et al., submitted). Regardless of the naïve PSC maintenance condition, we found that during capacitation continuous Wnt pathway inhibition markedly improves the efficiency and consistency of transition. This finding may relate to the known role of Wnt signalling in promoting axis formation and gastrulation (Huelsken et al., 2000; Liu et al., 1999; Morkel et al., 2003). Activation of the Wnt pathway via GSK3 inhibition is also a key component in two of the three lineage induction protocols for hPSC. Our findings indicate that prior to competence, transitional epiblast should be shielded from Wnt stimulation to avoid inappropriate gene induction and miscellaneous differentiation.

The formative transition is considerably faster in mice than in primates, both in vitro and in the embryo. One contribution to different developmental timing may be the appropriation of *TCF7L1* (TCF3) in mouse but not primate naïve epiblast (Boroviak et al., 2018). In mouse ES cells TCF3 acts as a potent repressor of naïve network transcription factors (Martello et al., 2012). Inhibition of glycogen synthase kinase 3 supports mouse ES cell self-renewal principally by abrogating the repressor activity of TCF3 (Wray et al., 2011) and TCF3 depletion substantially delays mouse ES cell exit from naïve pluripotency (Betschinger et al., 2013; Guo et al., 2011; Pereira et al., 2006). Without TCF3, therefore, human naïve cells lack the major accelerator of naïve state exit. Subsequent development of lineage competence is also slower in human, however, implicating additional determinants of overall transition timing.

The slower pace of formative transition in human may make the sequence of events easier to delineate and mechanistically dissect than in mouse. Indeed, whereas exit from the naïve state and gain of multilineage competence are difficult to dissociate in mouse ES cells (Mulas et al., 2017), they appear to be separated by several days in the human system. Accordingly, we identified two major waves of dynamic transcriptome change. Over the first 3 days the cells down-regulate a subset of naïve factors (KLF4, TFCP2L1) coincident with the reduction in ability to reform naïve colonies. Some post-implantation markers (TCF15, FGF2, HES1) are up-regulated early and oxidative phosphorylation components reduced. The end of this period marks the exit from naïve pluripotency for the bulk population. The second wave of transcriptional changes features loss of other naïve factors (KLF17, DPPA3 and DPPA5), marked up-regulation of multiple genes associated with epithelial function, and gain of TCF7L1 and TCF7L2, mediators in the canonical Wnt pathway. This complex profile indicates that the formative transition is wide-reaching and comprised of distinct steps.

Gene expression dynamics during capacitation of naïve hPSC exhibit similarity with epiblast progression in a non-human primate embryo, arguing that the transition path is not merely an in vitro phenomenon. In the embryo, epiblast undergoes profound cellular and molecular changes during the peri-implantation and early post-implantation period, preparatory to gastrulation (Acampora et al., 2016; Mohammed et al., 2017; Nakamura et al., 2016; Peng et al., 2016). Our transcriptome analyses indicate that hPSC acquire full competence for somatic cell differentiation once they reach a state that is similar to the late pre-gastrulation epiblast. During subsequent expansion in FGF2 with TGFβ or activin A, cells converge closer to conventional hPSC, consistent with elaboration of a growth factor driven stem cell phenotype.

In summary, when shielded from Wnt signalling but exposed to ERK1/2 and aPKC activity that are inhibited during self-renewal (Takashima et al., 2014), human naïve PSC convert efficiently to a pluripotent condition that is empowered for specification and commitment. This competent state lacks overt transcriptional lineage priming, but can be reliably induced to undergo productive differentiation into endodermal, mesodermal and neuronal cell types. Furthermore, once acquired, multilineage competence can be stably maintained, similarly to conventional human PSC. At the technical level, this simple and reliable transition system provides a platform for systematic evaluation of the differentiation propensity and consistency of human naïve PSC and thereby for rigorous comparison with other types of hPSC culture. More fundamentally, the naïve PSC transition provides a window into a crucial phase of human embryogenesis that cannot be accessed in vivo. It will be of great interest to characterise the sequence of molecular events and the mechanisms that underlie acquisition of embryonic lineage competence in this system.

MATERIALS AND METHODS

Human pluripotent stem cell lines

Experiments were performed in parallel throughout on the embryo derived naïve hPSC line HNES1 (Guo et al., 2016) and the reset naïve hPSC line cR-H9EOS (Guo et al., 2017). H9EOS cells, conventional H9 hESC (Thomson et al., 1998) carrying the EOS reporter (Guo et al., 2017), were used as reference. Validation experiments were carried out on two clonal embryo-derived naïve hPSC (HNES5c1 and HNES5c2; unpublished), and on reset hESC described in Guo et al., 2017 or reset iPSC (LQT1, (Chen et al., 2017; Moretti et al., 2010)) generated by the method of Guo et al., 2017. HNES cells were derived with informed consent under licence from the Human Embryology and Fertilisation Authority.

hPSC Maintenance

Cells were cultured throughout in a humidified incubator with 5% O_2 and 7% CO_2 at 37^{o}C.

Naïve hPSC were maintained on irradiated mouse embryonic fibroblasts (MEF) feeder cells in N2B27 (for details see Supplementary Materials and Methods) supplemented with 1μM PD0325901, 10ng/ml human LIF (produced in house), 2μM Gö6983 (Tocris Bio-Techne, Cat. 2285), and either 0.3-1μM CHIR99021 (Guo et al., 2016; Takashima et al., 2014) or 2μM XAV939 (Tocris Bio-Techne, Cat. 3748) (Guo et al., 2017). Geltrex (ThermoFisher Scientific, A1413302) was optionally added to the medium at a concentration 1μl/ml during re-plating. Cells were passaged using TrypLE (ThermoFisher Scientific, Cat. 12605028). ROCK inhibitor (10μM; Y-27632, Cat. 688000, Millipore) was added for 24 hours after passaging.

Conventional hPSC were cultured in E8 medium (prepared in house according to (Chen et al., 2011) on Geltrex pre-coated plates and passaged using 0.5mM EDTA in PBS.

Capacitation

Prior to capacitation, naïve hPSC were passaged once without feeders in naïve PSC medium plus Geltrex at 1μl/cm^2. For the formative transition, cells were dissociated with TrypLE and plated to Geltrex coated tissue culture plates at a seeding density 1.6x10^4/cm^2 in medium for naïve hPSC supplemented with 10μM ROCK inhibitor. After 48 hours, cells were washed with DMEM/F12 supplemented with 0.1% BSA. Capacitation was then performed in the following conditions: E8 medium; N2B27 medium without supplementation; N2B27 supplemented with one of the following WNT inhibitors: 2μM XAV939 (Tocris Bio-Techne, Cat. 3748), or 1μM IWP2 (Tocris Bio-Techne, Cat. 3533), or 1μM WNT-C59 (Tocris Bio-Techne, Cat. 5148). Medium was renewed every 1-2 days. Cells were passaged at a 1:2 ratio at confluency using TrypLE and10μM ROCK inhibitor. Cells were replated for lineage induction after 10 days, unless otherwise specified.

For expansion after capacitation, cells were cultured in either E8 or N2B27 supplemented with 2μM XAV939, 3ng/ml Activin A and 10ng/ml FGF2 (XAF medium, modified from (Sumi et al., 2013). The cells could also be maintained in capacitation conditions (2μM XAV939 in N2B27), but were prone to differentiation beyond 20 days. During expansion, cells were cultured on Geltrex pre-coated tissue culture plates and passaged by dissociation with either 0.5mM EDTA or TrypLE. In the latter case, 10μM ROCK inhibitor was added for 24 hours after dissociation.

Cells could be frozen during capacitation in N2B27 with 10% DMSO. ROCK inhibitor (10μM) was added for 24 hours after thawing. Cells were passaged after thawing before setting up differentiation assays.

Colony assay

For colony formation assay, cells were dissociated with TrypLE and plated to Gelterx coated 12-well plates at a density 1, 2 or $4x10^3$ cells/well in naïve medium supplemented with 10μM ROCK inhibitor. After 5-7 days colonies were fixed and stained for alkaline phosphatase (Sigma-Aldrich, Cat. 86R). Whole well images were acquired using Olympus IX51 and CellSens software and colonies were scored either manually or automatically using Ilastik and Fiji software.

***In vitro* differentiation**

Embryoid bodies. hPSC were aggregated in Aggrewell plates with 400μm microwells (Stem Cell Technologies, Cat. 34411) in N2B27 supplemented with 10μM ROCK inhibitor. After 2 days, aggregates were flushed from the wells in N2B27 medium and transferred to non-adhesive 6-well plates for further culture in suspension. On day 7 aggregates were optionally plated to tissue culture grade plates coated with Geltrex for outgrowth differentiation. RT-qPCR analysis was performed on day 14 day.

Neuroectoderm. Neuroectoderm was induced based on (Chambers et al., 2009) in N2B27 supplemented with 1μM A8301 (Tocris Bio-Techne, Cat. 2939) and 500nM LDN193189 (alternative name DM3189, Axon Medchem, Cat. 1509), for 10 days. For subsequent neuronal differentiation, cells were dissociated after 10 days, passaged to plates pre-coated with poly-L-ornithine and laminin (both from Sigma) in N2B27 at a density 10^5 cells/cm^2, and cultured for an additional 30 days.

Endoderm. Definitive endoderm was induced according to (Loh et al., 2014). Cells were cultured in CDM2 basal medium supplemented with 100ng/ml Activin A (produced in house), 100nM PI-103 (Tocris, Bio-Techne, Cat. 2930), 3μM CHIR99021, 10ng/ml FGF2 (produced in house), 3ng/ml BMP4 (Peprotech, Cat. 120-05ET), 10μg/ml heparin (Sigma-Aldrich, H3149) for one day; then for the next 2 days the following supplements were applied: 100ng/ml Activin A, 100nM PI-103, 20ng/ml FGF2, 250nM LDN193189, 10μg/ml heparin. Further induction of foregut progenitors was performed according to (Rezania et al., 2014) with analysis at the S4 stage.

Paraxial mesoderm. Differentiation to paraxial mesoderm and myotubes was performed according to (Chal et al., 2016). Cells were cultured in 3μM CHIR99021 and 500nM LDN193189 for 6 days, with addition of 20ng/ml FGF2 from day 3-6. Multi-step induction of myotubes was continued up to 40 days of differentiation.

Quantitative RT-PCR

Total RNA was extracted using Reliaprep RNA Miniprep and 200-500ng used for reverse transcription using GoScript Reverse Transcription system (both from Promega). Quantitative PCR was performed with GoTaq qPCR Master Mix (Promega) using Universal Probe Library (Roche) or Taqman probes (ThermoFisher Scientific) for detection. Primer sequences and Taqman probes are listed in Tables S2 and S3. GraphPad Prism software was used for graphic representation. Each analysis was performed in parallel on separate cell lines and data are presented as means from technical duplicates for each line.

Flow cytometry

Cells were dissociated using Accutase and washed using PBS with 2% FCS. For surface marker staining, cells were incubated with directly conjugated antibodies (Table S4) diluted in PBS with 2% FCS for 30min at +4°C, followed by washing and resuspending in PBS. For intracellular marker staining, the cells were fixed with Fixation Buffer (00-8222-49, ThermoFisher Scientific) for 30min at +4°C, washed with Permeabilization Buffer (00-8333-56, ThermoFisher Scientific), and incubated

with antibodies diluted with Permeabilization Buffer and 5% donkey serum (Sigma-Aldrich) for 1 hour at +4°C. Detection was performed using a BD LSRFortessa instrument (BD Biosciences) with analysis using FlowJo software.

Immunofluorescence

Fixation was performed with 4% formaldehyde in PBS for 15min, permeabilization with 0.5% Triton X-100 in PBS for 10min, blocking with 3% BSA and 0.1% Tween-20 in PBS for 30min, all at room temperature. Antibodies (Table S5) were diluted in PBS with 0.1% Triton X-100 and 3% donkey serum. Incubation with primary antibodies was carried out overnight at +4°C, then secondary antibodies were added for 1 hour at room temperature. Slides were mounted using Prolong Diamond Antifade Mountant (Life Technologies ThermoFisher Scientific, Cat. P36970).

Assessment of mitochondrial membrane potential.

TMRM (Tetramethylrhodamine, methyl ester) was applied to cells in culture medium at 100nM for 30min at 37°C. Fluorescence was measured using flow cytometry, then an uncoupling agent of mitochondrial oxidative phosphorylation FCCP (carbonyl cyanide-p-trifluoromethoxyphenylhydrazone) was added to cell suspension at 1µM for 5 min and the measurement was repeated.

Microscopy, image processing and quantification

Immunofluorescent imaging was performed on a Leica DMI4000 microscope, or Andor Revolution XD spinning disk system for myotube visualisation. Immunofluorescent images of myotubes were deconvolved using Autoquant X3 (MediaCybernetics), and Maximum Intensity Projections created using Imaris (Bitplane). For quantification of immunofluorescent staining for KLF17, NANOG, KLF4, OCT4, mean intensity of staining over identified DAPI nuclei was measured using Cell Profiler (Broad Institute). Movies were acquired using Leica DMI4000 microscope, saved as .avi using Fiji software and then converted to .mp4 format.

Library preparation and RNA sequencing

Total RNA was prepared using Reliaprep RNA Miniprep kit (Promega). rRNA was depleted using Illumina's Ribozero HMR kit following the manufacturers protocol. Libraries were prepared using the KAPA Stranded mRNA-Seq Kit (Kapa Biosystems, Roche) on an Agilent Bravo liquid handling system and MJ thermocyclers. Libraries were sequenced on HiSeq 2500, single end 50bp reads.

Transcriptome analysis

Reads were aligned to human genome build GRCh38/hg38 with STAR 2.5.2b (Dobin et al., 2013) and human gene annotation from Ensembl release 87 (Yates et al., 2016). Htseq-count (Anders et al., 2015) was used to quantify expression to gene loci. Mouse samples were compiled from an earlier study (Mohammed et al., 2017) and analysed as described in (Boroviak et al., 2018). Macaque FPKM expression table from the dataset published in (Nakamura et al., 2016) was provided by the authors. Orthology 1-to-1 for cross-species comparison was used.

Cluster analysis and principal component analysis were computed using log2 FPKM values with the Bioconductor packages DESeq2 (Love et al., 2014) or FactoMineR (Lê et al., 2008) in addition to custom scripts. SCDE R package (Kharchenko et al., 2014), was used to perform differential expression analysis. Fractional identity between the sample of the time-course and the selected embryo stages was determined via quadratic programming using the R package DeconRNASeq (Gong and Szustakowski, 2013). Average expression levels of cells comprising distinct macaque embryo stages were used as the ‘signature’ dataset, and the relative identity of each time-course sample was computed by quadratic programming. Enrichment of KEGG pathways was computed with custom script using the KEGG

database resource (https://www.genome.jp/kegg/). Soft clusters were computed with R package MFuzz (Futschik and Carlisle, 2005; Kumar and E.M., 2007) and elbow method was used to determine the appropriate number of clusters. In order to detect genes with the greatest expression variability, a non-linear regression curve was fitted between log2 FPKM expression and the square of coefficient of variation. Thresholds were applied along the x-axis (log2FPKM) and y-axis (logCV2) to identify the most variable genes.

ACKNOWLEDGEMENTS

We are grateful to Ziad Al Tanouri, Olivier Pourquie, Balazs Varga, Valeria Orlova, and Christine Mummery for advice on differentiation assays. We thank Ge Guo, Milena Bellin and Christine Mummery for providing cell lines. Rosalind Drummond provided excellent technical support and James Clarke lab management. Wojciech Szlachcic helped with experiments during the initial phase of the project. We acknowledge the core facilities of the Cambridge Stem Cell Institute for support, help and advice. Peter Humphries helped with imaging and performed quantification of images; Maike Paramor and Vicki Murray prepared libraries for RNAseq. We thank Marko Hyoven for in-house growth factor provision. We are grateful to Tomonori Nakamura and Mitinori Saitou for providing the cynomolgus monkey embryo FPKM expression table from their published dataset. The MF20 monoclonal antibody, developed by Fischman, D.A., Weill Cornell Medical College, was obtained from the Developmental Studies Hybridoma Bank, created by the NICHD of the NIH and maintained at The University of Iowa, Department of Biology, Iowa City, IA 52242.

COMPETING INTERESTS

AS is an inventor on a patent application filed by the University of Cambridge relating to human naïve pluripotent stem cells

FUNDING

This research was funded by the Medical Research Council of the United Kingdom (G1001028 and MR/P00072X/1), the European Commission Framework 7 (HEALTH-F4-2013-602423, PluriMes) and the UK Regenerative Medicine Platform (MR/L012537/1). The Cambridge Stem Cell Institute receives core funding from the Wellcome Trust and the Medical Research Council. AS is a Medical Research Council Professor.

REFERENCES

Acampora, D., Omodei, D., Petrosino, G., Garofalo, A., Savarese, M., Nigro, V., Di Giovannantonio, L. G., Mercadante, V. and Simeone, A. (2016). Loss of the Otx2-Binding Site in the Nanog Promoter Affects the Integrity of Embryonic Stem Cell Subtypes and Specification of Inner Cell Mass-Derived Epiblast. *Cell Rep* **15**, 2651-2664.

Anders, S., Pyl, P. T. and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169.

Bedzhov, I. and Zernicka-Goetz, M. (2014). Self-organizing properties of mouse pluripotent cells initiate morphogenesis upon implantation. *Cell* **156**, 1032-1044.

Betschinger, J., Nichols, J., Dietmann, S., Corrin, Philip D., Paddison, Patrick J. and Smith, A. (2013). Exit from Pluripotency Is Gated by Intracellular Redistribution of the bHLH Transcription Factor Tfe3. *Cell* **153**, 335-347.

Blakeley, P., Fogarty, N. M., 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**, 3613.

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

Boroviak, T., Loos, R., Lombard, P., Okahara, J., Behr, R., Sasaki, E., Nichols, J., Smith, A. and Bertone, P. (2015). Lineage-Specific Profiling Delineates the Emergence and Progression of Naive Pluripotency in Mammalian Embryogenesis. *Dev Cell* **35**, 366-382.

Boroviak, T., Stirparo, G. G., Dietmann, S., Herraez, I. H., 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* **In press**.

Bredenkamp, N., Stirparo, G. G., Smith, A. and Guo, G. (submitted). Cell Surface Marker Sushi Containing Domain 2 Facilitates Establishment Of Human Naïve Pluripotent Stem Cells. *Stem Cell Reports*.

Brons, I. G., Smithers, L. E., Trotter, M. W., 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.

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

Buecker, C., Srinivasan, R., Wu, Z., Calo, E., Acampora, D., Faial, T., Simeone, A., Tan, M., Swigut, T. and Wysocka, J. (2014). Reorganization of enhancer patterns in transition from naive to primed pluripotency. *Cell Stem Cell* **14**, 838-853.

Chal, J., Al Tanoury, Z., Hestin, M., Gobert, B., Aivio, S., Hick, A., Cherrier, T., Nesmith, A. P., Parker, K. K. and Pourquie, O. (2016). Generation of human muscle fibers and satellite-like cells from human pluripotent stem cells in vitro. *Nat Protoc* **11**, 1833-1850.

Chambers, S. M., Fasano, C. A., Papapetrou, E. P., Tomishima, M., Sadelain, M. and Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat Biotechnol* **27**, 275-280.

Chen, B., Dodge, M. E., Tang, W., Lu, J., Ma, Z., Fan, C. W., Wei, S., Hao, W., Kilgore, J., Williams, N. S., et al. (2009). Small molecule-mediated disruption of Wnt-dependent signaling in tissue regeneration and cancer. *Nat Chem Biol* **5**, 100-107.

Chen, G., Gulbranson, D. R., Hou, Z., Bolin, J. M., Ruotti, V., Probasco, M. D., Smuga-Otto, K., Howden, S. E., Diol, N. R., Propson, N. E., et al. (2011). Chemically defined conditions for human iPSC derivation and culture. *Nat Methods* **8**, 424-429.

Chen, Z., Xian, W., Bellin, M., Dorn, T., Tian, Q., Goedel, A., Dreizehnter, L., Schneider, C. M., Ward-van Oostwaard, D., Ng, J. K., et al. (2017). Subtype-specific promoter-

driven action potential imaging for precise disease modelling and drug testing in hiPSC-derived cardiomyocytes. *Eur Heart J* **38**, 292-301.

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.

Enver, T., Pera, M., Peterson, C. and Andrews, P. W. (2009). Stem cell states, fates, and the rules of attraction. *Cell Stem Cell* **4**, 387-397.

Evans, M. J. and Kaufman, M. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-156.

Fiorenzano, A., Pascale, E., D'Aniello, C., Acampora, D., Bassalert, C., Russo, F., Andolfi, G., Biffoni, M., Francescangeli, F., Zeuner, A., et al. (2016). Cripto is essential to capture mouse epiblast stem cell and human embryonic stem cell pluripotency. *Nat Commun* **7**, 12589.

Futschik, M. E. and Carlisle, B. (2005). Noise-robust soft clustering of gene expression time-course data. *J Bioinform Comput Biol* **3**, 965-988.

Gokhale, P. J., Au-Young, J. K., Dadi, S., Keys, D. N., Harrison, N. J., Jones, M., Soneji, S., Enver, T., Sherlock, J. K. and Andrews, P. W. (2015). Culture adaptation alters transcriptional hierarchies among single human embryonic stem cells reflecting altered patterns of differentiation. *PLoS One* **10**, e0123467.

Gong, T. and Szustakowski, J. D. (2013). DeconRNASeq: a statistical framework for deconvolution of heterogeneous tissue samples based on mRNA-Seq data. *Bioinformatics* **29**, 1083-1085.

Guo, G., Huang, Y., Humphreys, P., Wang, X. and Smith, A. (2011). A *PiggyBac*-Based Recessive Screening Method to Identify Pluripotency Regulators. *PLoS One* **6**, e18189.

Guo, G., von Meyenn, F., Rostovskaya, M., Clarke, J., Dietmann, S., Baker, D., Sahakyan, A., Myers, S., Bertone, P., Reik, W., et al. (2017). Epigenetic resetting of human pluripotency. *Development* **144**, 2748-2763.

Guo, G., von Meyenn, F., Santos, F., Chen, Y., Reik, W., Bertone, P., Smith, A. and Nichols, J. (2016). Naive Pluripotent Stem Cells Derived Directly from Isolated Cells of the Human Inner Cell Mass. *Stem Cell Reports* **6**, 437-446.

Hackett, J. A., Kobayashi, T., Dietmann, S. and Surani, M. A. (2017). Activation of Lineage Regulators and Transposable Elements across a Pluripotent Spectrum. *Stem Cell Reports* **8**, 1645-1658.

Hackett, J. A. and Surani, M. A. (2014). Regulatory principles of pluripotency: from the ground state up. *Cell Stem Cell* **15**, 416-430.

Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S. and Saitou, M. (2011). Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* **146**, 519-532.

Hoffman, J. A., Wu, C. I. and Merrill, B. J. (2013). Tcf7l1 prepares epiblast cells in the gastrulating mouse embryo for lineage specification. *Development* **140**, 1665-1675.

Hough, S. R., Thornton, M., Mason, E., Mar, J. C., Wells, C. A. and Pera, M. F. (2014). Single-cell gene expression profiles define self-renewing, pluripotent, and lineage primed states of human pluripotent stem cells. *Stem Cell Reports* **2**, 881-895.

Huang, S. M., Mishina, Y. M., Liu, S., Cheung, A., Stegmeier, F., Michaud, G. A., Charlat, O., Wiellette, E., Zhang, Y., Wiessner, S., et al. (2009). Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. *Nature* **461**, 614-620.

Huelsken, J., Vogel, R., Brinkmann, V., Erdmann, B., Birchmeier, C. and Birchmeier, W. (2000). Requirement for beta-catenin in anterior-posterior axis formation in mice. *J Cell Biol* **148**, 567-578.

Kalkan, T., Olova, N., Roode, M., Mulas, C., Lee, H. J., Nett, I., Marks, H., Walker, R., Stunnenberg, H. G., Lilley, K. S., et al. (2017). Tracking the embryonic stem cell transition from ground state pluripotency. *Development* **144**, 1221-1234.

Kalkan, T. and Smith, A. (2014). Mapping the route from naive pluripotency to lineage specification. *Phil Trans R Soc B* **369**.

Kharchenko, P. V., Silberstein, L. and Scadden, D. T. (2014). Bayesian approach to single-cell differential expression analysis. *Nat Methods* **11**, 740-742.

Kilens, S., Meistermann, D., Moreno, D., Chariau, C., Gaignerie, A., Reignier, A., Lelièvre, Y., Casanova, M., Vallot, C., Nedellec, S., et al. (2018). Parallel derivation of isogenic human primed and naive induced pluripotent stem cells. *Nature Communications* **9**, 360.

Kojima, Y., Kaufman-Francis, K., Studdert, J. B., Steiner, K. A., Power, M. D., Loebel, D. A., Jones, V., Hor, A., de Alencastro, G., Logan, G. J., et al. (2014). The transcriptional and functional properties of mouse epiblast stem cells resemble the anterior primitive streak. *Cell Stem Cell* **14**, 107-120.

Kumar, L. and E.M., F. (2007). Mfuzz: a software package for soft clustering of microarray data. *Bioinformation* **2**, 5-7.

Lê, S., Josse, J. and Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. *2008* **25**, 18.

Lehtio, L., Chi, N. W. and Krauss, S. (2013). Tankyrases as drug targets. *FEBS J* **280**, 3576-3593.

Liu, P., Wakamiya, M., Shea, M. J., Albrecht, U., Behringer, R. R. and Bradley, A. (1999). Requirement for Wnt3 in vertebrate axis formation. *Nat Genet* **22**, 361-365.

Liu, X., Nefzger, C. M., Rossello, F. J., Chen, J., Knaupp, A. S., Firas, J., Ford, E., Pflueger, J., Paynter, J. M., Chy, H. S., et al. (2017). Comprehensive characterization of distinct states of human naive pluripotency generated by reprogramming. *Nature Methods* **14**, 1055-1062.

Loh, K. M., Ang, L. T., Zhang, J., Kumar, V., Ang, J., Auyeong, J. Q., Lee, K. L., Choo, S. H., Lim, C. Y., Nichane, M., et al. (2014). Efficient endoderm induction from human pluripotent stem cells by logically directing signals controlling lineage bifurcations. *Cell Stem Cell* **14**, 237-252.

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.

Martello, G., Sugimoto, T., Diamanti, E., Joshi, A., Hannah, R., Ohtsuka, S., Gottgens, B., Niwa, H. and Smith, A. (2012). Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. *Cell Stem Cell* **11**, 491-504.

Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci. USA* **78**, 7634-7638.

Mohammed, H., Hernando-Herraez, I., Savino, A., Scialdone, A., Macaulay, I., Mulas, C., Chandra, T., Voet, T., Dean, W., Nichols, J., et al. (2017). Single-Cell Landscape of Transcriptional Heterogeneity and Cell Fate Decisions during Mouse Early Gastrulation. *Cell Rep* **20**, 1215-1228.

Moretti, A., Bellin, M., Welling, A., Jung, C. B., Lam, J. T., Bott-Flugel, L., Dorn, T., Goedel, A., Hohnke, C., Hofmann, F., et al. (2010). Patient-specific induced pluripotent stem-cell models for long-QT syndrome. *N Engl J Med* **363**, 1397-1409.

Morgani, S., Nichols, J. and Hadjantonakis, A. K. (2017). The many faces of Pluripotency: in vitro adaptations of a continuum of in vivo states. *BMC Dev Biol* **17**, 7.

Morkel, M., Huelsken, J., Wakamiya, M., Ding, J., van de Wetering, M., Clevers, H., Taketo, M. M., Behringer, R. R., Shen, M. M. and Birchmeier, W. (2003). Beta-catenin regulates Cripto- and Wnt3-dependent gene expression programs in mouse axis and mesoderm formation. *Development* **130**, 6283-6294.

Mulas, C., Kalkan, T. and Smith, A. (2017). NODAL Secures Pluripotency upon Embryonic Stem Cell Progression from the Ground State. *Stem Cell Reports* **9**, 77-91.

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.

Nichols, J. and Smith, A. (2009). Naive and primed pluripotent states. *Cell Stem Cell* **4**, 487-492.

Peng, G., Suo, S., Chen, J., Chen, W., Liu, C., Yu, F., Wang, R., Chen, S., Sun, N., Cui, G., et al. (2016). Spatial Transcriptome for the Molecular Annotation of Lineage Fates and Cell Identity in Mid-gastrula Mouse Embryo. *Dev Cell* **36**, 681-697.

Pereira, L., Yi, F. and Merrill, B. J. (2006). Repression of nanog gene transcription by tcf3 limits embryonic stem cell self-renewal. *Mol Cell Biol* **26**, 7479-7491.

Plusa, B., Piliszek, A., Frankenberg, S., Artus, J. and Hadjantonakis, A. K. (2008). Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. *Development* **135**, 3081-3091.

Proffitt, K. D., Madan, B., Ke, Z., Pendharkar, V., Ding, L., Lee, M. A., Hannoush, R. N. and Virshup, D. M. (2013). Pharmacological inhibition of the Wnt acyltransferase PORCN prevents growth of WNT-driven mammary cancer. *Cancer Res* **73**, 502-507.

Rezania, A., Bruin, J. E., Arora, P., Rubin, A., Batushansky, I., Asadi, A., O'Dwyer, S., Quiskamp, N., Mojibian, M., Albrecht, T., et al. (2014). Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. *Nat Biotechnol* **32**, 1121-1133.

Rossant, J. (2015). Mouse and human blastocyst-derived stem cells: vive les differences. *Development* **142**, 9-12.

Rossant, J. and Tam, P. P. L. (2017). New Insights into Early Human Development: Lessons for Stem Cell Derivation and Differentiation. *Cell Stem Cell* **20**, 18-28.

Sheng, G. (2015). Epiblast morphogenesis before gastrulation. *Developmental Biology* **401**, 17-24.

Smith, A. (2013). Nanog heterogeneity: tilting at windmills? *Cell Stem Cell* **13**, 6-7.

Smith, A. (2017). Formative pluripotency: the executive phase in a developmental continuum. *Development* **144**, 365-373.

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**.

Sumi, T., Oki, S., Kitajima, K. and Meno, C. (2013). Epiblast ground state is controlled by canonical Wnt/beta-catenin signaling in the postimplantation mouse embryo and epiblast stem cells. *PLoS One* **8**, e63378.

Takahashi, K. and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676.

Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficz, G., Krueger, F., Oxley, D., Santos, F., Clarke, J., Mansfield, W., et al. (2014). Resetting transcription factor control circuitry toward ground-state pluripotency in human. *Cell* **158**, 1254-1269.

Tesar, P. J. (2005). Derivation of germ-line-competent embryonic stem cell lines from preblastocyst mouse embryos. *Proc Natl Acad Sci U S A* **102**, 8239-8244.

Theunissen, T. W., Friedli, M., He, Y., Planet, E., O'Neil, R. C., Markoulaki, S., Pontis, J., Wang, H., Iouranova, A., Imbeault, M., et al. (2016). Molecular Criteria for Defining the Naive Human Pluripotent State. *Cell Stem Cell* **19**, 502-515.

Theunissen, T. W., Powell, B. E., Wang, H., Mitalipova, M., Faddah, D. A., Reddy, J., Fan, Z. P., Maetzel, D., Ganz, K., Shi, L., et al. (2014). Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. *Cell Stem Cell* **15**, 471-487.

Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S. and Jones, J. M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145-1147.

Tsakiridis, A., Huang, Y., Blin, G., Skylaki, S., Wymeersch, F., Osorno, R., Economou, C., Karagianni, E., Zhao, S., Lowell, S., et al. (2014). Distinct Wnt-driven primitive streak-like populations reflect in vivo lineage precursors. *Development* **141**, 1209-1221.

Wray, J., Kalkan, T., Gomez-Lopez, S., Eckardt, D., Cook, A., Kemler, R. and Smith, A. (2011). Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. *Nat Cell Biol* **13**, 838-845.

Yamaguchi, T. P., Takada, S., Yoshikawa, Y., Wu, N. and McMahon, A. P. (1999). T (Brachyury) is a direct target of Wnt3a during paraxial mesoderm specification. *Genes Dev* **13**, 3185-3190.

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.

Yates, A., Akanni, W., Amode, M. R., Barrell, D., Billis, K., Carvalho-Silva, D., Cummins, C., Clapham, P., Fitzgerald, S., Gil, L., et al. (2016). Ensembl 2016. *Nucleic Acids Res* **44**, D710-716.

Zhou, W., Choi, M., Margineantu, D., Margaretha, L., Hesson, J., Cavanaugh, C., Blau, C. A., Horwitz, M. S., Hockenbery, D., Ware, C., et al. (2012). HIF1alpha induced switch from bivalent to exclusively glycolytic metabolism during ESC-to-EpiSC/hESC transition. *EMBO J* **31**, 2103-2116.

Zimmerlin, L., Park, T. S., Huo, J. S., Verma, K., Pather, S. R., Talbot, C. C., Agarwal, J., Steppan, D., Zhang, Y. W., Considine, M., et al. (2016). Tankyrase inhibition promotes a stable human naïve pluripotent state with improved functionality. *Development* **143**, 4368-4380.

FIGURE LEGENDS

Figure 1. Naïve hPSC do not respond to lineage differentiation cues
(A) Aggregates cultured for 7 days in N2B27, unstained and stained with Trypan Blue; bright field, 10x. (B) RT-qPCR analysis of marker expression after 14 days of aggregation culture. Here and in all qPCR measurements, error bars are standard deviations of technical duplicates. (C) Flow cytometry analysis of intracellular marker staining after application of dual SMAD inhibition for neuroectoderm induction. (D) RT-qPCR analysis following dual SMAD inhibition. (E) Flow cytometry analysis of cell surface (CXCR4) and intracellular (SOX17) marker expression after application of definitive endoderm induction conditions. (F) RT-qPCR analysis in definitive endoderm induction conditions. (G) Cell counts in paraxial mesoderm induction conditions. Error bars are derived from 2 independent experiments. (H) RT-qPCR analysis following application of paraxial mesoderm induction protocol.
Un – undifferentiated, EB – embryoid bodies.
NE – neuroectoderm, DE – definitive endoderm, PM – paraxial mesoderm.

Figure 2. Transition of naive hPSC in N2B27 only
(A) Bright field images of naïve hPSC and cultures in N2B27 for 7 or 12 days, or in E8 for 14 days; 10x. (B) Colony counts after culture in N2B27 or E8 and replating at low density in naïve PSC conditions. (C) RT-qPCR analysis of pluripotency marker expression during culture in N2B27 or E8. (D) Flow cytometry analysis of neuroectoderm and definitive endoderm markers following induction after 7 days in N2B27 or E8. (E) Summary of efficiencies of differentiation after 7 days in N2B27 or E8, from multiple independent experiments. (F) RT-qPCR analysis of mesoendodermal marker expression after 7 days in N2B27 or E8. (G) Immunostaining for TBXT and NANOG on day 7 in the indicated conditions.

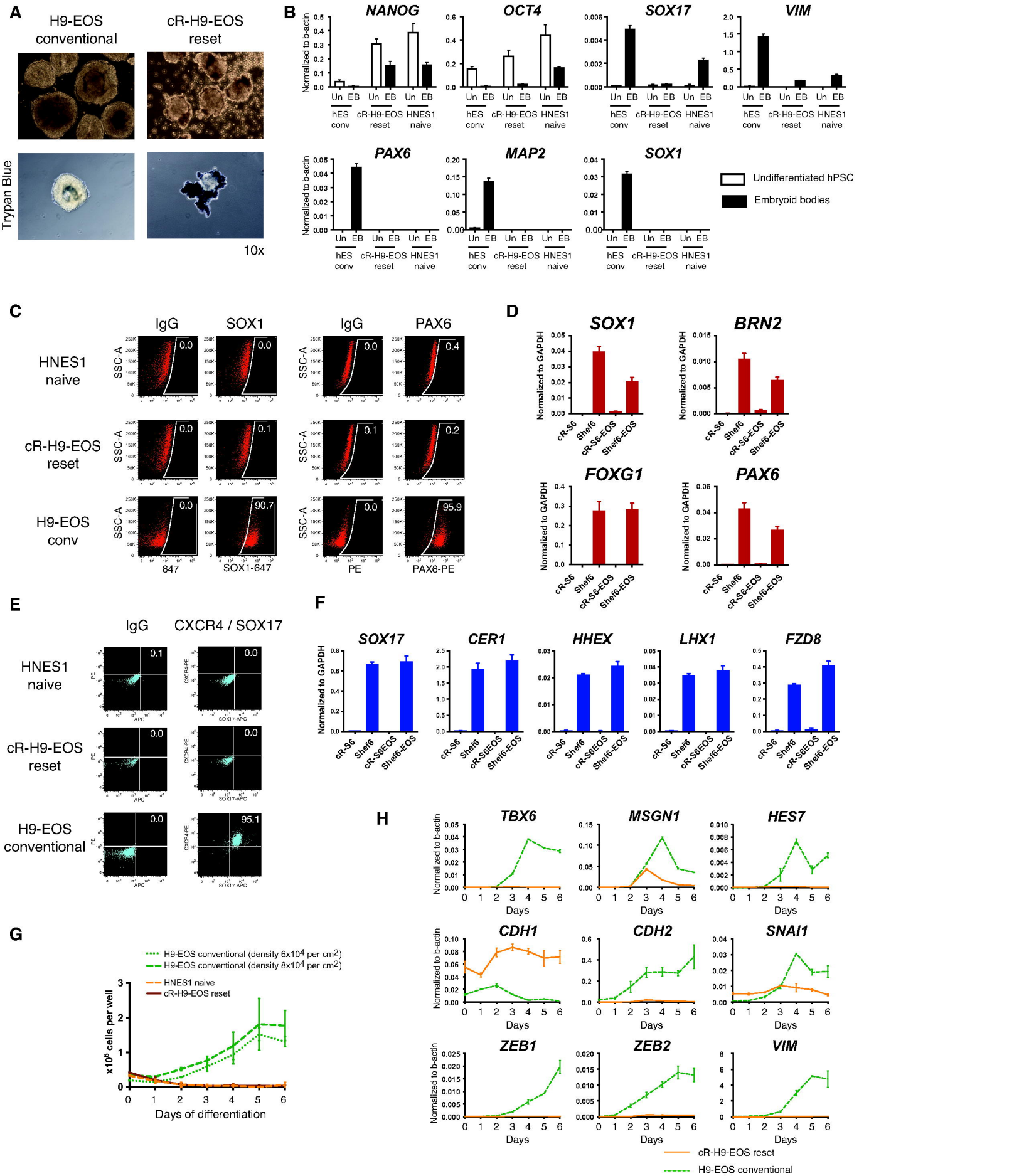
Figure 3. Transition of naïve hPSC is facilitated by XAV939
(A) Conventional hPSC and naive hPSC before and after culture in N2B27 plus XAV939 for 14 days; bright field, 10x. (B) Cell counts during transition in XAV939; error bars are standard deviations from 3 independent experiments. (C) Colony assay after culture in XAV939; performed as in Fig 2B. (D) RT-qPCR assay of pluripotency gene expression during transition in XAV939. (E) Immunofluorescent staining for pluripotency markers during transition in XAV939. H9 conventional hPSC provide reference staining. (F) Quantification of intensity of immunostaining. Mean intensity of cells stained with the secondary antibodies only was subtracted from measurement of each cell stained with specific antibodies. The resulted values of intensity above control are represented as boxplots.

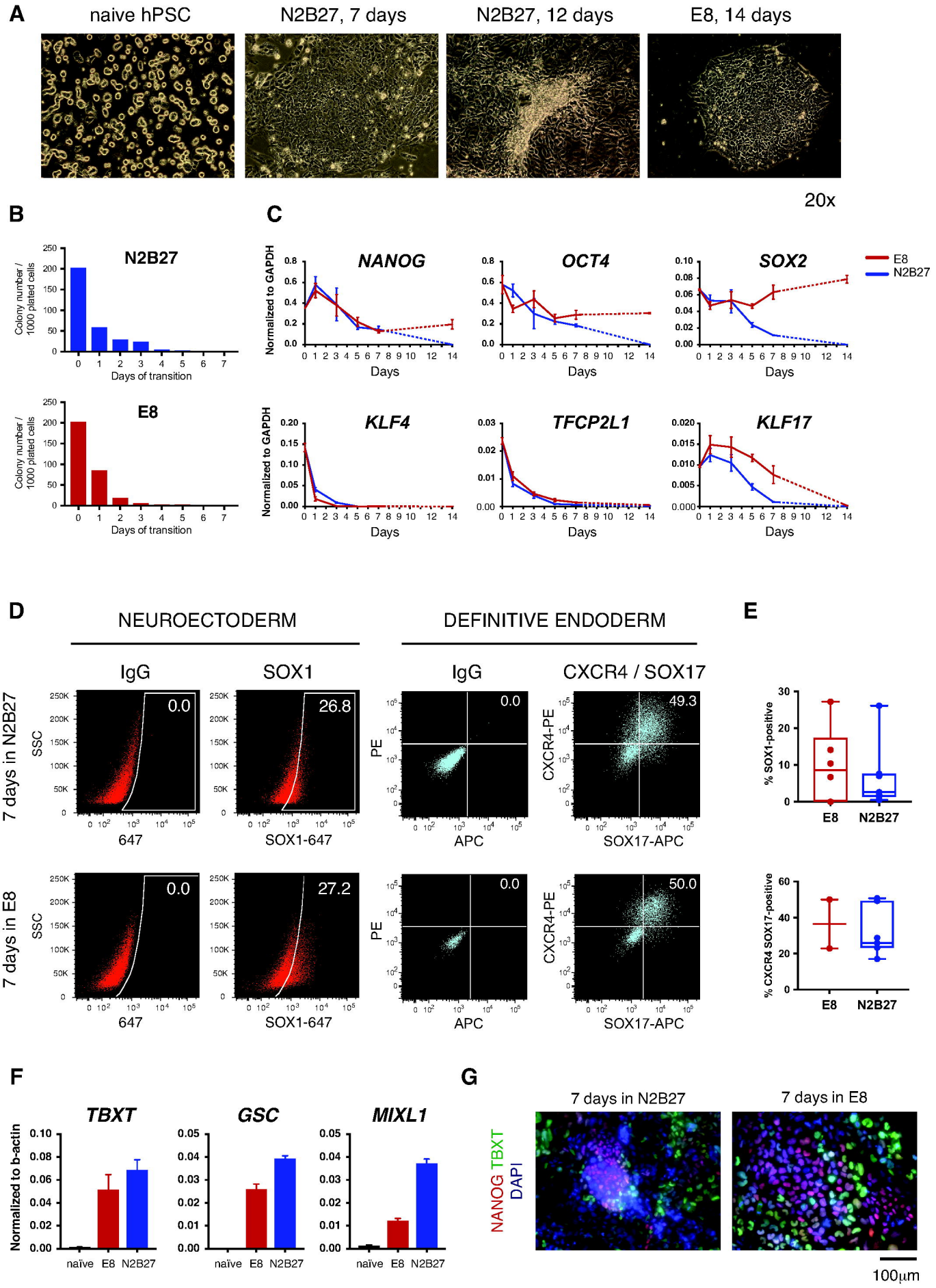
Figure 4. Differentiation competence of hPSC after formative transition in XAV939
(A-C) Neuroectoderm induction after transition in XAV939, parallel treatment of conventional hPSC and naïve hPSC after capacitation. Neuroectoderm markers were examined by: (A) flow cytometry; (B) immunostaining, (C) RT-qPCR. (D-F) Definitive endoderm induction after transition in XAV939. Markers were examined by: (D) immunostaining; (E) flow cytometry; (F) RT-qPCR. (G-H) Differentiation to paraxial mesoderm after transition in XAV939. Markers were analysed by: (G) RT-qPCR; (H) immunostaining for TBX6 and CDX2.

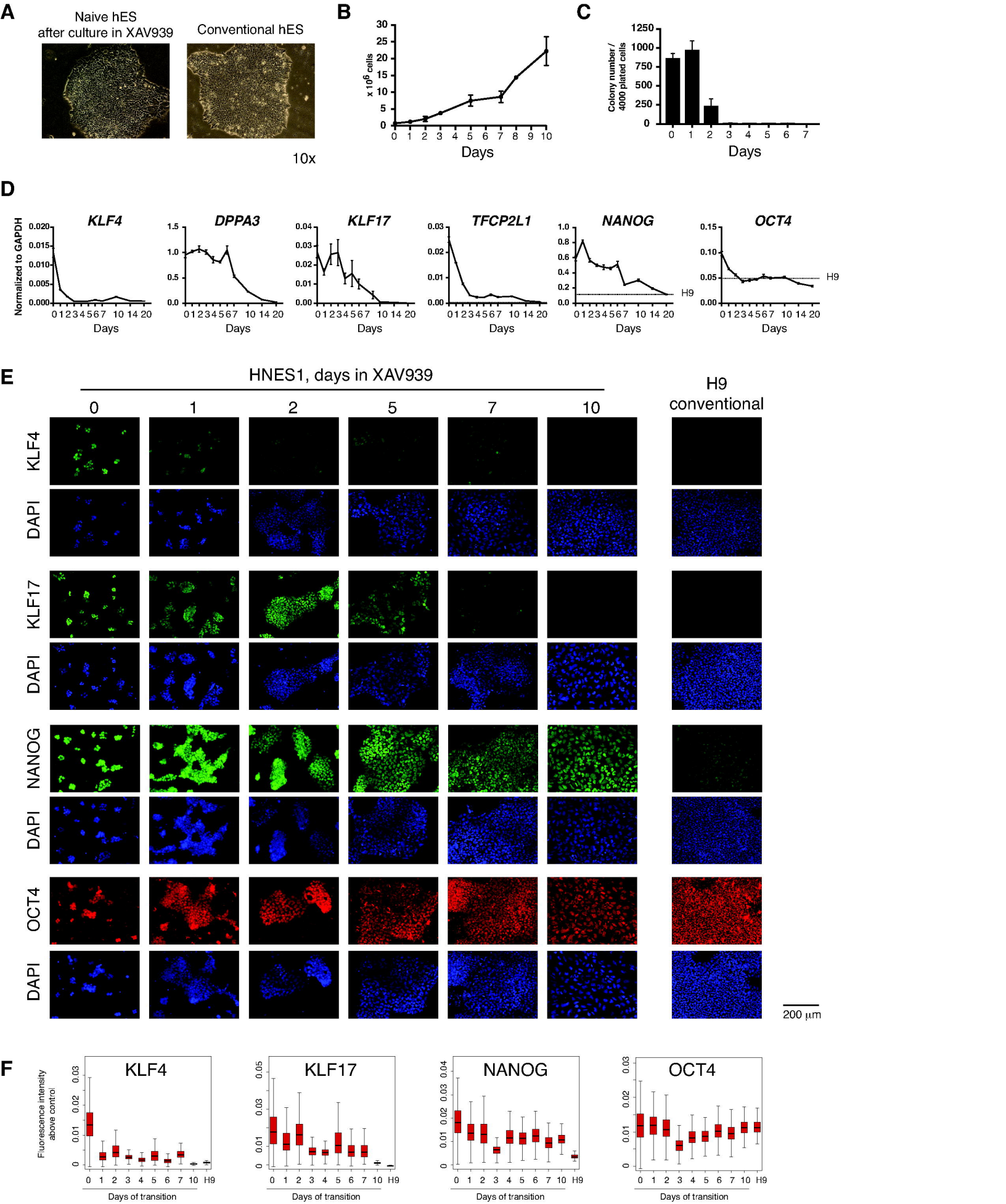
NE – neuroectoderm, DE – definitive endoderm, PM – paraxial mesoderm, Un – undifferentiated (conventional and capacitated hPSC that were not induced to differentiation).

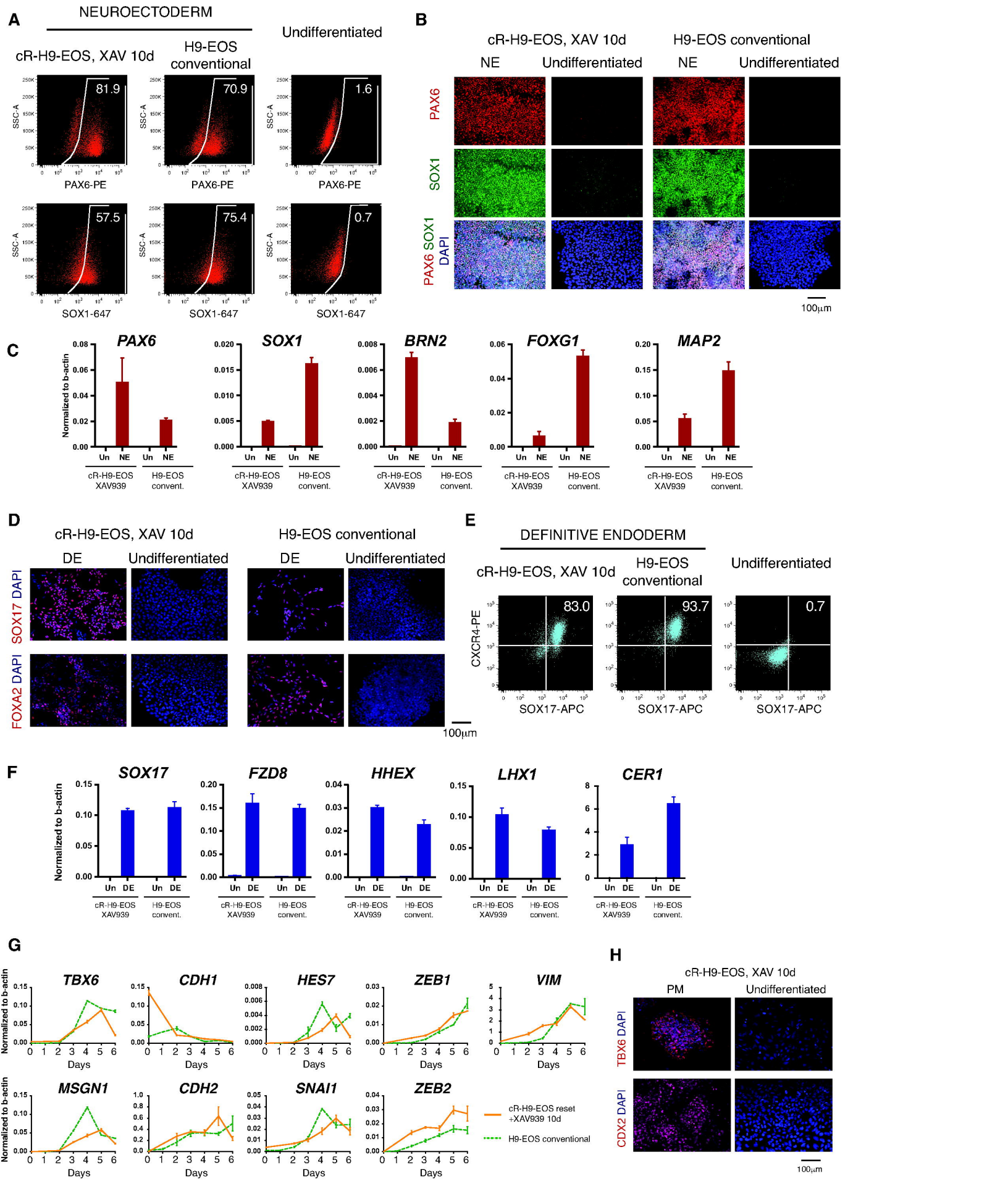
Figure 5. Global transcriptome analysis during formative transition
(A) PCA of HNES1 and cR-H9EOS during formative transition. For each cell line three independent experiments are represented. (B) Heatmaps showing expression of variable genes assigned to 5 dynamic clusters, derived separately for HNES1 and cR-H9-EOS. (C) Numbers of protein coding genes, and of TF/coF/REM (transcription factors, co-factors and epigenetic remodellers), within the 5 dynamic clusters. (D) Comparison of gene content of clusters between HNES1 and cR-H9-EOS; 61.7% of genes belong to the same clusters and 83.9% belong to the same+similar clusters. (E) Heatmap of RNAseq expression values for selected genes during formative transition.

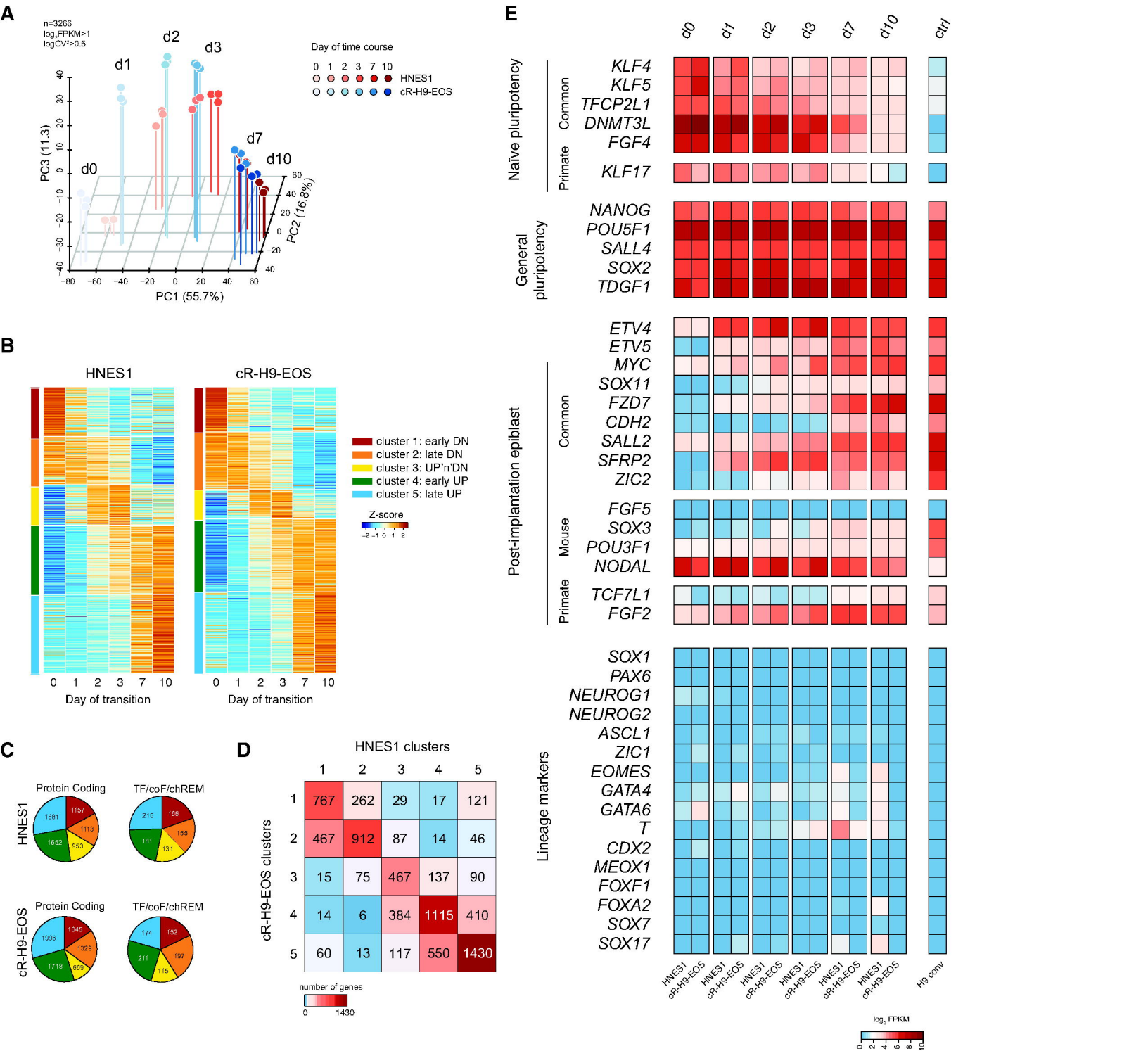
Figure 6. Comparison of gene expression during formative transition of hPSC *in vitro* and *M. fascicularis* embryonic epiblast *in utero*.
(A) Soft clustering analysis of variable genes during developmental progression of mouse and macaque epiblast *in utero*, and transition of hPSC *in vitro*; average level of gene expression per cluster is indicated. (B) Comparison of clusters of variable genes between progression of *M. musculus* or *M. fascicularis* epiblast *in utero* and hPSC *in vitro*. (C) Expression of selected genes during progression of mouse and macaque epiblast *in utero* and hPSC *in vitro*. (D) PCA of macaque epiblast single cells at different stages of development *in utero* and hPSC populations during formative transition *in vitro*. (E-F) fractions of similarity of hPSC during formative transition to embryonic stages EPI, post-E and post-L of the macaque embryo: (E) 3D plot; (F) 2D projections.

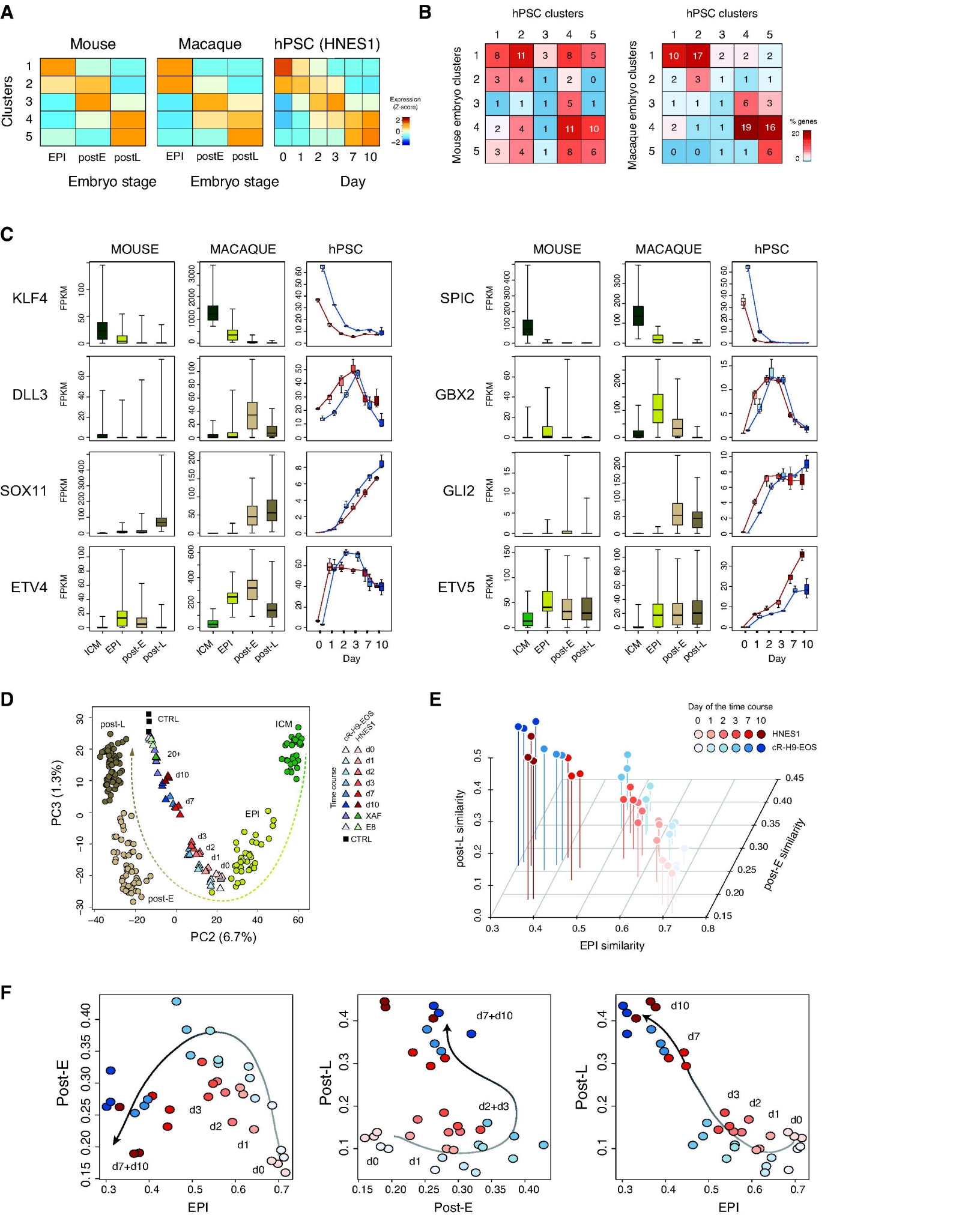

A
H9-EOS conventional
cR-H9-EOS reset
Trypan Blue
10x
B
NANOG
OCT4
SOX17
VIM
PAX6
MAP2
SOX1
Normalized to b-actin
Un EB
hES conv
cR-H9-EOS reset
HNES1 naive
Undifferentiated hPSC
Embryoid bodies
C
IgG
SOX1
PAX6
HNES1 naive
cR-H9-EOS reset
H9-EOS conv
SSC-A
647
SOX1-647
PE
PAX6-PE
D
SOX1
BRN2
FOXG1
PAX6
Normalized to GAPDH
cR-S6
Shef6
cR-S6-EOS
Shef6-EOS
E
IgG
CXCR4 / SOX17
HNES1 naive
cR-H9-EOS reset
H9-EOS conventional
F
SOX17
CER1
HHEX
LHX1
FZD8
G
H9-EOS conventional (density 6x10^4 per cm^2)
H9-EOS conventional (density 8x10^4 per cm^2)
HNES1 naive
cR-H9-EOS reset
x10^6 cells per well
Days of differentiation
H
TBX6
MSGN1
HES7
CDH1
CDH2
SNAI1
ZEB1
ZEB2
VIM
Days
cR-H9-EOS reset
H9-EOS conventional


A
naive hPSC
N2B27, 7 days
N2B27, 12 days
E8, 14 days
20x
B
N2B27
E8
Colony number / 1000 plated cells
Days of transition
C
NANOG
OCT4
SOX2
KLF4
TFCP2L1
KLF17
Normalized to GAPDH
Days
E8
N2B27
D
NEUROECTODERM
DEFINITIVE ENDODERM
IgG
SOX1
IgG
CXCR4 / SOX17
7 days in N2B27
7 days in E8
SSC
647
SOX1-647
PE
APC
CXCR4-PE
SOX17-APC
0.0
26.8
49.3
27.2
50.0
E
% SOX1-positive
% CXCR4 SOX17-positive
E8
N2B27
F
TBXT
GSC
MIXL1
Normalized to b-actin
naïve
E8
N2B27
G
7 days in N2B27
7 days in E8
NANOG TBXT
DAPI
100μm


A
Naive hES
after culture in XAV939
Conventional hES
10x
B
x 10^6 cells
Days
C
Colony number / 4000 plated cells
Days
D
KLF4
DPPA3
KLF17
TFCP2L1
NANOG
OCT4
Normalized to GAPDH
H9
Days
E
HNES1, days in XAV939
0
1
2
5
7
10
H9
conventional
KLF4
DAPI
KLF17
DAPI
NANOG
DAPI
OCT4
DAPI
200 µm
F
KLF4
KLF17
NANOG
OCT4
Fluorescence intensity above control
Days of transition


A
NEUROECTODERM
Undifferentiated
cR-H9-EOS, XAV 10d
H9-EOS conventional
81.9
70.9
1.6
57.5
75.4
0.7
SSC-A
PAX6-PE
SOX1-647
B
cR-H9-EOS, XAV 10d
H9-EOS conventional
NE
Undifferentiated
PAX6
SOX1
PAX6 SOX1 DAPI
100μm
C
PAX6
SOX1
BRN2
FOXG1
MAP2
Normalized to b-actin
Un
NE
cR-H9-EOS XAV939
H9-EOS convent.
D
cR-H9-EOS, XAV 10d
H9-EOS conventional
DE
Undifferentiated
SOX17 DAPI
FOXA2 DAPI
100μm
E
DEFINITIVE ENDODERM
cR-H9-EOS, XAV 10d
H9-EOS conventional
Undifferentiated
83.0
93.7
0.7
CXCR4-PE
SOX17-APC
F
SOX17
FZD8
HHEX
LHX1
CER1
Normalized to b-actin
Un
DE
cR-H9-EOS XAV939
H9-EOS convent.
G
TBX6
CDH1
HES7
ZEB1
VIM
MSGN1
CDH2
SNAI1
ZEB2
Normalized to b-actin
Days
cR-H9-EOS reset +XAV939 10d
H9-EOS conventional
H
cR-H9-EOS, XAV 10d
PM
Undifferentiated
TBX6 DAPI
CDX2 DAPI
100μm


A
n=3266
log2FPKM>1
logCV2>0.5
d0
d1
d2
d3
d7
d10
PC1 (55.7%)
PC2 (16.8%)
PC3 (11.3)
Day of time course
0 1 2 3 7 10
HNES1
cR-H9-EOS
B
HNES1
cR-H9-EOS
cluster 1: early DN
cluster 2: late DN
cluster 3: UP'n'DN
cluster 4: early UP
cluster 5: late UP
Z-score
Day of transition
C
HNES1
cR-H9-EOS
Protein Coding
TF/coF/chREM
D
HNES1 clusters
cR-H9-EOS clusters
number of genes
E
d0
d1
d2
d3
d7
d10
ctrl
Naïve pluripotency
Common
Primate
KLF4
KLF5
TFCP2L1
DNMT3L
FGF4
KLF17
General pluripotency
NANOG
POU5F1
SALL4
SOX2
TDGF1
Post-implantation epiblast
Common
ETV4
ETV5
MYC
SOX11
FZD7
CDH2
SALL2
SFRP2
ZIC2
Mouse
FGF5
SOX3
POU3F1
NODAL
Primate
TCF7L1
FGF2
Lineage markers
SOX1
PAX6
NEUROG1
NEUROG2
ASCL1
ZIC1
EOMES
GATA4
GATA6
T
CDX2
MEOX1
FOXF1
FOXA2
SOX7
SOX17
HNES1
cR-H9-EOS
H9 conv
log2 FPKM


A
Mouse
Macaque
hPSC (HNES1)
Clusters
EPI postE postL
Embryo stage
0 1 2 3 7 10
Day
Expression (Z-score)
B
hPSC clusters
Mouse embryo clusters
Macaque embryo clusters
% genes
C
MOUSE
MACAQUE
hPSC
KLF4
DLL3
SOX11
ETV4
SPIC
GBX2
GLI2
ETV5
FPKM
ICM EPI post-E post-L
D
PC3 (1.3%)
PC2 (6.7%)
post-L
CTRL
ICM
20+
d10
d7
EPI
d3
d2
d1
d0
post-E
cR-H9-EOS
HNES1
Time course
XAF
E8
E
Day of the time course
HNES1
cR-H9-EOS
post-L similarity
post-E similarity
EPI similarity
F
Post-E
EPI
d7+d10
d3
d2
d1
d0
Post-L
Post-E
d2+d3
Post-L
EPI
d10
d7
d3
d2
d1
d0