Modeling human extraembryonic mesoderm cells using naive pluripotent stem cells

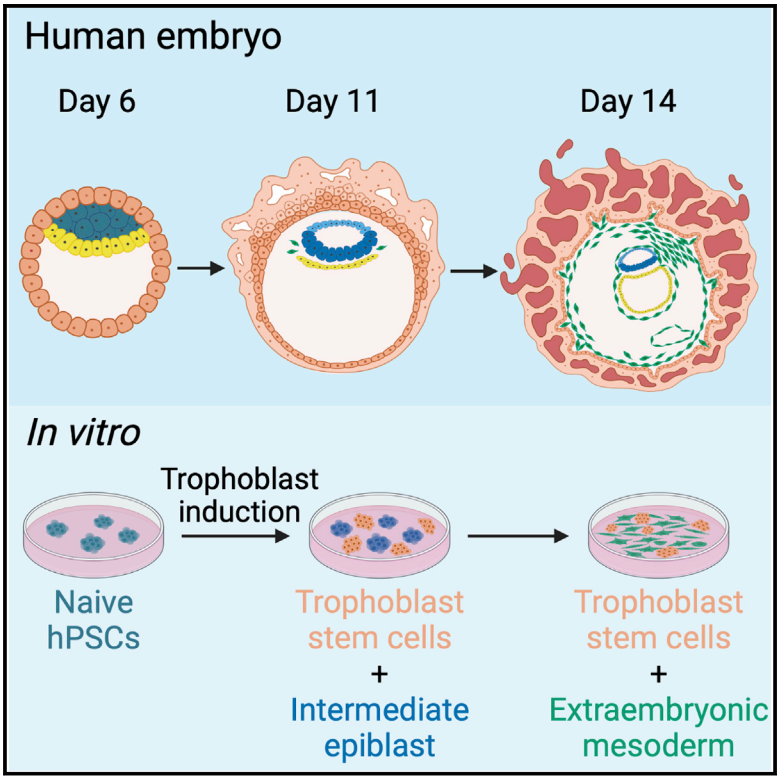
Thi Xuan Ai Pham, Amitesh Panda, Harunobu Kagawa, San Kit To, Cankat Ertekin, Grigorios Georgolopoulos, Sam S.F.A. van Knippenberg, Ryan Nicolaas Allsop, Alexandre Bruneau, Jonathan Sai-Hong Chui, et al.

Cell Stem Cell

Article

Modeling human extraembryonic mesoderm cells using naive pluripotent stem cells

Graphical abstract



Highlights

- Human naive to trophoblast or primitive endoderm conversion induces EXMCs
- EXMCs represent primate postimplantation extraembryonic mesoderm
- Differentiation of EXMCs from an epiblast intermediate
- EXMCs are the main off-target cells specified during human blastoid generation

Authors

Thi Xuan Ai Pham, Amitesh Panda, Harunobu Kagawa, ..., Nicolas Rivron, Bradley Philip Balaton, Vincent Pasque

Correspondence

bradley.balaton@kuleuven.be (B.P.B.), vincent.pasque@kuleuven.be (V.P.)

In brief

Little is known about human embryos just after implantation. Pasque et al. established a human pluripotent stem cell-based model that recapitulates primate-specific differentiation into extraembryonic mesoderm cells, which represent a little-understood but major cell type thought to develop into several important lineages in the embryo.

Pham et al., 2022, Cell Stem Cell *29*, 1346–1365
September 1, 2022
https://doi.org/10.1016/j.stem.2022.08.001

CellPress

CellPress
OPEN ACCESS

Cell Stem Cell

Article

Modeling human extraembryonic mesoderm cells using naive pluripotent stem cells

Thi Xuan Ai Pham,[1,10] Amitesh Panda,[1,10] Harunobu Kagawa,[2] San Kit To,[1] Cankat Ertekin,[1] Grigorios Georgolopoulos,[1,6] Sam S.F.A. van Knippenberg,[1] Ryan Nicolaas Allsop,[1] Alexandre Bruneau,[3] Jonathan Sai-Hong Chui,[1] Lotte Vanheer,[1] Adrian Janiszewski,[1] Joel Chappell,[1,7] Michael Oberhuemer,[1,8] Raissa Songwa Tchinda,[1,9] Irene Talon,[1] Sherif Khodeer,[1] Janet Rossant,[4] Frederic Lluis,[1] Laurent David,[3,5] Nicolas Rivron,[2] Bradley Philip Balaton,[1,*] and Vincent Pasque[1,11,*]

[1]Department of Development and Regeneration, Leuven Stem Cell Institute, Leuven Institute for Single-cell Omics (LISCO), KU Leuven-University of Leuven, 3000 Leuven, Belgium
[2]Institute of Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), Vienna BioCenter (VBC), 1030 Vienna, Austria
[3]Nantes Université, CHU Nantes, Inserm, CR2TI, UMR 1064, F-44000, Nantes, France
[4]Program in Developmental and Stem Cell Biology, Hospital for Sick Children, Toronto, ON M5V 0B1, Canada
[5]Nantes Université, CHU Nantes, Inserm, CNRS, BioCore, F-44000 Nantes, France
[6]Present address: Genevia Technologies Oy, 33100 Tampere, Finland
[7]Present address: Bit.bio, Dorothy Hodgkin Building Babraham Research, Cambridge CB22 3FH, United Kingdom
[8]Present address: Max Perutz Labs Vienna, Vienna Biocenter (VBC), University of Vienna, 1030 Vienna, Austria
[9]Present address: Department of Biological Sciences, University of Toledo, Toledo, OH 43606, USA
[10]These authors contributed equally
[11]Lead contact
*Correspondence: bradley.balaton@kuleuven.be (B.P.B.), vincent.pasque@kuleuven.be (V.P.)
https://doi.org/10.1016/j.stem.2022.08.001

SUMMARY

A hallmark of primate postimplantation embryogenesis is the specification of extraembryonic mesoderm (EXM) before gastrulation, in contrast to rodents where this tissue is formed only after gastrulation. Here, we discover that naive human pluripotent stem cells (hPSCs) are competent to differentiate into EXM cells (EXMCs). EXMCs are specified by inhibition of Nodal signaling and GSK3B, are maintained by mTOR and BMP4 signaling activity, and their transcriptome and epigenome closely resemble that of human and monkey embryo EXM. EXMCs are mesenchymal, can arise from an epiblast intermediate, and are capable of self-renewal. Thus, EXMCs arising via primate-specific specification between implantation and gastrulation can be modeled *in vitro*. We also find that most of the rare off-target cells within human blastoids formed by triple inhibition (Kagawa et al., 2021) correspond to EXMCs. Our study impacts our ability to model and study the molecular mechanisms of early human embryogenesis and related defects.

INTRODUCTION

The extraembryonic mesoderm (EXM) is an important tissue with essential roles in development. EXM is implicated in primitive erythropoiesis and extracellular matrix formation; becomes an integral part of the amnion, yolk sac, allantois, and chorion; and forms the primitive umbilical cord (Enders and King, 1988; Luckett, 1978; Nahaboo et al., 2022; Ross and Boroviak, 2020; Sadler, 2012; Saykali et al., 2019; Shepard, 1989; Spencer Chapman et al., 2021). Intriguingly, the mechanisms of EXM specification differ dramatically between species of mammals. In rodents, the EXM is specified only after gastrulation from the primitive streak (Arnold and Robertson, 2009; Saykali et al., 2019; Tam and Beddington, 1987). In primates however, the EXM starts to emerge earlier, before gastrulation, around Carnegie stage 5 in humans (Cui et al., 2022; Enders and King, 1988; Gasser, 1975; Kinder et al., 1999; Luckett, 1978; Ross and Boroviak, 2020; Shepard, 1989). The EXM arises in close proximity to the epiblast, primitive endoderm (PrE), and trophoblast (TB) (Enders and King, 1988; Ross and Boroviak, 2020), then spreads to line the inner surface of the cytotrophoblast (CTB) and the outer surface of the primitive yolk sac and amnion. It forms a connecting stalk between the CTB and the amnion, epiblast disc, and PrE by day 13, which forms the primitive umbilical cord (Sadler, 2012). Later on, EXMCs fill chorionic villi (Enders and King, 1988; Hertig et al., 1956; O'Rahilly and Müller, 1987; Ross and Boroviak, 2020; Rossant and Tam, 2022; Spencer Chapman et al., 2021; Yang et al., 2021). However, despite the importance of the EXM, our understanding of its cellular and molecular regulation in humans remains limited. Moreover, there are no *in vitro* models for primate EXM development (Ross and Boroviak, 2020). An exciting prospect is the derivation of human *in vitro* extraembryonic mesoderm cells (EXMCs) to model EXM development.

The lineage origin of the EXM in humans and other primates is unknown and subject to considerable uncertainty, with multiple

CellPress
OPEN ACCESS

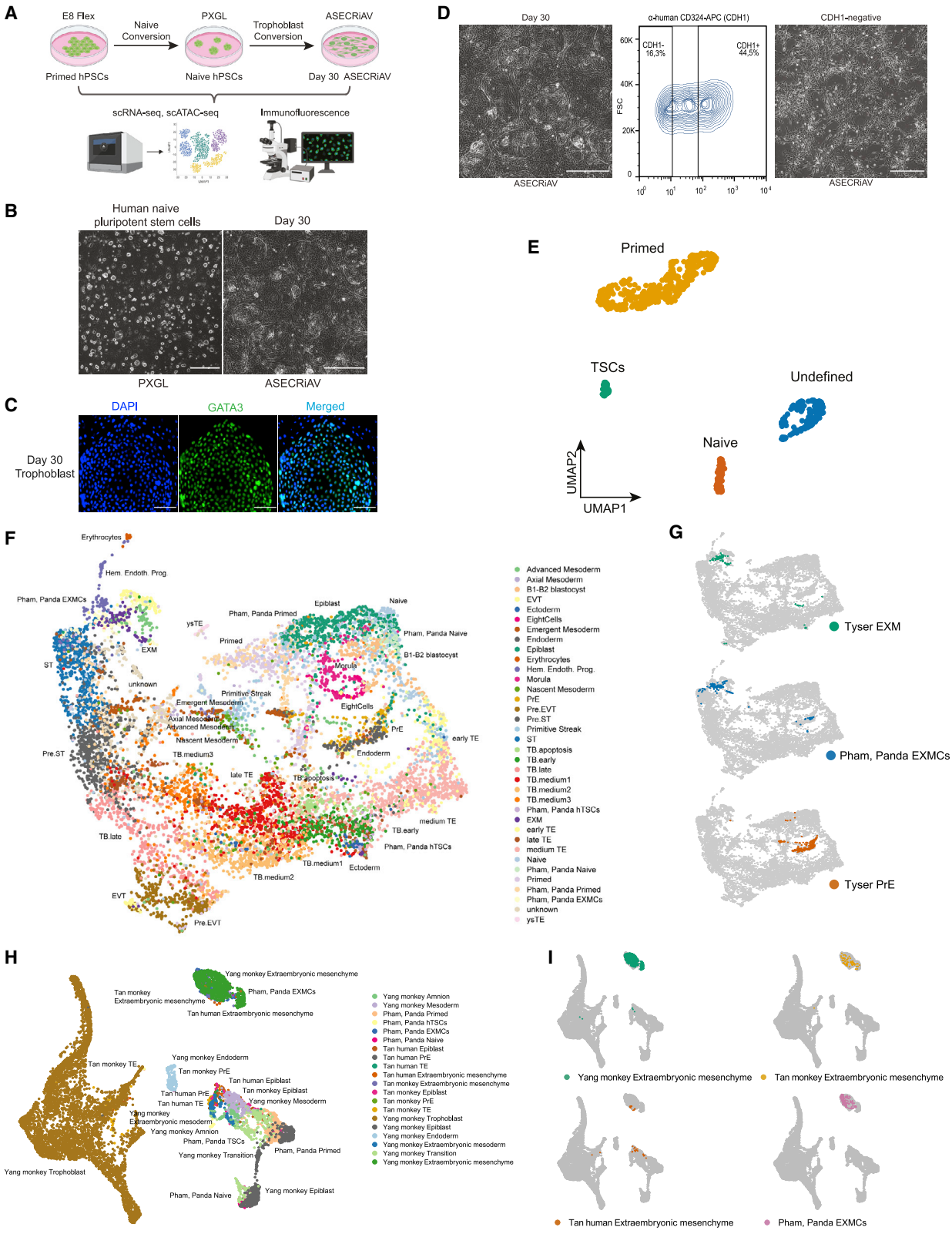


Figure 1. Derivation of EXMCs from naive hPSCs

(A) Experimental strategy. Created with Biorender.

(B) Bright-field microscopy images showing ICSIG-1 naive hPSCs and converted cells under ASECRiAV. Scale bar 500 μm.

(legend continued on next page)

sources proposed (Pera and Rossant, 2021; Rossant and Tam, 2022). Early studies suggested that the EXM is derived from the trophoblast, due to its location and emergence prior to formation of the primitive streak (Hertig, 1935; Hertig and Rock, 1941; Hertig et al., 1956). Others suggested that the EXM originates from the early primitive streak, due to its appearance in a similar region of the epiblast (Hill, 1932; Luckett, 1978). The EXM in mice and other species originates from the primitive streak at gastrulation (Saykali et al., 2019). However, in primates, the EXM is found prior to primitive streak formation, therefore the primitive streak cannot be the only source of EXMCs (Enders and King, 1988). An epiblast origin has been suggested as cells expressing mesoderm genes align closely to epiblast cells in a monkey embryo scRNA-seq dataset, prior to the emergence of the primitive streak (Yang et al., 2021). EXM was also suggested to originate from PrE based on electron microscopy images (Luckett, 1978), their shared gene expression with PrE (Nakamura et al., 2016), and lineage tracing using mutations (Spencer Chapman et al., 2021). A combination of origins is also possible (Ross and Boroviak, 2020; Yang et al., 2021). The regulatory elements underlying EXM identity in humans are also unknown.

Human embryo development is difficult to study because it occurs *in utero*. Advances have enabled culturing human embryos *ex utero* up to 14 days (Deglincerti et al., 2016; Shahbazi and Zernicka-Goetz, 2018; Shahbazi et al., 2016; Xiang et al., 2020; Zhou et al., 2019). However, obtaining human embryos for research remains a challenge due to extensive ethical and legal restrictions (Lovell-Badge et al., 2021). As a result, our understanding of early human development remains limited. To fill this gap, a number of human stem cell-based embryo models have been developed to recapitulate specific stages of human embryogenesis (Fan et al., 2021; Kagawa et al., 2021; Liu et al., 2021; Moris et al., 2020; Shao et al., 2017; Simunovic et al., 2022; Sozen et al., 2021; Veenvliet et al., 2020; Yanagida et al., 2021; Yu et al., 2021; Zhao et al., 2021). Naive human pluripotent stem cells (hPSCs) represent the preimplantation naive pluripotent epiblast and have the ability to differentiate into embryonic lineages as well as extraembryonic PrE and TB lineages, including human trophoblast stem cells (hTSCs) and amnion (Castel et al., 2020; Cinkornpumin et al., 2020; Dong et al., 2020; Guo et al., 2021; Io et al., 2021; Karvas et al., 2022; Linneberg-Agerholm et al., 2019; Rostovskaya et al., 2022). However, whether naive hPSCs have the ability to form additional extraembryonic lineages such as the EXM is unknown.

Attempts to form human blastoids recapitulating aspects of blastocyst development and cellular composition have been made (Fan et al., 2021; Kagawa et al., 2021; Liu et al., 2021; Yanagida et al., 2021; Yu et al., 2021; Zhao et al., 2021). The power of such models to predict development depends on their ability to form cells reflecting the blastocyst stage. Blastoids generate different extents of off-target cells depending on the initial cell state and molecules used to stimulate their formation. The lineage identity and developmental stages of the cells generated remain heavily debated and were proposed to correspond to postimplantation epiblast, primitive streak, amnion, mesoderm-like cells, and EXMCs in humans (Kagawa et al., 2021; Yanagida et al., 2021; Zhao et al., 2021) and to embryonic mesoderm or EXM in mice (Posfai et al., 2021).

Here, we report the discovery of EXMC specification from naive hPSC cultures. We surmise that modeling EXMC specification could help understand cell fate specification events in human peri-implantation embryogenesis, defects of which may cause developmental failure. Our work demonstrates that naive hPSC cultures can differentiate into EXMCs and establish a model that allows the study and manipulation of early human postimplantation development *in vitro*.

RESULTS

Derivation of EXMCs from naive hPSCs

We sought to derive hTSCs from naive hPSCs (PXGL) exposed to hTSC media ASECRiAV consisting of A83-01 and SB431542 (TGF-β type I receptor ALK4, 5, 7 inhibitors), hEGF, CHIR99021 (GSK-3 inhibitor), Y-27632 (ROCK inhibitor), insulin transferrin selenium ethanolamine (ITS-X) and valproic acid (histone deacetylase inhibitor) (Figure 1A) (See STAR Methods; Castel et al., 2020; Dong et al., 2020; Okae et al., 2018). By day 30 of conversion, we observed colonies with hTSC morphology and GATA3 expression (Figures 1B and 1C). These results suggested induction of hTSCs, as expected.

Unexpectedly, another cell type with mesenchymal morphology and mostly lacking GATA3 expression was present in day 30: ASECRiAV cultures (Figure 1B). We consistently obtained both hTSCs and the other cell type in all conversion attempts (>35 experiments) and the ratio of hTSCs to the other cell types varied. hTSCs, which express the epithelial marker CDH1 (Okae et al., 2018), could be separated from CDH1− cells by fluorescence-activated cell sorting (FACS) (Figure 1D). The CDH1− cells appeared to self-renew and expand for over 14 passages (70 days). Similar results were obtained with two other naive hPSC lines (H9 and WIBR2-MGT) grown in two different naive media (Figure S1A). In summary, we found that differentiation of naive hPSCs using hTSC conditions gives rise to an unexpected, CDH1− mesenchymal cell type.

To establish the identity of all cell types obtained by ASECRiAV conversion and compare them to the human embryo, we applied single-cell RNA sequencing (scRNA-seq) on day 30 of conversion with naive and primed hPSCs as controls. We obtained the transcriptome of 629 single cells. Uniform manifold approximation and projection (UMAP) revealed 4 clusters, each

(C) IF for the indicated marker in PXGL and ASECRiAV. Scale bar 100 μm.
(D) Flow cytometry contour plot of day 30 ASECRiAV cells analyzed for CDH1. Microscopy images of naive hPSCs converted under ASECRiAV for 30 days and cells sorted for lack of CDH1. Scale bar 500 μm.
(E) UMAP of day 30 TB conversion, naive and primed hPSCs scRNA-seq data.
(F) UMAP of integrated datasets of published human embryos, reference hPSCs, and this study.
(G) Selected cell type annotations from (F).
(H) UMAP of integrated datasets from monkey embryos, human-monkey chimera, and data from this study.
(I) Selected cell type annotations from (H). See also Figure S1, Tables S1, and S5.

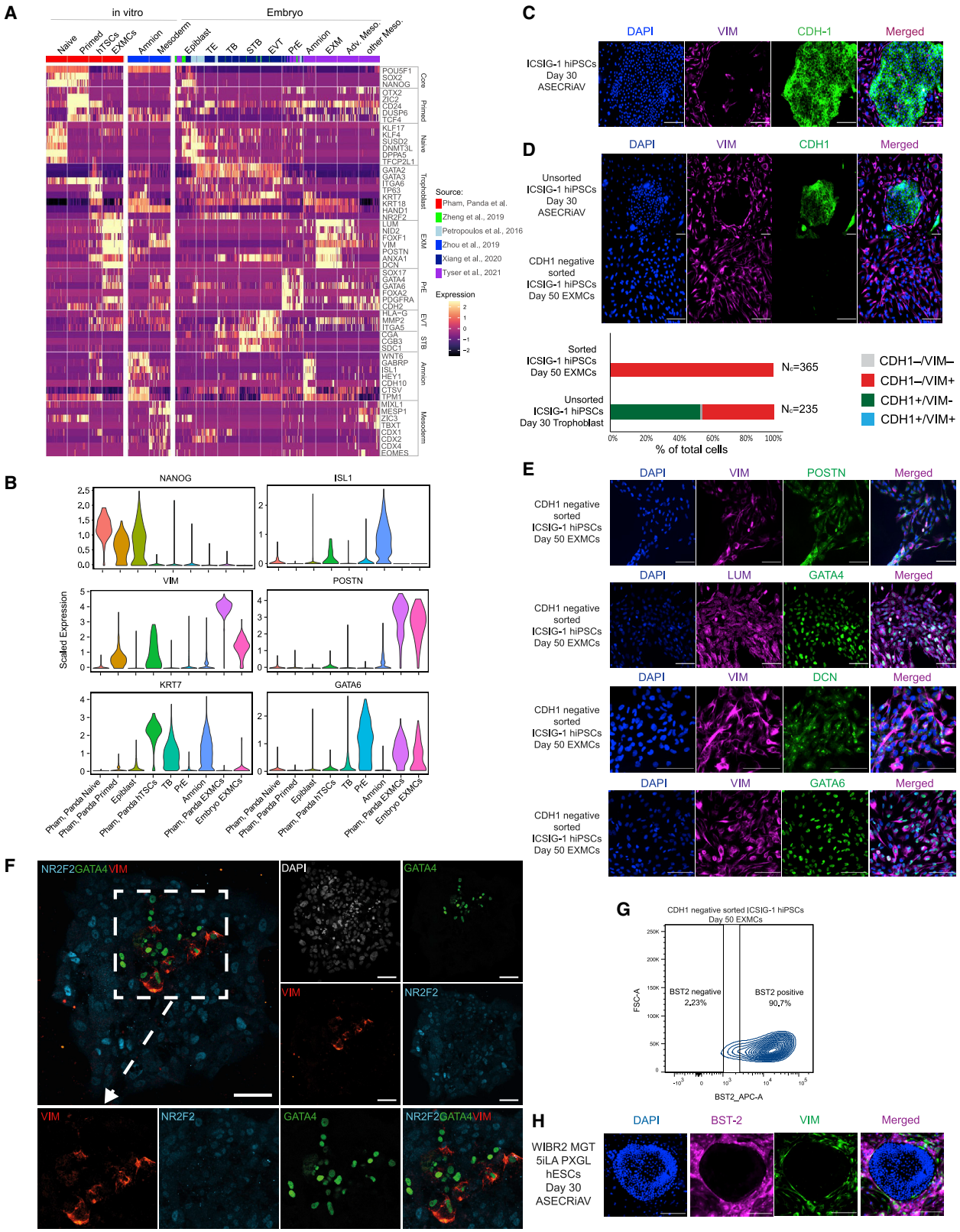


Figure 2. Characterization of EXMCs

(A) Marker genes expression heatmap.

(B) Marker gene expression violin plots.

(legend continued on next page)

corresponding to a specific cell type (Figure 1E). To determine if cells correspond to cells of the human embryo, we integrated the scRNA-seq data with a reference human embryo atlas, which included datasets of preimplantation, postimplantation, and a Carnegie stage 7 human embryo (Petropoulos et al., 2016; Tyser et al., 2021; Zhou et al., 2019), as well as reference hPSCs (Messmer et al., 2019) (Figures 1F and S1B). UMAP showed that cells were arranged according to developmental progression; hPSCs and hTSCs overlapped with their embryo counterparts, namely epiblast and early postimplantation trophoblast, respectively (Figure S1C). Unexpectedly, the undefined mesenchymal cells aligned mostly to the EXM of the Carnegie stage 7 human embryo, suggesting that they reflect EXM (Figure 1G). A few unidentified mesenchymal cells aligned to the PrE. We verified this analysis by integrating the scRNA-seq data with monkey embryo data, which contains day 14 monkey EXM (Tan et al., 2021; Yang et al., 2021). Naive and primed hPSCs aligned with monkey epiblast, as expected (Figure 1H). The unidentified mesenchymal cells aligned to the EXM (Figures 1H and 1I), which were annotated as extraembryonic mesenchyme cells by Yang et al. (2021). Extraembryonic mesenchyme and extraembryonic mesoderm have been used interchangeably in the literature (Rossant and Tam, 2022; Yang et al., 2021). Analysis confirmed a high correlation of the mesenchymal cells with human and monkey embryo EXM (Figures S1D and S1E, Table S1). These results suggest that the unexpected, mesenchymal cell type obtained by differentiation of naive hPSC cultures are EXMCs.

In vitro EXMCs recapitulate the gene expression profile of postimplantation human and monkey embryo EXM

To determine if the gene expression profile of primate EXM in the embryo is recapitulated *in vitro*, we analyzed the expression of known marker genes for each cell type (Table S2). Primed hPSCs expressed high levels of ZIC2 and CD24; naive hPSCs expressed KLF17, KLF4, DNMT3L, and DPPA5; hTSCs expressed GATA2, GATA3, and KRT7 (Figure 2A). Importantly, *in vitro* EXMCs expressed several primate embryo EXM marker genes, including POSTN, VIM, and NID2 (Figures 2A and S2A) (Niu et al., 2019; Tyser et al., 2021; Yang et al., 2021). Most EXM marker genes were not expressed in primed and naive hPSCs nor hTSCs. Comparing *in vitro* EXMCs to other *in vitro* and embryo data revealed that *in vitro* and embryo EXM are highly similar to each other and differ from other cell types (Figure 2A). Although several EXMC genes such as VIM, LUM and POSTN mark the amnion (Ma et al., 2019; Sozen et al., 2021; Tyser et al., 2021), we found these genes to be more highly expressed in the EXM compared to the amnion (Figure 2B). Immunofluorescence (IF) imaging of day 30 ASECRiAV cultures and of EXMCs confirmed the presence of VIM, LUM, POSTN, DCN, GATA4, and GATA6 and the absence of CDH1 in EXMCs, whereas hTSCs lacked VIM and expressed CDH1 (Figures 2C, 2E and S2B). These results confirm that EXMCs obtained by conversion from naive hPSC cultures transcriptionally match human and monkey embryo EXM and express specific key proteins.

EXMCs did not represent another cell type, although there was partial overlap in gene expression with other cell types. EXMCs lacked PrE marker FOXA2 and hTSC markers GATA2, TP63, and KRT7, but expressed SOX17, GATA4, GATA6 and PDGFRA (PrE), KRT18, HAND1, and NR2F2 (TSC). EXMCs did not express key mesoderm gene T/TBXT/BRACHYURY, in line with their pregastrulation origin and the lack of T expression in monkey embryo EXMCs (Cui et al., 2022), but unlike mouse EXMCs, which do express *T/Brachyury* due to their gastrulation origin (Peng et al., 2020). Multiple mesoderm markers including MESP1, GSC, and EOMES were absent from EXMCs (Figure 2A), indicating that EXMCs do not represent embryonic mesoderm. EXMCs also did not express amnion marker genes such as ISL1, HEY1, CDH10, or CTSV.

Differential gene expression analysis between naive hPSCs and day 30 EXMCs revealed 38 genes that were significantly increased in EXMCs (>5.6-fold, p value <0.05) (Figure S2C, Table S3) including key human and monkey EXM genes such as VIM, POSTN, DCN, GATA4, LUM, and H19 (Figure S2C). Gene ontology (GO) analysis of EXMCs identified an enrichment of genes in the mesodermal commitment pathway (Figure S2D).

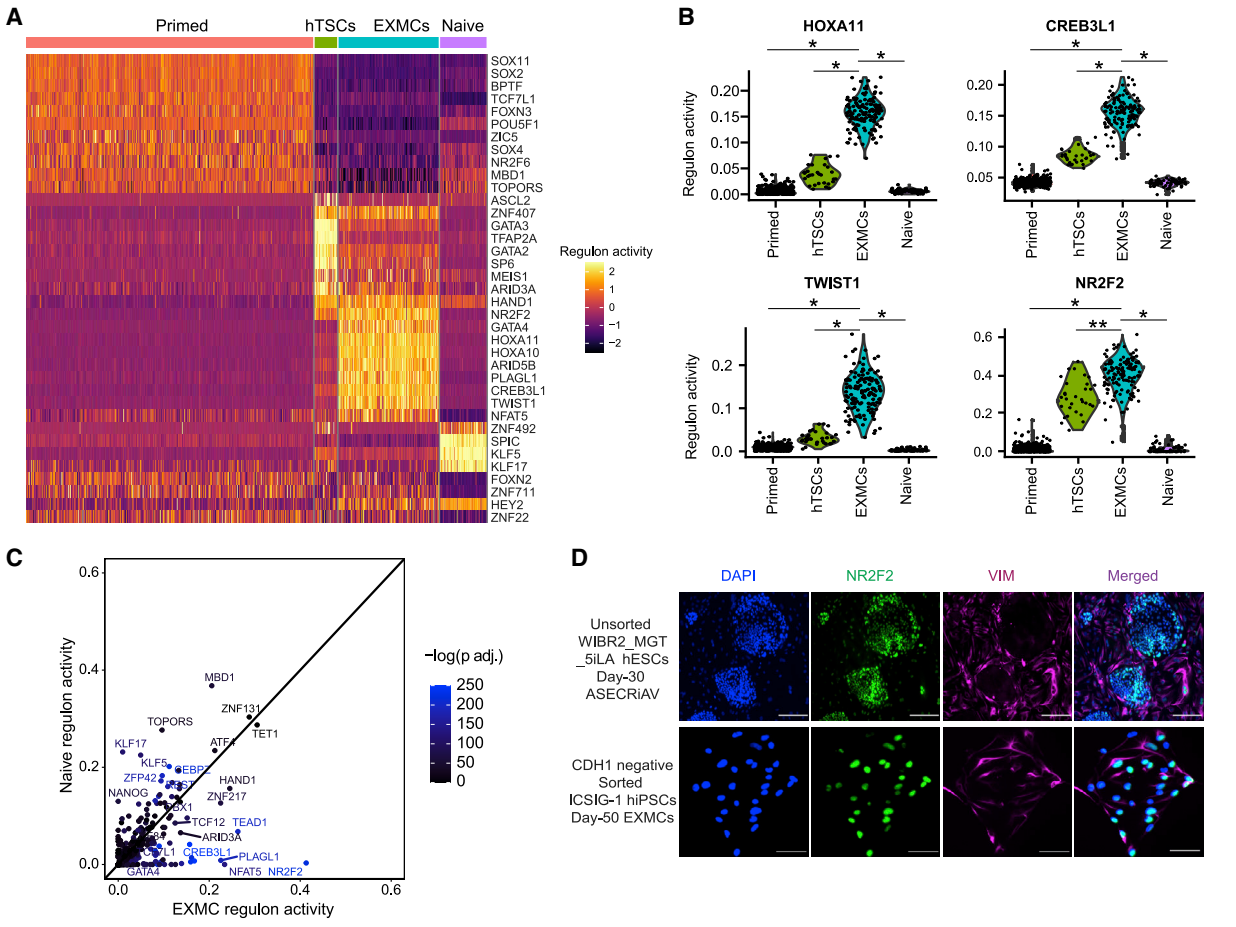
In summary, these results suggest that exposing naive hPSC cultures to human hTSC culture conditions results in the induction of hTSCs as well as EXMCs, and that EXMCs *in vitro* are similar to EXMCs in early postimplantation primate embryos and distinct from other cell types. Altogether, we captured *in vitro* a primate-specific postimplantation human embryo cell type, making it accessible for experimentation.

Conversion of naive hPSCs to EXMCs models, an epithelial-to-mesenchymal transition

It has been proposed that in monkey embryos, EXMCs are the first cells to undergo EMT during embryogenesis (Enders and King, 1988). Thus, we sought to characterize epithelial-to-mesenchymal (EMT) transitions in EXMCs. EXMCs showed mesenchymal morphology (Figures 1B and 1C). As EXMCs lost CDH1 expression, they gained expression of the mesenchymal marker CDH2 (Figure S2E). EXMCs expressed mesenchymal marker VIM (Figures 2A–2C) along with TWIST1, SNAI2, and ZEB2, which promote EMT (Figure S2F) (Cui et al., 2022; Dongre and Weinberg, 2019). Trophectoderm (TE) maturation marker NR2F2 (Meistermann et al., 2021), which is expressed in EXMCs, also regulates expression of multiple key transcription factors (TFs) that promote EMT, such as ZEB1/2 and PRRX1 (Mauri et al., 2021). Enriched GO terms in EXMCs included signaling for EMT, suggesting signaling might be implicated in EMT in EXMCs (Figure S2D). We conclude that EXMCs acquire features of mesenchymal cells consistent with an EMT.

To examine EMT and EXM specification in the embryo, we immunostained day 3 postimplantation human embryos

(C) IF for the indicated markers in day 30 ASECRiAV cells. Scale bar 200 μm.
(D) As in 2C for the indicated cell types. Scale bar 200 μm; bottom: quantification; nc: total nuclei count.
(E) IF for the indicated markers and cell types. Scale bar 100 μm.
(F) IF for the indicated markers in a day 10 human embryo. Scale bar 100 μm.
(G) BST2 flow cytometry of sorted CDH1- EXMCs.
(H) IF for the indicated markers in day 30 ASECRiAV cells. Scale bar 200 μm. See also Figure S2, Videos S1 and S2, and Table S2.



(legend on next page)

(day 10 post fertilization) grown *ex vivo* with antibodies against VIM, NR2F2 and GATA4. These embryos all had NANOG+ epiblast, GATA4+ PrE and NR2F2+ TB (Figure S2G). Excitingly, in two out of eight embryos, we found VIM+ cells, suggesting EMT may be initiated around day 10 of human embryogenesis (Figures 2F and Videos S1 and S2). Moreover, VIM+ cells were restricted to a subset of cells marked by GATA4 and low NR2F2 expression. As *in vitro* EXMCs also express this unique combination of marker genes (Figure 2A), VIM+ cells in day 10 embryos may represent EXM, an early precursor, or another cell type. Altogether, these results suggest that VIM+ cells in day 10 embryos may be undergoing EMT and/or EXM specification.

Extracellular matrix gene expression in EXMCs

A hallmark of EXMCs in the early primate embryo, and of stromal cells in general, is production of an extracellular matrix (Bonnans et al., 2014; Enders and King, 1988; Nakamura et al., 2016; Ross and Boroviak, 2020). We therefore analyzed the expression of key monkey and human embryo EXM extracellular matrix genes such as FN1, specific collagens (COL1A1, COL1A2, COL3A1, COL4A1, COL6A1, and COL6A3) and laminins (LAMB1, LAMC1), which were increased in EXMCs relative to naive hPSCs and, to a large extent, also to hTSCs (Figures S2C and S2H, Table S3). Additionally, GO analysis identified "focal adhesion" and "extracellular matrix receptor interactions" as top enriched pathways in EXMCs (Figure S2D). Therefore, EXMCs derived from naive hPSC cultures possess another hallmark of the EXM: the expression of a specific set of extracellular matrix genes.

BST2 as a cell surface marker of EXMCs

We aimed to identify EXMC cell surface markers by examining genes with differential gene expression between EXMCs and hTSCs (adjusted p value <0.01 and expression fold change >2). This identified 11 candidate cell surface genes, of which BST2 was chosen as it was also expressed in human embryo EXM, but not in most other embryonic cell types, with the exception of epiblast and mesoderm (Figure S2I). Flow cytometry for BST2 on EXMCs revealed that 90.7% of EXMCs were positive for BST2 (Figure 2G), which was further confirmed by IF (Figure 2H). Therefore, we identified BST2 as a cell surface marker of EXMCs.

Gene regulatory networks in EXMCs

TFs control cell fate specification by binding to *cis*-regulatory regions, thus forming gene regulatory networks (GRNs) (Davidson and Levin, 2005), yet the GRNs of human EXMCs remain undefined. Therefore, we used single-cell regulatory network inference and clustering (SCENIC) analysis (Aibar et al., 2017) to reconstruct GRNs and predict TFs with regulatory activity in EXMCs (Figure 3A). SCENIC measures TF regulatory activity by combining expression of TFs and their candidate target genes that are co-expressed and have TF binding motifs. The expected regulons for KLF17, SOX11, and GATA3 were found active in naive, primed hPSCs and hTSCs, respectively. TFs recently shown to be essential regulators of human hTSC identity including ARID3A, GATA2, and ZNF407 (Dong et al., 2022; Rhee et al., 2017) had high regulatory activity in hTSCs, supporting the notion that the approach can identify critical regulators of cell identity. We detected high regulatory activity in EXMCs for TFs including ARID5B, PLAGL1, CREB3L1, GATA4, and FOXF1 (Figure 3A and Table S4), in line with their reported expression or regulatory activity in monkey EXM (Nakamura et al., 2016; Niu et al., 2019; Yang et al., 2021). TWIST1, an important mesoderm regulator (Qin et al., 2012), also showed regulatory activity in EXMCs. Several HOX genes including HOXA10, HOXA11, HOXA9, and HOXA13 had high activity in EXMCs (Figures 3B and S3A), in line with the recently reported regulatory activity of HOXA11 in the monkey EXM (Cui et al., 2022). CREB3L1, which was previously reported as a human-specific epiblast factor (Boroviak et al., 2018) and has reported roles in extracellular matrix formation (Chen et al., 2014), was also active in EXMCs (Figure 3B). We detected high activity of TFs shared between hTSCs and EXMCs, including HAND1 and NR2F2 (Figures 3B and S3A). The regulatory activity of NR2F2 was unexpectedly higher (adjusted p value $3.9*10^{-8}$) in EXMCs than in hTSCs and naive hPSCs (Figures 3B and 3C). Surprisingly, although we could confirm the expression of NR2F2 in EXMCs by IF, NR2F2 appeared more highly expressed at the protein level in hTSCs than in EXMCs (Figure 3D). In summary, EXMCs possess regulatory activity for a unique combination of TFs typically associated with TB, PrE, and mesoderm as well as TFs expressed in human and monkey embryo EXM, providing a valuable resource for future studies.

Single-cell chromatin accessibility profiling of EXMCs

To further characterize EXMCs and identify regulatory elements that may underlie EXMC identity, we performed single-cell ATAC sequencing (scATAC-seq). We obtained the chromatin accessibility landscape of 1,133 cells comprising naive hPSCs, primed hPSCs, hTSCs, and EXMCs, which clustered into 4 populations (Figure S3B). 40,314 peaks were called for the EXMC population, with 22,444 being unique (Figure S3C). scATAC-seq revealed that EXMCs had high chromatin accessibility in the vicinity of most EXM marker genes (NID2, FOXF1, POSTN) and lacked chromatin accessibility at regions associated with several genes of alternative cell fates (Figures 3E and S3D). TF motifs including

Figure 3. Gene regulatory networks and scATAC-seq profiles of EXMCs

(A) Activity of the top differentially active regulons (SCENIC).

(B) Regulon activity for indicated TFs. Significant difference between regulon activity, Wilcoxon rank-sum test, *adjusted p < $2*10^{-16}$, ** adjusted p = $3.9*10^{-8}$. See Table S5 for the n number of cells included in each cell type.

(C) Median regulon activity in naive hPSCs versus EXMCs. Points colored by –log(Bonferroni adjusted p value). See Table S5 for the n number of cells included in each cell type.

(D) IF for the indicated markers in the indicated cell types. Scale bar 100 μm.

(E) Dot plot of marker genes chromatin accessibility (scATAC-seq).

(F) scATAC-seq motif enrichment. See also Figure S3 and Table S4.

HOXA10, HOXA13, TWIST1, CREB3L1, and NR2F2 were found to be highly enriched in accessible chromatin of EXMCs, corroborating SCENIC analyses (Figure 3F). Our results identify candidate *cis*-regulatory elements and TFs underlying the human EXM gene regulatory program. Of note, we also report as a resource single-cell epigenomic data for primed and naive hPSCs as well as hTSCs.

Contribution of ASECRiAV components to EXMC induction

We aimed to define the ASECRiAV components that are most important for EXMC induction. We removed ASECRiAV components individually or in combinations (SB431542/A83-01; ITS-X/hEGF), then induced EXMC conversion, and assessed induction by immunostaining for GATA4 and VIM at day 12. Despite high variability, all conditions contained GATA4/VIM double-positive cells, suggesting that no single ASECRiAV component is strictly required for EXMC induction. However, ITS-X and hEGF removal decreased EXMC induction, especially when removed together (Figures S4A, S4B, and S4C). Insulin and the insulin-like growth factor activate the PI3/AKT and mTOR pathway (Budi et al., 2015; Taniguchi et al., 2006; Wamaitha et al., 2020), suggesting insulin might activate the mTOR pathway during EXMC induction, in addition to the recognized role of this pathway in human pluripotency (Wamaitha et al., 2020). Other factors may also contribute to EXMC induction, although to a smaller extent. Collectively, these results suggest that ITS-X and hEGF are the most important factors in ASECRiAV for EXMC induction.

BMP4 and mTOR signaling in EXMCs

We aimed to identify signaling pathways that maintain EXMCs. GO analysis revealed that the TGF-β superfamily and mTOR signaling pathways are enriched in EXMCs (Figure S2D). TGF-β is surprising since ASECRiAV contains two TGF-β inhibitors, A83-01 and SB431542. However, these only target ALK4, 5, and 7, which are receptors activated by ligands of one branch of the TGF-β superfamily, namely ACTIVIN/NODAL/TGF-β (Li et al., 2021; Osnato et al., 2021; Wakefield and Hill, 2013). The other branch of the TGF-β superfamily is activated by the BMP signaling pathway, including BMP4, which acts through ALK1, 2, 3, or 6 and is not targeted by inhibitors in ASECRiAV (Shi and Massagué, 2003). Accordingly, higher expression of BMP4 and of its downstream target genes ID2 and ID3 in EXMCs, both *in vitro* and *in vivo* as compared to naive hPSCs, suggest that this pathway is active in EXMCs (Figures S4D and S4E). Several genes related to the mTOR pathway were also highly expressed in EXMCs, suggesting that the mTOR pathway is active (Figures S4D and S4E).

To determine if the BMP4, TGF-β, and mTOR pathways are active in EXMCs, we used western blot analysis for phosphorylation of key signal transduction proteins in these pathways. SMAD1/5/9 was highly phosphorylated in EXMCs, but not in naive hPSCs, indicating that the BMP4 pathway is active in EXMCs (Figures 4A and 4B). Increased SMAD1/5/9 phosphorylation was validated by immunostaining in EXMCs (Figure 4C). Western blot analysis against phosphorylated SMAD2, an effector specific to the ACTIVIN/NODAL/TGF-β branch (Derynck et al., 1998), confirmed that this branch of the TGF-β pathway is active in naive hPSCs and repressed in EXMCs (Figures 4A and 4B). The mTOR pathway was active both in EXMCs and in naive hPSCs, as suggested by P-4EBP1 analysis (Figures 4A–4C). Collectively, these results show that the BMP4 and mTOR signaling pathways, but not the ACTIVIN/NODAL/TGF-β pathway, are active in EXMCs.

To determine whether the BMP4 and mTOR pathways are required for EXMC maintenance, we inhibited the BMP4 pathway using the inhibitor LDN-193189 (LDN), targeting receptors ALK1, 2, 3, and 6. EXMCs were treated daily for 10 days. Western blot analysis confirmed a decrease of phosphorylated SMAD1/5/9 upon inhibition by LDN (Figure 4A), which correlated with reduced EXMC growth (Figures 4D and S4F). These results show that the BMP4 pathway is needed for EXMC growth. To determine the effect of the mTOR pathway on EXMC maintenance, we inhibited the mTOR pathway using inhibitor GSK1059615 targeting PI3Kα/β/δ/γ and mTOR. GSK1059615 inhibited cell growth and induced cell death (Figures 4D and S4F). Thus, the mTOR pathway is required for EXMC growth and survival. Altogether, we conclude that the BMP4 and mTOR pathways are implicated in EXMC maintenance with the monkey EXM (Niu et al., 2019).

Single-cell transcriptome analysis of differentiation kinetics

To create a differentiation trajectory, we used single-cell RNA-seq time course analysis. We collected samples at different timepoints during ASECRiAV conversion (Figure 5A). Cells were passaged on days 5, 10, and 15. We included sorted day 70 EXMCs. After quality control, we obtained data for 12,977 cells with an average of 1,622 cells per sample (Table S5). UMAP analysis showed the presence of distinct populations at different times during conversion (Figures 5B and 5C). At day 0, the majority of cells were naive hPSCs and expressed pluripotency genes, with a small fraction of cells corresponding to 8-cell-like cells (8CLCs) marked by expression of 8-cell stage embryo genes (Figures 5C–5D), as recently reported (Mazid et al., 2022; Taubenschmid-Stowers et al., 2022). Naive cells were only detected at day 0 (Figures 5B–5E). By day 1, most cells appeared to progress into an epiblast intermediate (Figure 5C), marked by decreased expression of naive pluripotency genes DNMT3L and KLF4 and continued expression of several formative pluripotency genes such as DPPA2, GDF3, ZNF728, and ZNF729 (Kinoshita et al., 2021) (Figures 5D and 5F). This intermediate epiblast population formed 3 clusters, which we termed epiblast intermediates 1, 2, and 3, reflecting their progression during differentiation (Figures 5B and 5C). hTSCs represented a low fraction of the cells up to day 4 but became the most abundant cell type at day 8 (Figure 5E). Their abundance declined at day 13 and was minimal by day 18.

The first EXMCs appeared on day 8 and clustered with EXMCs from day 13 and day 18; we termed this cluster early EXMCs (Figures 5C and 5E). In early EXMCs, pluripotency genes were silenced and EXMC genes LUM, NID2, FOXF1, VIM, and POSTN were activated (Figure 5F). At day 18, the majority of cells were EXMCs (Figures 5C–5E). Intriguingly, day 70 EXMCs isolated by FACS formed a distinct population, which we termed late EXMCs (Figures 5C–5E). These cells clearly expressed multiple EXMC genes such as LUM, NID2, VIM, and POSTN (Figure 5F). Several genes including NID2, POSTN, and NR2F2

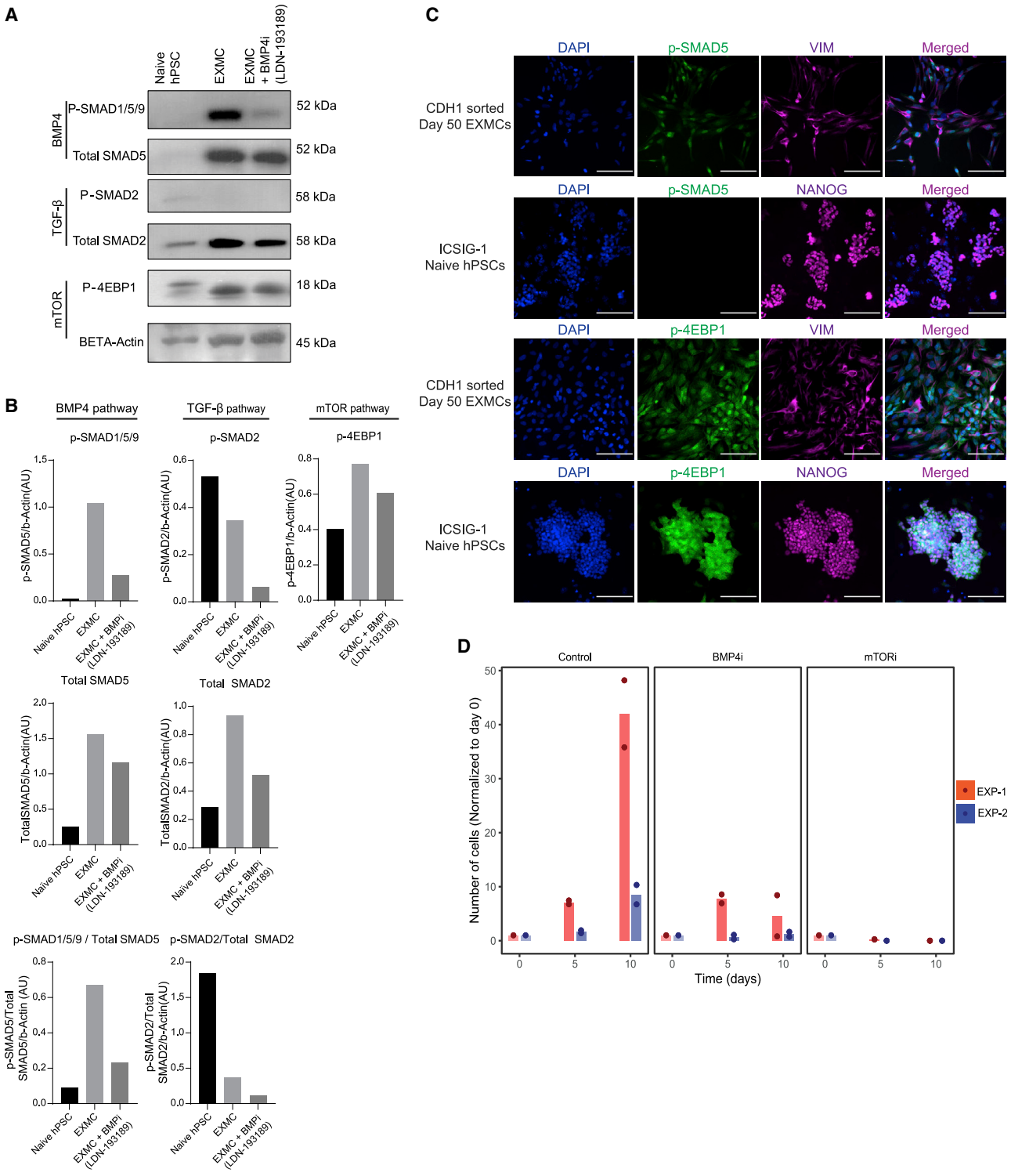


Figure 4. Signaling pathways in EXMCs

(A) Western blot analysis for the indicated proteins in naive ICSIG-1 hPSCs and EXMCs with and without BMP4 inhibition (LDN-193189).

(B) Quantification of (A). AU: arbitrary unit. Quantification from n = one experiment.

(C) IF for the indicated markers in the indicated cell types. Scale bar 200 μm.

(D) Number of EXMCs grown for 5 or 10 days under either BMP4 inhibition (1 μM LDN-193189) or mTOR inhibition (1 μM GSK1059615). n = 2 experiments. Biological replicates are shown as individual data points. See also Figure S4.

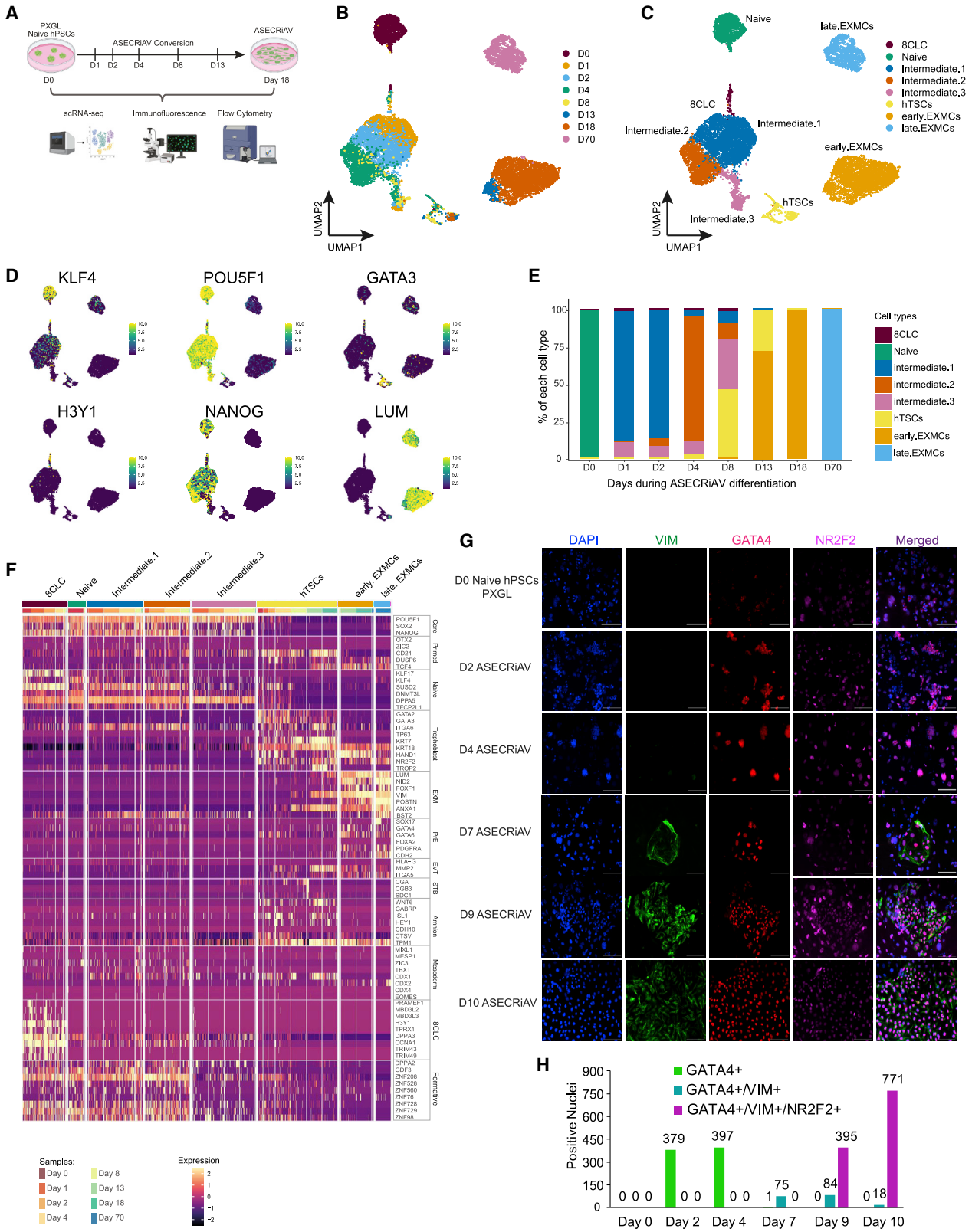


Figure 5. Single-cell time course analysis

(A) Experimental strategy. Image created with Biorender.

(B) UMAP of time course scRNA-seq data during differentiation of naive hPSCs under ASECRiAV condition as well as day 70 EXMCs, colored by days.

(legend continued on next page)

were more highly expressed in late EXMCs compared with early EXMCs. Together, these results show that EXMCs can be derived *in vitro* by day 8, and that their identity can be maintained for at least 70 days, with expression of EXMC marker genes increasing over time.

To relate these changes in cell identity to embryo development, we integrated the time course scRNA-seq data with our initial day 30 data and human embryo atlas (Messmer et al., 2019; Petropoulos et al., 2016; Tyser et al., 2021; Xiang et al., 2020; Zhou et al., 2019). The integration UMAP showed good correspondence between *in vitro* cell types and the embryo (Figures S5A–S5C). In particular, naive hPSCs reflected the naive epiblast, and intermediate epiblast cells aligned between the naive and primed epiblast. hTSCs reflected early TB, while early EXMCs reflected embryo EXM. Embryo EXM generally showed high correlation with *in vitro* EXMCs at all time points (Figure S5C). Day 14 embryo EXM (Chhabra and Warmflash, 2021; Xiang et al., 2020) correlated best to days 8 and 13 *in vitro* EXMCs, while Carnegie stage 7 embryo EXM (embryonic days 16–19) (Tyser et al., 2021) correlated best to day 18 *in vitro* EXMCs (Figure S5C and Table S1). These results suggest that *in vitro* EXMC differentiation follows a progression that resembles that of the embryo EXM, with earlier *in vitro* cells correlating better with the earlier *ex vivo* embryo time point and later *in vitro* cells correlating better with later embryo stages. As the day 70 EXMCs clustered separately from the other *in vitro* EXMCs and embryo EXM, they may represent an as yet unstudied later stage of EXM development.

The scRNA-seq experiments above suggested that acquisition of the EXMC fate is a sequential process, marked by gradual acquisition of gene expression. While early EXMCs activated the expression of GATA4, VIM, ANXA1, COL4A1, COL4A2, and BMP4, late EXMCs showed higher expression of POSTN and PTX3 compared to early EXMCs and appeared to show increased NR2F2 expression (Figures S5D and 5F). Time course IF analysis confirmed the sequential activation of GATA4, then VIM, and finally NR2F2 and the absence of VIM+/GATA4-cells during naive to ASECRiAV conversion (Figures 5G and 5H). Collectively, these results provide insights into the sequence and progression of EXMC differentiation starting from naive hPSCs.

Origin of EXMCs

To determine the origin of EXMCs, we first investigated whether EXMCs arise from pre-differentiated cells in naive hPSC cultures by sorting SUSD2+ cells, which enrich for naive hPSCs (Bredenkamp et al., 2019a, 2019b; Wojdyla et al., 2020), and SUSD2- cells, to enrich pre-differentiated cells from naive hPSC cultures (Figures 6A and 6B). After replating in PXGL for 24 h, then switching to ASECRiAV for 12 days, both populations gave rise to EXMCs (Figures 6C and 6D). Hence, EXMCs do not appear to preferentially arise from SUSD2-pre-differentiated cells in naive hPSC cultures.

PrE cells were absent in the ASECRiAV time course (Figures 5C and 5F), so *in vitro* EXMCs are unlikely to be specified via a PrE intermediate. In the embryo, the EXM appears several days after the 8-cell stage (Luckett, 1978; Ross and Boroviak, 2020). Thus, a direct 8CLC to EXMC origin is unlikely and would be devoid of developmental logic, although it cannot be excluded.

In the scRNA-seq time course, naive hPSCs were seen only at day 0, while EXMCs were detected starting at day 8. Hence, EXMCs likely do not originate directly from naive hPSCs. Instead, EXMCs may arise from an intermediate cell state between naive and EXMC states (Figure 5D). Intriguingly, at day 4 of the time course, the majority of cells possessed an intermediate epiblast state. These results suggest that intermediate epiblast cells might be a source of EXMCs.

To test the ability of intermediate epiblast cells to give rise to EXMCs, we used cell surface marker BST2 (Figures 2G and 2H), which is also expressed in intermediate epiblast cells prior to EXMC specification (Figure S6A). We also used the TB lineage marker TROP2 (Figure S6B) (Io et al., 2021). We sorted cells at day 6 of conversion and grew these for another 6 days in ASECRiAV (Figure S6B). BST2+/TROP2- epiblast intermediate cells gave rise to EXMCs, indicating that epiblast intermediate cells are competent to differentiate into EXMCs (Figures 6E and 6F). hTSCs were also obtained (Figures 6G and 6H). These results suggest that epiblast intermediate cells are not irreversibly committed to the embryonic epiblast lineage, but instead are competent to differentiate into extraembryonic cell types, including EXMCs and hTSCs.

BST2–/TROP2+ cells isolated at day 6 of differentiation and grown further into ASECRiAV were enriched for cells that form hTSCs, as expected (Figures 6G and 6H). However, EXMCs were also obtained (Figures 6E and 6F). These results suggest that BST2–/TROP2+ cells at day 6 of differentiation may not yet be irreversibly committed to the TB lineage. However, in scRNA-seq data, we detected a small population of intermediate epiblast cells expressing TROP2, hence EXMCs may arise from a few BST2–/TROP2+ epiblast intermediate cells (Figure 5F). BST2/TROP2 double-negative cells also gave rise to both EXMCs and hTSCs (Figures 6E–6H). Intriguingly, BST2/TROP2 double-positive cells died and did not give rise to EXMCs or hTSCs (Figures 6E–6H). Collectively, these results suggest that epiblast intermediate cells marked by TROP2+ or BST2+ are both competent to differentiate into EXMCs and hTSCs. We propose a model in which naive hPSCs give rise to epiblast intermediate cells from which EXMCs originate.

Human naive pluripotent stem cell to PrE conversion

Several reports suggested a PrE origin of EXMCs (Luckett, 1978; Nakamura et al., 2016; Spencer Chapman et al., 2021). Thus, we tested if differentiation of naive hPSCs to the PrE fate induces

(C) As in (B), colored by cell types.
(D) Expression of selected markers in (B).
(E) Proportion of cell types during ASECRiAV conversion in (B). Cell numbers can be found in Table S5.
(F) Expression of selected genes during conversion.
(G) IF for the indicated markers during naive hPSC to ASECRiAV conversion. Scale bar 100 μm. Please note feeders have background staining for NR2F2 in D0-D4 images.
(H) Quantification of (G). n = two rounds of differentiation. See also Figure S5 and Table S5.

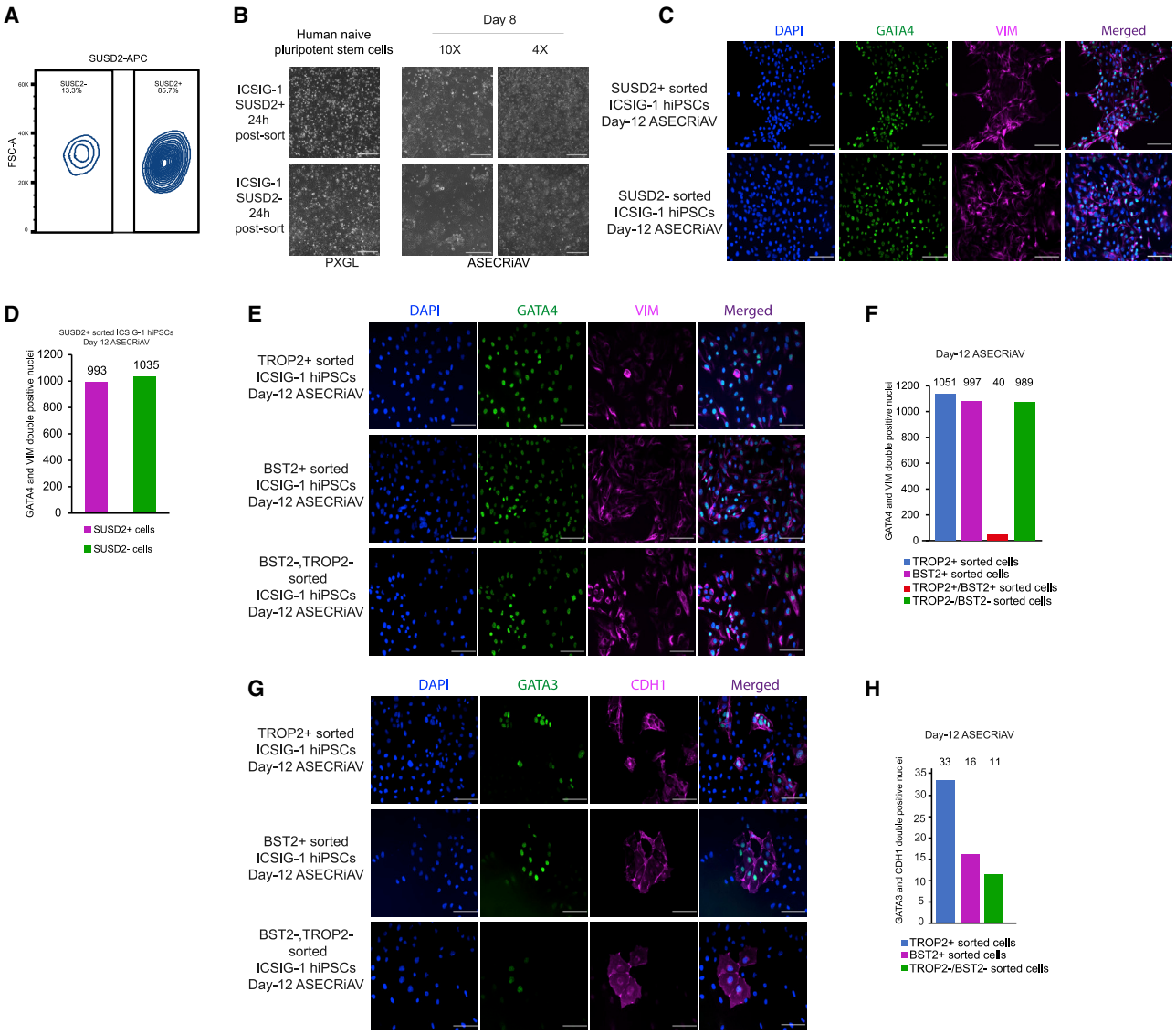


Figure 6. Origins of EXMCs

(A) Flow cytometry analysis of naive hPSCs (ICSIG-1).

(B) Bright-field microscopy images of SUSD2+ and –ICSIG-1 naive hPSCs 24 h after sorting and cultured back in PXGL (Top) and 8 days after switching to ASECRiAV (Bottom). Scale bar 500 μm.

(C) IF for the indicated markers in day 12 ASECRiAV cells converted from SUSD2 sorted naive cells. Scale bar 100 μm.

(D) Quantification of C. Nuclei counted from 5 random fields.

(E) IF for the indicated markers in TROP2+/BST2– sorted ICSIG-1 hPSCs, BST2+/TROP2– sorted ICSIG-1 hPSCs and TROP2–/BST2– sorted cells at day 12 of ASECRiAV conversion. Scale bar 100 μm.

(F) Quantification of E. Nuclei counted from 5 random fields.

(G) As in (E) in TROP2+/BST2– sorted ICSIG-1 hPSCs, BST2+/TROP2– sorted ICSIG-1 hPSCs and TROP2–/BST2– sorted cells at day 12 of ASECRiAV conversion. Scale bar, 100 μm.

(H) As in (F) but for (G). See also Figure S6.

EXMCs. We subjected naive hPSCs (PXGL) to RACL medium (See STAR Methods) for 6 days to induce a PrE fate (Linneberg-Agerholm et al., 2019), followed by an additional 24 days in NACL medium to induce a naive extraembryonic endoderm (nXEN) fate. To examine cell identity, we reconstructed single-cell transcriptomes at day 8 of RACL conversion and day 24 of NACL conversion (Figure S6C). 593 and 75 cells passed quality control for RACL and NACL samples, respectively. The cells organized into 6 clusters (Figures S6E and S6F). In addition to PrE cells, RACL conversion unexpectedly induced EXMCs. RACL EXMCs had low pluripotency gene expression and expressed EXMC genes including NID2, FOXF1, VIM, ANXA1, but not FOXA2 and LUM (Figures S6E and S6F). By integration with human embryo and ASECRiAV conversion data, they

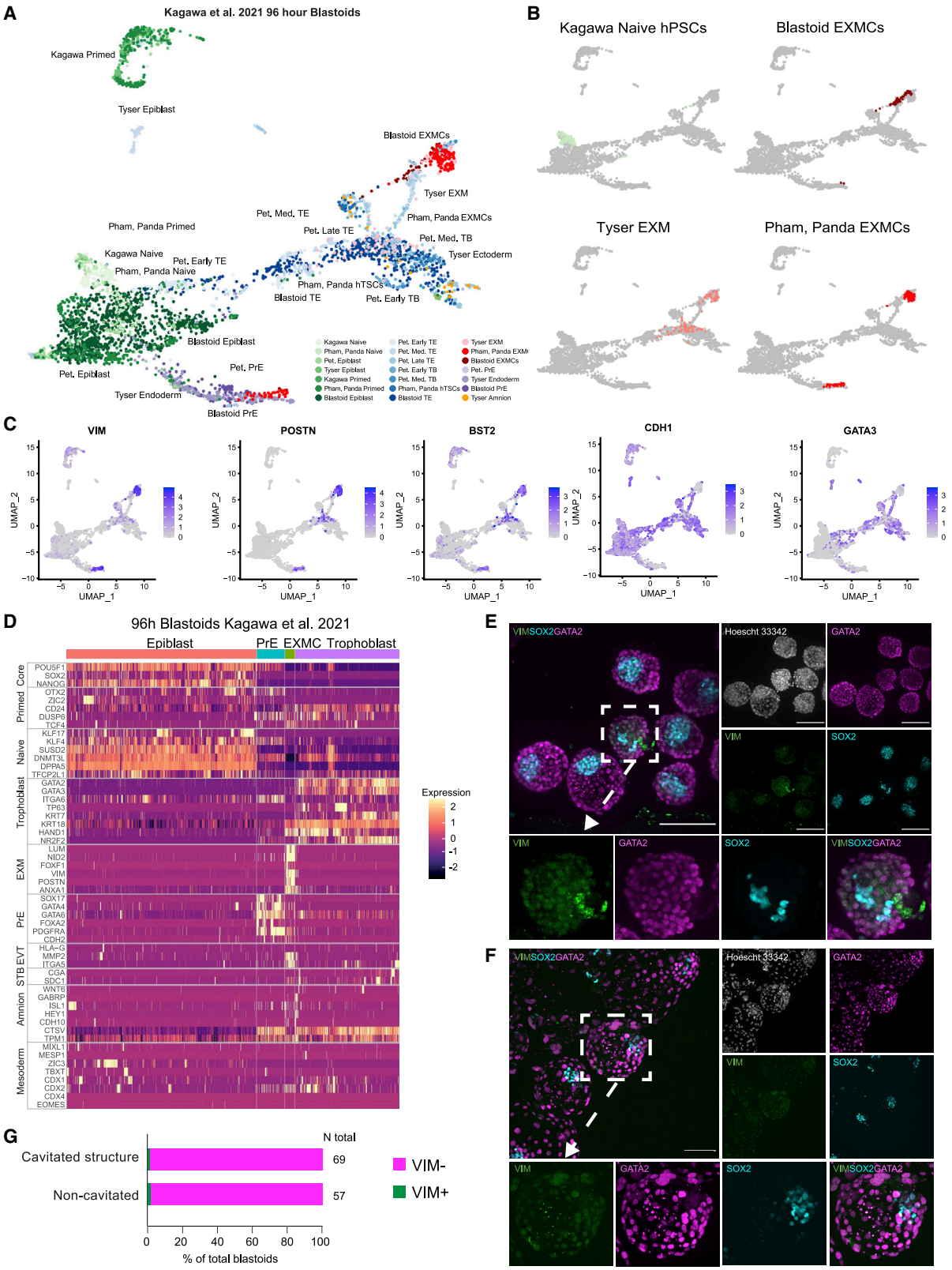


Figure 7. Blastoids contain EXM-like cells

(A) UMAP of the integration of 96 h blastoid data (Kagawa et al., 2021) with embryonic data sets (Petropoulos et al., 2016; Tyser et al., 2021) and our ASECRiAV conversion containing TB and EXMCs from Figure 1.

(legend continued on next page)

resemble embryo early EXMCs (Chhabra and Warmflash, 2021; Xiang et al., 2020) and correspond to early EXMCs obtained by ASECRiAV conversion (Figures S6G and S6H). Thus, unexpectedly, RACL conversion also induces the EXMC fate.

Interestingly, RACL also induced 3 intermediate populations, each of which were similar to the three corresponding epiblast intermediate populations obtained by ASECRiAV conversion, and expressed pluripotency genes POU5F1, NANOG and SOX2, as well as formative genes DPPA2, GDF3, ZNF208 and ZNF729 (Figures S6F–S6H). These results raise the possibility that, in RACL media, EXMCs also arise via an epiblast intermediate state, a PrE state, or both. In summary, we found that in addition to PrE-like cells and EXMCs, RACL induces epiblast intermediate cells.

We next sought to test the cell types obtained after culture of RACL cells in NACL media reported to induce a nXEN state (Linneberg-Agerholm et al., 2019). Unexpectedly, most NACL cells comprised EXMCs, but this time the cells corresponded to late EXMCs with similar gene expression profile as obtained in the ASECRiAV experiments above (Figures S6F–S6H). The EXMCs found in both RACL and NACL media had correlated gene expression with both early and late EXMCs found in ASECRiAV media and embryo EXM (Figure S6I, Table S1). These transcriptome analyses show that naive hPSCs grown in RACL and NACL generate several cell types, including EXMCs. As hTSCs were not present in RACL culture and PrE was not present in ASECRiAV culture, this further supports that EXMCs can be generated from an epiblast intermediate.

EXMCs and human blastoids

Attempts to form human blastoids have generated different extents of non-blastocyst-stage cell types, which would impair the potency of embryo models to predict development (Kagawa et al., 2021; Liu et al., 2021; Sozen et al., 2021; Yu et al., 2021; Zhao et al., 2021). We wondered whether non-blastocyst-stage cell types might correspond to EXMCs and re-analyzed data from human blastoids (Kagawa et al., 2021). Clustering of fully developed blastoids (96 h post-induction) with primed and naive hPSCs revealed 6 distinct populations of cells (Figure S7A). We then integrated datasets of fully developed blastoids and human embryos (Petropoulos et al., 2016; Tyser et al., 2021), along with *in vitro* EXMCs (Figures 7A, 7B, and S7B). We found that 1.6% (15/920) of the cells in fully developed blastoids did not align with the blastocyst-stage but rather with embryo EXM and *in vitro* EXMCs (Figures 7B and S7C). These non-blastocyst-stage cells expressed key EXM marker genes and had reduced expression of marker genes of alternative lineages including epiblast, PrE, TB, and mesoderm (Figures 7C and 7D). The majority of these non-blastocyst-stage cells corresponded to EXMCs (93%, 14/15 cells), while 7% (1/15 cells) also expressed amnion markers and may reflect amnion cells (Kagawa et al., 2021; Zhao et al., 2021) (Figure S7D). These results suggested that human blastoids (Kagawa et al., 2021; Liu et al., 2021; Sozen et al., 2021; Yu et al., 2021; Zhao et al., 2021) form <2% non-blastocyst-stage cells that, in majority, are EXMCs. Analysis of blastoids generated through another method (Yu et al., 2021) identified 3% of cells as EXMC-like cells, while we identified 9% of cells as EXMC-like cells in iBlastoids (Liu et al., 2021). Both of these other datasets included an additional 45% of cells not aligning to any of the four cell types examined here (blastocyst-stage epiblast, PrE, TB, and EXM).

To investigate the origin of EXMCs in blastoids, we analyzed earlier time points in blastoids (Kagawa et al., 2021). We found that 1.4% (2/139) of cells in the initial naive hPSC culture clustered with EXMCs (Figures S7E and S7F), either expressing PITX1 or PITX1 and NID2, but importantly lacking expression of most other EXMC marker genes, including VIM, with both cells expressing amnion markers, including ISL1 (Figure S7G). These pre-differentiated cells are therefore not mature EXMCs, but still may contribute to the population of EXMCs that we found in fully formed (96 h) blastoids.

We then examined the presence of EXMC-like cells during the course of blastoid formation. We observed cells that aligned with EXMCs at 24 h (2%, 4/200 cells) and 60 h (7.2%, 30/418 cells) of blastoid formation (Figure S7F). Thus, EXMC-like cells became progressively more abundant during early blastoid formation but remained rare in fully developed blastoids (1.6%). Importantly, contrary to cells harvested from fully formed blastoids (the latter accounting for ~75% of the total number of structures, see morphological criteria and efficiency for blastoids in Kagawa et al. [2021]), cells harvested at 24 and 60 h also included the 25% of structures that will not form blastoids (Kagawa et al., 2021). Given the increased proportion of EXMC-like cells at 60 h over 96 h, this raised the possibility that EXMC-like cells might appear preferentially in the ~25% of structures that do not form blastoids. We thus determined the context of emergence of VIM+ cells during blastoid formation using IF analysis for VIM, SOX2, and GATA2. We found that VIM+ cells were rare, with only 1 in 69 cavitated blastoids (<0.02% of cells) and 1 in 57 non-blastoid structures (Figures 7E–7G). These VIM+ EXMC-like cells are thus equally prominent in blastoids and non-blastoids and less abundant than the ones detected by scRNA-seq. Because the EXM arises after implantation, we conclude that the presence of rare VIM+ EXMC-like cells in a blastocyst model is inappropriate. Altogether, we propose that understanding human EXMC specification will help to improve stem cell-based embryo models and enable us to gain insights into mechanisms of early human embryogenesis.

DISCUSSION

To study and manipulate molecular processes underlying cell fate specification after implantation, an *in vitro* model of human

(B) Selected cell type annotations from (A).
(C) Selected gene expression from (A).
(D) Selected gene expression. Expression scaled as *Z* score, 0 is mean expression per gene across all cells, and 1 indicates cells with expression 1 standard deviation higher than the mean.
(E and F) IF for the indicated markers after 96 h of blastoid generation. VIM+ structures indicated with white box (top left), enlarged at the bottom. Scale bar 200 μm. Non-cavitated (E, box) and cavitated structure (F, box).
(G) Quantification of (E) and (F). See also Figure S7.

extraembryonic mesoderm development is needed. This model would not only enable us to study postimplantation functionally but would also improve our understanding of human reproduction and diseases.

Here we discovered that naive hPSC cultures can differentiate into EXMCs that transcriptionally resemble human and monkey embryo EXM. We propose that naive hPSCs could provide a useful model for EXMC specification, enabling the study of a major cell type of the early postimplantation primate embryo that arises in early postimplantation human embryogenesis.

Here we separated EXMCs and hTSCs by sorting. Purified EXMCs provide an ideal model for future studies investigating the function of EXMCs *in vitro*. For example, EXMCs could be used to investigate the ability of human EXM to contribute to the first wave of hematopoiesis in the human embryo (Ross and Boroviak, 2020). Co-culture of EXMCs with other embryonic cell types will help understand the role that EXM plays during human embryogenesis.

The lineage origin of EXMCs in human embryos is unknown (Rossant and Tam, 2022). An epiblast origin of EXMCs in human embryos remains to be tested but is supported by several lines of evidence. Human naive epiblast stem cells can differentiate into EXMCs *in vitro* at least via an intermediate epiblast state. Primate EXMCs first arise in close proximity to the epiblast at a time when the epiblast is in an intermediate state between pre-implantation naive epiblast and post-implantation primed epiblast (Enders and King, 1988; Gasser, 1975; Luckett, 1978; Rostovskaya et al., 2022). Recent monkey embryo scRNA-seq analyses suggest that EXMCs arise from the epiblast (Yang et al., 2021) and EXMCs are specified from the epiblast in mice, although later in development (Saykali et al., 2019).

Recent work linked the extended epiblast peri-implantation period between the naive and primed epiblast states, which lasts 10 days in humans and 2 days in mice, to multiple distinct waves of amniogenesis (Rostovskaya et al., 2022). Similarly, it is possible that different waves of extraembryo mesogenesis (the development of EXM) exist in primates. Indeed, several human epiblast stem cell states appear to be competent for EXMC induction. Human extended pluripotent stem cells in monkey chimeras can give rise to EXMCs (Tan et al., 2021). Additionally, several studies have suggested that primed hPSCs may be competent to generate EXMCs under certain conditions, although this remains to be strictly demonstrated (Io et al., 2021; Markouli et al., 2021; Simunovic et al., 2022; Tietze et al., 2020). More work is needed to understand the mechanisms regulating the plasticity of pluripotent stem cell states in mammals.

While the primate epiblast may have the capacity to activate the EXM developmental program, a TB origin has been proposed (Hertig, 1935; Hertig and Rock, 1941; Hertig et al., 1956). Since EXMCs were also obtained from TROP2+ cells in our experiments, a TB origin of EXMCs, although deemed unlikely in the embryo (Ross and Boroviak, 2020), cannot be excluded. On the other hand, there are several lines of evidence suggesting an origin of EXMCs from the PrE (Bianchi et al., 1993; Enders and King, 1988; Nakamura et al., 2016). Moreover, PrE conversion generated EXMCs. Although the intermediate epiblast route is also possible in these experiments, a PrE origin of EXMCs cannot be excluded. The absence of PrE cells in ASECRiAV cultures suggest it is unlikely that PrE is the sole origin of EXMCs *in vitro*. A combination of origins remains possible and understanding the origin and window of competence for EXMC differentiation in human development remains an exciting prospect for future studies.

We showed that VIM+ EXMCs are present at low frequency (<0.02%) during the generation of human blastoids generated from triply inhibited PXGL hPSCs. Others have shown that the starting cells and medium used to conduct blastoid specification strongly impacts the faithfulness of the model (see proposed features in Kagawa et al., 2021; Zhao et al., 2021). Inadequacies in these parameters cause cells to follow an abnormal sequence of events, specify slowly and inefficiently and generate abnormal lineages and stages (Zhao et al., 2021). Here we show the formation of putative EXM progenitors during the early steps of blastoid organization. A suboptimal blastoid medium or defects during blastoid formation could lead to the precocious activation or acceleration of EXMC specification. However, human blastoids may still be competent to initiate a *bona fide* postimplantation EXMC specification program later, when grown to postimplantation stages, which warrants future investigations.

Altogether, we have discovered that naive hPSC cultures can specify the EXMC fate, which provides a model to molecularly and functionally characterize EXM specification *in vitro*. This system is of particular interest given that in humans, EXM specification takes place after implantation and starts before gastrulation and is therefore inaccessible for experimentation. The induction and maintenance of EXMCs from multiple naive hPSC lines will enable the study of EXM in culture and allow molecular, genetic, and epigenetic manipulations. EXMCs may also allow the development of improved integrated stem cell embryo models in combination with TB, epiblast, and PrE lineage-derived cell types.

Limitations of the study

Limitations of this study include our limited knowledge of human postimplantation embryos, the restricted number of human and monkey embryos that were sequenced, and the scarce molecular characterization of EXMCs in human embryos. More sequencing data of primate embryos are needed for additional, higher resolution analyses. Additionally, all features of embryonic development may not be recapitulated *in vitro* and in embryos grown *ex vivo*.

Another limitation is that the efficiency and timing of EXMC induction under ASECRiAV conditions are variable and may be influenced by genetic background, naive culture conditions, the quality of naive hPSC lines, or other factors. Further optimization may reduce such variability. Despite the advances reported here, the precise origin(s) of EXMCs *in vitro* and in the embryo remains uncertain. We examined the origin of EXMCs during TSC differentiation, but the origin of EXMCs during PrE differentiation remains to be determined. Additionally, the origin of EXMCs may differ *in vitro* and in embryos. Further analyses are required to determine the developmental potential of *in vitro* EXMCs. In blastoids, we have found EXMCs, but the factors contributing to EXMC differentiation in blastoids and the effects of precocious EXMCs on blastoid development remain

unknown. Whether mis-specified cells (e.g., VIM+ cells) may appear in blastocysts is unknown. These factors are important to fully understand EXMC and blastocyst development.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Ethics statement
  - Human embryos extended culture
  - Human blastoids
  - Cell lines
- METHOD DETAILS
  - Immunofluorescence and microscopy
  - Immunohistochemistry and microscopy on blastoids
  - Immunohistochemistry and microscopy on embryos
  - Flow cytometry
  - Western Blot
  - Single-cell RNA sequencing
  - Single-cell RNA-seq analysis
  - Single-cell gene expression analysis of merged datasets
  - Gene regulatory network interference
  - Single-cell ATAC-seq cell preparation and sequencing
  - Single-cell ATAC-seq analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.stem.2022.08.001.

ACKNOWLEDGMENTS

We thank Dr. Rudolf Jaenisch and Dr. Thorold W. Theunissen for the WIBR2-29M-GP26-TN9 hESC line and Dr. Catherine Verfaillie for the Sigma hiPSC line. We are grateful to the Vlaams Supercomputer Center Leuven (https://www.vscentrum.be) for computing, in particular Jan Ooghe; the VIB/KU Leuven Center for Brain and Disease Research, in particular to Kristofer Davie; and Stein Aerts and the Aerts lab for help with bioinformatics. We also thank the Genomics Core Leuven (http://genomicscore.be) for high-throughput sequencing, Jean Christophe Marine's lab for the use of Tapestation and 10X Chromium Controller, and in particular Greet Bervoets for helpful feedback. We are also grateful to Susan Schlenner and the KU Leuven FACS core team for providing the facility and especially Reena Chinaraj who helped us during flow cytometry experiments. Research in the Pasque laboratory was supported by the Research Foundation–Flanders (FWO; Odysseus Return Grant G0F7716N to V.P.; FWO grants G0C9320N and G0B4420N to V.P.), the KU Leuven Research Fund (C1 grant C14/21/19 to V.P.), and FWO PhD fellowships to A.J. (1158318N), I.T. (1S72719N), S.K.T. (1S75720N), L.V. (1S29419N), R.N.A. (11L0722N), and T.X.A.P. (11N3122N). This project has also received funding from the ANR "BOOSTIVF". Laurent David thanks the iPSCDTC and MicroPICell core facilities (endorsed by IBiSA and Biogenouest). The Lluis lab is supported by FWO grant G073622N and KU Leuven Research Fund C14/21/115 to F.L. and FWO PhD fellowships 1S65321N to J.C. Research in the Rivron lab is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (ERC-Co grant agreement No.101002317, "BLASTOID: a discovery platform for early human embryogenesis").

AUTHOR CONTRIBUTIONS

Conceptualization: T.X.A.P., A.P., N.R., B.P.B., and V.P.; data curation: T.X.A.P., A.P., and B.P.B.; formal analysis: T.X.A.P., A.P., and B.P.B.; funding acquisition: F.L., L.D., N.R., and V.P.; investigation (omics experiments): T.X.A.P., A.P., S.K.T., C.E., G.G., R.N.A., A.J., I.T., L.V, B.P.B. and V.P.; investigation (stem cell experiments): A.P., T.X.A.P., H.K., C.E., S.S.F.A.v.K., J.S-H.C., G.G., L.V., J.C., M.O., R.S.T., S.K. and B.P.B.; investigation (blastoid experiments): H.K.; and N.R.; investigation (human embryo staining): A.B., and L.D.; methodology: T.X.A.P., A.P., F.L., L.D., N.R., B.P.B., and V.P.; project administration: T.X.A.P., A.P., N.R., B.P.B., and V.P.; resources: F.L., L.D., N.R., and V.P.; supervision: B.P.B., and V.P.; validation: T.X.A.P., A.P., H.K., C.E., S.S.F.A.v.K., A.B., J.S-H.C., L.V., L.D., N.R., B.P.B., and V.P.; visualization: T.X.A.P., A.P., L.D., N.R., B.P.B., and V.P.; writing, reviewing, and editing of the manuscript: T.X.A.P., A.P., J.R., F.L., L.D., N.R., B.P.B., and V.P.

DECLARATION OF INTERESTS

The Institute for Molecular Biotechnology, Austrian Academy of Sciences has filed patent application EP21151455.9 describing the protocols for human blastoid formation. H.K. and N.R. are the inventors of this patent. All other authors declare no competing interests. J.R. is a member of the *Cell Stem Cell* advisory board.

INCLUSION AND DIVERSITY

One or more of the authors of this paper self-identifies as a member of the LGBTQ+ community.

Received: December 21, 2021
Revised: June 8, 2022
Accepted: August 5, 2022
Published: September 1, 2022

REFERENCES

Aibar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.-C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. Nat. Methods *14*, 1083–1086. https://doi.org/10.1038/nmeth.4463.

Angerer, P., Haghverdi, L., Büttner, M., Theis, F.J., Marr, C., and Buettner, F. (2016). *destiny*: diffusion maps for large-scale single-cell data in R. Bioinformatics *32*, 1241–1243. https://doi.org/10.1093/bioinformatics/btv715.

Arnold, S.J., and Robertson, E.J. (2009). Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. Nat. Rev. Mol. Cell Biol. *10*, 91–103. https://doi.org/10.1038/nrm2618.

Arora, S., Morgan, M., Carlson, M., and Pagès, H. (2022). GenomeInfoDb: Utilities for manipulating chromosome names, including modifying them to follow a particular naming style.

Bianchi, D.W., Wilkins-Haug, L.E., Enders, A.C., and Hay, E.D. (1993). Origin of extraembryonic mesoderm in experimental animals: Relevance to chorionic mosaicism in humans. Am. J. Med. Genet. *46*, 542–550. https://doi.org/10.1002/ajmg.1320460517.

Bonnans, C., Chou, J., and Werb, Z. (2014). Remodelling the extracellular matrix in development and disease. Nat. Rev. Mol. Cell Biol. *15*, 786–801. https://doi.org/10.1038/nrm3904.

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. https://doi.org/10.1242/dev.167833.

Bredenkamp, N., Stirparo, G.G., Nichols, J., Smith, A., and Guo, G. (2019a). The Cell-Surface Marker Sushi Containing Domain 2 Facilitates Establishment of Human Naive Pluripotent Stem Cells. Stem Cell Rep. *12*, 1212–1222. https://doi.org/10.1016/j.stemcr.2019.03.014.

Bredenkamp, N., Yang, J., Clarke, J., Stirparo, G.G., von Meyenn, F., Dietmann, S., Baker, D., Drummond, R., Ren, Y., Li, D., et al. (2019b). Wnt Inhibition Facilitates RNA-Mediated Reprogramming of Human Somatic Cells to Naive Pluripotency. Stem Cell Rep. *13*, 1083–1098. https://doi.org/10.1016/j.stemcr.2019.10.009.

Budi, E.H., Muthusamy, B.-P., and Derynck, R. (2015). The insulin response integrates increased TGF-β signaling through Akt-induced enhancement of cell surface delivery of TGF-β receptors. Sci. Signal. *8*, ra96. https://doi.org/10.1126/scisignal.aaa9432.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. *36*, 411–420. https://doi.org/10.1038/nbt.4096.

Carlson, M. (2019). org.Hs.eg.db: Genome wide annotation for Human.

Castel, G., Meistermann, D., Bretin, B., Firmin, J., Blin, J., Loubersac, S., Bruneau, A., Chevolleau, S., Kilens, S., Chariau, C., et al. (2020). Induction of Human Trophoblast Stem Cells from Somatic Cells and Pluripotent Stem Cells. Cell Rep. *33*, 108419. https://doi.org/10.1016/j.celrep.2020.108419.

Chen, Q., Lee, C.-E., Denard, B., and Ye, J. (2014). Sustained induction of collagen synthesis by TGF-β requires regulated intramembrane proteolysis of CREB3L1. PLoS One *9*, e108528. https://doi.org/10.1371/journal.pone.0108528.

Chhabra, S., and Warmflash, A. (2021). BMP-treated human embryonic stem cells transcriptionally resemble amnion cells in the monkey embryo. Biol. Open *10*, bio058617. https://doi.org/10.1242/bio.058617.

Cinkornpumin, J.K., Kwon, S.Y., Guo, Y., Hossain, I., Sirois, J., Russett, C.S., Tseng, H.-W., Okae, H., Arima, T., Duchaine, T.F., et al. (2020). Naive Human Embryonic Stem Cells Can Give Rise to Cells with a Trophoblast-like Transcriptome and Methylome. Stem Cell Rep. *15*, 198–213. https://doi.org/10.1016/j.stemcr.2020.06.003.

Cui, G., Feng, S., Yan, Y., Wang, L., He, X., Li, X., Duan, Y., Chen, J., Tam, P.P.L., Tang, K., et al. (2022). Spatial and Molecular Anatomy of Germ Layers in the Gastrulating Cynomolgus Monkey Embryo.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., et al. (2021). Twelve years of SAMtools and BCFtools. GigaScience *10*. https://doi.org/10.1093/gigascience/giab008.

Davidson, E., and Levin, M. (2005). Gene regulatory networks. Proc. Natl. Acad. Sci. USA *102*, 4935. https://doi.org/10.1073/pnas.0502024102.

Deglincerti, A., Croft, G.F., Pietila, L.N., Zernicka-Goetz, M., Siggia, E.D., and Brivanlou, A.H. (2016). Self-organization of the in vitro attached human embryo. Nature *533*, 251–254. https://doi.org/10.1038/nature17948.

Derynck, R., Zhang, Y., and Feng, X.H. (1998). Smads: transcriptional activators of TGF-beta responses. Cell *95*, 737–740. https://doi.org/10.1016/s0092-8674(00)81696-7.

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. https://doi.org/10.1093/bioinformatics/bts635.

Dong, C., Beltcheva, M., Gontarz, P., Zhang, B., Popli, P., Fischer, L.A., Khan, S.A., Park, K.-M., Yoon, E.-J., Xing, X., et al. (2020). Derivation of trophoblast stem cells from naïve human pluripotent stem cells. Elife *9*, e52504. https://doi.org/10.7554/eLife.52504.

Dong, C., Fu, S., Karvas, R.M., Chew, B., Fischer, L.A., Xing, X., Harrison, J.K., Popli, P., Kommagani, R., Wang, T., et al. (2022). A genome-wide CRISPR-Cas9 knockout screen identifies essential and growth-restricting genes in human trophoblast stem cells. Nat. Commun. *13*, 2548. https://doi.org/10.1038/s41467-022-30207-9.

Dongre, A., and Weinberg, R.A. (2019). New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer. Nat. Rev. Mol. Cell Biol. *20*, 69–84. https://doi.org/10.1038/s41580-018-0080-4.

Enders, A.C., and King, B.F. (1988). Formation and differentiation of extraembryonic mesoderm in the rhesus monkey. Am. J. Anat. *181*, 327–340. https://doi.org/10.1002/aja.1001810402.

Fan, Y., Min, Z., Alsolami, S., Ma, Z., Zhang, E., Chen, W., Zhong, K., Pei, W., Kang, X., Zhang, P., et al. (2021). Generation of human blastocyst-like structures from pluripotent stem cells. Cell Discov. *7*, 81. https://doi.org/10.1038/s41421-021-00316-8.

Garnier, S., Ross, N., Rudis, B., Filipovic-Pierucci, A., Galili, T., timelyportfolio, Greenwell, B., Sievert, C., Harris, D.J., and Chen, J.J. (2021). sjmgarnier/viridis: viridis 0.6.0 (pre-CRAN release) (Zenodo).

Gasser, R.F. (1975). Atlas of Human Embryos (HarperCollins Publishers).

Guo, G., Stirparo, G.G., Strawbridge, S.E., Spindlow, D., Yang, J., Clarke, J., Dattani, A., Yanagida, A., Li, M.A., Myers, S., et al. (2021). Human naive epiblast cells possess unrestricted lineage potential. Cell Stem Cell *28*, 1040–1056.e6. https://doi.org/10.1016/j.stem.2021.02.025.

Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., 3rd, Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573–3587.e29. https://doi.org/10.1016/j.cell.2021.04.048.

Henry, L., and Wickham, H. (2022). Purrr: Functional Programming Tools.

Hertig, A.T. (1935). Angiogenesis in the Early Human Chorion and in the Primary Placenta of the Macaque Monkey.

Hertig, A.T., and Rock, J.C. (1941). Two Human Ova of the Pre-villous Stage: Having an Ovulation Age of about Eleven and Twelve Days Respectively.

Hertig, A.T., Rock, J., and Adams, E.C. (1956). A description of 34 human ova within the first 17 days of development. Am. J. Anat. *98*, 435–493. https://doi.org/10.1002/aja.1000980306.

Hill, J.P. (1932). Croonian lecture. - the developmental history of the primates. Philos. Trans. R. Soc. Lond. - Ser. B Contain. Pap. a Biol. Character *221*, 45–178. https://doi.org/10.1098/rstb.1932.0002.

Io, S., Kabata, M., Iemura, Y., Semi, K., Morone, N., Minagawa, A., Wang, B., Okamoto, I., Nakamura, T., Kojima, Y., et al. (2021). Capturing human trophoblast development with naive pluripotent stem cells in vitro. Cell Stem Cell *28*, 1023–1039.e13. https://doi.org/10.1016/j.stem.2021.03.013.

Kagawa, H., Javali, A., Khoei, H.H., Sommer, T.M., Sestini, G., Novatchkova, M., Scholte op Reimer, Y., Castel, G., Bruneau, A., Maenhoudt, N., et al. (2022). Human blastoids model blastocyst development and implantation. Nature *601*, 600–605. https://doi.org/10.1038/s41586-021-04267-8.

Karvas, R.M., Khan, S.A., Verma, S., Yin, Y., Kulkarni, D., Dong, C., Park, K.-M., Chew, B., Sane, E., Fischer, L.A., et al. (2022). Stem-cell-derived trophoblast organoids model human placental development and susceptibility to emerging pathogens. Cell Stem Cell *29*, 810–825.e8. https://doi.org/10.1016/j.stem.2022.04.004.

Kimmelman, J., Hyun, I., Benvenisty, N., Caulfield, T., Heslop, H.E., Murry, C.E., Sipp, D., Studer, L., Sugarman, J., and Daley, G.Q. (2016). Policy: Global standards for stem-cell research. Nature *533*, 311–313. https://doi.org/10.1038/533311a.

Kinder, S.J., Tsang, T.E., Quinlan, G.A., Hadjantonakis, A.K., Nagy, A., and Tam, P.P. (1999). The orderly allocation of mesodermal cells to the extraembryonic structures and the anteroposterior axis during gastrulation of the mouse embryo. Development *126*, 4691–4701.

Kinoshita, M., Barber, M., Mansfield, W., Cui, Y., Spindlow, D., Stirparo, G.G., Dietmann, S., Nichols, J., and Smith, A. (2021). Capture of Mouse and Human Stem Cells with Features of Formative Pluripotency. Cell Stem Cell *28*, 2180. https://doi.org/10.1016/j.stem.2021.11.002.

Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R., and Pfister, H. (2014). UpSet: Visualization of Intersecting Sets. IEEE Trans. Vis. Comput. Graph. *20*, 1983–1992. https://doi.org/10.1109/TVCG.2014.2346248.

Li, Y., Yan, J., Chang, H.-M., Chen, Z.-J., and Leung, P.C.K. (2021). Roles of TGF-β Superfamily Proteins in Extravillous Trophoblast Invasion. Trends Endocrinol. Metab. *32*, 170–189. https://doi.org/10.1016/j.tem.2020.12.005.

Linneberg-Agerholm, M., Wong, Y.F., Romero Herrera, J.A., Monteiro, R.S., Anderson, K.G.V., and Brickman, J.M. (2019). Naïve human pluripotent stem cells respond to Wnt, Nodal and LIF signalling to produce expandable naïve

extra-embryonic endoderm. Development *146*, dev180620. https://doi.org/10.1242/dev.180620.

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. Nat. Methods *14*, 1055–1062. https://doi.org/10.1038/nmeth.4436.

Liu, X., Tan, J.P., Schröder, J., Aberkane, A., Ouyang, J.F., Mohenska, M., Lim, S.M., Sun, Y.B.Y., Chen, J., Sun, G., et al. (2021). Modelling human blastocysts by reprogramming fibroblasts into iBlastoids. Nature *591*, 627–632. https://doi.org/10.1038/s41586-021-03372-y.

Lovell-Badge, R., Anthony, E., Barker, R.A., Bubela, T., Brivanlou, A.H., Carpenter, M., Charo, R.A., Clark, A., Clayton, E., Cong, Y., et al. (2021). ISSCR Guidelines for Stem Cell Research and Clinical Translation: The 2021 update. Stem Cell Rep. *16*, 1398–1408. https://doi.org/10.1016/j.stemcr.2021.05.012.

Luckett, W.P. (1978). Origin and differentiation of the yolk sac and extraembryonic mesoderm in presomite human and rhesus monkey embryos. Am. J. Anat. *152*, 59–97. https://doi.org/10.1002/aja.1001520106.

Ma, H., Zhai, J., Wan, H., Jiang, X., Wang, X., Wang, L., Xiang, Y., He, X., Zhao, Z.-A., Zhao, B., et al. (2019). In vitro culture of cynomolgus monkey embryos beyond early gastrulation. Science *366*, eaax7890. https://doi.org/10.1126/science.aax7890.

Markouli, C., De Deckersberg, E.C., Dziedzicka, D., Regin, M., Franck, S., Keller, A., Gheldof, A., Geens, M., Sermon, K., and Spits, C. (2021). Sustained intrinsic WNT and BMP4 activation impairs hESC differentiation to definitive endoderm and drives the cells towards extra-embryonic mesoderm. Sci. Rep. *11*, 8242. https://doi.org/10.1038/s41598-021-87547-7.

Mauri, F., Schepkens, C., Lapouge, G., Drogat, B., Song, Y., Pastushenko, I., Rorive, S., Blondeau, J., Golstein, S., Bareche, Y., et al. (2021). NR2F2 controls malignant squamous cell carcinoma state by promoting stemness and invasion and repressing differentiation. Nat. Cancer *2*, 1152–1169. https://doi.org/10.1038/s43018-021-00287-5.

Mazid, M.A., Ward, C., Luo, Z., Liu, C., Li, Y., Lai, Y., Wu, L., Li, J., Jia, W., Jiang, Y., et al. (2022). Rolling back human pluripotent stem cells to an eight-cell embryo-like stage. Nature *605*, 315–324. https://doi.org/10.1038/s41586-022-04625-0.

Meistermann, D., Bruneau, A., Loubersac, S., Reignier, A., Firmin, J., François-Campion, V., Kilens, S., Lelièvre, Y., Lammers, J., Feyeux, M., et al. (2021). Integrated pseudotime analysis of human pre-implantation embryo single-cell transcriptomes reveals the dynamics of lineage specification. Cell Stem Cell *28*, 1625–1640.e6. https://doi.org/10.1016/j.stem.2021.04.027.

Messmer, T., von Meyenn, F., Savino, A., Santos, F., Mohammed, H., Lun, A.T.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.e4. https://doi.org/10.1016/j.celrep.2018.12.099.

Moris, N., Anlas, K., van den Brink, S.C., Alemany, A., Schröder, J., Ghimire, S., Balayo, T., van Oudenaarden, A., and Martinez Arias, A. (2020). An in vitro model of early anteroposterior organization during human development. Nature *582*, 410–415. https://doi.org/10.1038/s41586-020-2383-9.

Müller, K., and Wickham, H. (2022). Tibble: Simple Data Frames.

Nahaboo, W., Eski, S.E., Despin-Guitard, E., Vermeersch, M., Versaevel, M., Saykali, B., Monteyne, D., Gabriele, S., Magin, T.M., Schwarz, N., et al. (2022). Keratin filaments mediate the expansion of extra-embryonic membranes in the post-gastrulation mouse embryo. EMBO J. *41*, e108747. https://doi.org/10.15252/embj.2021108747.

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. https://doi.org/10.1038/nature19096.

Niu, Y., Sun, N., Li, C., Lei, Y., Huang, Z., Wu, J., Si, C., Dai, X., Liu, C., Wei, J., et al. (2019). Dissecting primate early post-implantation development using long-term in vitro embryo culture. Science *366*, eaaw5754. https://doi.org/10.1126/science.aaw5754.

Okae, H., Toh, H., Sato, T., Hiura, H., Takahashi, S., Shirane, K., Kabayama, Y., Suyama, M., Sasaki, H., and Arima, T. (2018). Derivation of human trophoblast stem cells. Cell Stem Cell *22*, 50–63.e6. https://doi.org/10.1016/j.stem.2017.11.004.

O'Rahilly, R., and Müller, F. (1987). Stages in early human development. Future Aspects in Human In Vitro Fertilization, 238–244. https://doi.org/10.1007/978-3-642-71412-2_33.

Osnato, A., Brown, S., Krueger, C., Andrews, S., Collier, A.J., Nakanoh, S., Quiroga Londoño, M., Wesley, B.T., Muraro, D., Brumm, A.S., et al. (2021). TGFβ signalling is required to maintain pluripotency of human naïve pluripotent stem cells. Elife *10*, e67259. https://doi.org/10.7554/eLife.67259.

Pasque, V., Tchieu, J., Karnik, R., Uyeda, M., Sadhu Dimashkie, A., Case, D., Papp, B., Bonora, G., Patel, S., Ho, R., et al. (2014). X chromosome reactivation dynamics reveal stages of reprogramming to pluripotency. Cell *159*, 1681–1697. https://doi.org/10.1016/j.cell.2014.11.040.

Peng, G., Suo, S., Chen, J., Chen, W., Liu, C., Yu, F., Wang, R., Chen, S., Sun, N., Cui, G., et al. (2020). Spatial transcriptome for the molecular annotation of lineage fates and cell identity in mid-gastrula mouse embryo. Dev. Cell *55*, 802–804. https://doi.org/10.1016/j.devcel.2020.11.018.

Pera, M.F., and Rossant, J. (2021). The exploration of pluripotency space: Charting cell state transitions in peri-implantation development. Cell Stem Cell *28*, 1896–1906. https://doi.org/10.1016/j.stem.2021.10.001.

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. https://doi.org/10.1016/j.cell.2016.03.023.

Posfai, E., Schell, J.P., Janiszewski, A., Rovic, I., Murray, A., Bradshaw, B., Yamakawa, T., Pardon, T., El Bakkali, M., Talon, I., et al. (2021). Evaluating totipotency using criteria of increasing stringency. Nat. Cell Biol. *23*, 49–60. https://doi.org/10.1038/s41556-020-00609-2.

Qin, Q., Xu, Y., He, T., Qin, C., and Xu, J. (2012). Normal and disease-related biological functions of Twist1 and underlying molecular mechanisms. Cell Res. *22*, 90–106. https://doi.org/10.1038/cr.2011.144.

Rhee, C., Edwards, M., Dang, C., Harris, J., Brown, M., Kim, J., and Tucker, H.O. (2017). ARID3A is required for mammalian placenta development. Dev. Biol. *422*, 83–91. https://doi.org/10.1016/j.ydbio.2016.12.003.

Rivron, N.C., Vrij, E.J., Rouwkema, J., Le Gac, S., van den Berg, A., Truckenmüller, R.K., and van Blitterswijk, C.A. (2012). Tissue deformation spatially modulates VEGF signaling and angiogenesis. Proc. Natl. Acad. Sci. USA *109*, 6886–6891. https://doi.org/10.1073/pnas.1201626109.

Ross, C., and Boroviak, T.E. (2020). Origin and function of the yolk sac in primate embryogenesis. Nat. Commun. *11*, 3760. https://doi.org/10.1038/s41467-020-17575-w.

Rossant, J., and Tam, P.P.L. (2022). Early human embryonic development: Blastocyst formation to gastrulation. Dev. Cell *57*, 152–165. https://doi.org/10.1016/j.devcel.2021.12.022.

Rostovskaya, M., Andrews, S., Reik, W., and Rugg-Gunn, P.J. (2022). Amniogenesis occurs in two independent waves in primates. Cell Stem Cell *29*, 744–759.e6. https://doi.org/10.1016/j.stem.2022.03.014.

RStudio Team (2020). RStudio: Integrated Development for R.

Sadler, T.W. (2012). Langman's Medical Embryology (Lippincott Williams & Wilkins).

Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nat. Biotechnol. *33*, 495–502. https://doi.org/10.1038/nbt.3192.

Satpathy, A.T., Granja, J.M., Yost, K.E., Qi, Y., Meschi, F., McDermott, G.P., Olsen, B.N., Mumbach, M.R., Pierce, S.E., Corces, M.R., et al. (2019). Massively parallel single-cell chromatin landscapes of human immune cell development and intratumoral T cell exhaustion. Nat. Biotechnol. *37*, 925–936. https://doi.org/10.1038/s41587-019-0206-z.

Saykali, B., Mathiah, N., Nahaboo, W., Racu, M.-L., Hammou, L., Defrance, M., and Migeotte, I. (2019). Distinct mesoderm migration phenotypes in

extra-embryonic and embryonic regions of the early mouse embryo. Elife *8*, e42434. https://doi.org/10.7554/eLife.42434.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods *9*, 671–675. https://doi.org/10.1038/nmeth.2089.

Shahbazi, M.N., and Zernicka-Goetz, M. (2018). Deconstructing and reconstructing the mouse and human early embryo. Nat. Cell Biol. *20*, 878–887. https://doi.org/10.1038/s41556-018-0144-x.

Shahbazi, M.N., Jedrusik, A., Vuoristo, S., Recher, G., Hupalowska, A., Bolton, V., Fogarty, N.N.M., Campbell, A., Devito, L., Ilic, D., et al. (2016). Self-organization of the human embryo in the absence of maternal tissues. Nat. Cell Biol. *18*, 700–708. https://doi.org/10.1038/ncb3347.

Shao, Y., Taniguchi, K., Townshend, R.F., Miki, T., Gumucio, D.L., and Fu, J. (2017). A pluripotent stem cell-based model for post-implantation human amniotic sac development. Nat. Commun. *8*, 208. https://doi.org/10.1038/s41467-017-00236-w.

Shepard, T.H. (1989). Developmental stages in human embryos. R. O'Rahilly and F. Müller (eds), Carnegie Institution of Washington, Washington, DC, 1987, 306 pp., $52. Teratology *40*. 85–85. https://doi.org/10.1002/tera.1420400111.

Shi, Y., and Massagué, J. (2003). Mechanisms of TGF-beta signaling from cell membrane to the nucleus. Cell *113*, 685–700. https://doi.org/10.1016/s0092-8674(03)00432-x.

Simunovic, M., Siggia, E.D., and Brivanlou, A.H. (2022). In vitro attachment and symmetry breaking of a human embryo model assembled from primed embryonic stem cells. Cell Stem Cell *29*, 962–972.e4. https://doi.org/10.1016/j.stem.2022.05.001.

Sozen, B., Jorgensen, V., Weatherbee, B.A.T., Chen, S., Zhu, M., and Zernicka-Goetz, M. (2021). Reconstructing aspects of human embryogenesis with pluripotent stem cells. Nat. Commun. *12*, 5550. https://doi.org/10.1038/s41467-021-25853-4.

Spencer Chapman, M., Ranzoni, A.M., Myers, B., Williams, N., Coorens, T.H.H., Mitchell, E., Butler, T., Dawson, K.J., Hooks, Y., Moore, L., et al. (2021). Lineage tracing of human development through somatic mutations. Nature *595*, 85–90. https://doi.org/10.1038/s41586-021-03548-6.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., 3rd, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell *177*, 1888–1902.e21. https://doi.org/10.1016/j.cell.2019.05.031.

Stuart, T., Srivastava, A., Madad, S., Lareau, C.A., and Satija, R. (2021). Single-cell chromatin state analysis with Signac. Nat. Methods *18*, 1333–1341. https://doi.org/10.1038/s41592-021-01282-5.

Tam, P.P., and Beddington, R.S. (1987). The formation of mesodermal tissues in the mouse embryo during gastrulation and early organogenesis. Development *99*, 109–126.

Tan, T., Wu, J., Si, C., Dai, S., Zhang, Y., Sun, N., Zhang, E., Shao, H., Si, W., Yang, P., et al. (2021). Chimeric contribution of human extended pluripotent stem cells to monkey embryos ex vivo. Cell *184*, 3589. https://doi.org/10.1016/j.cell.2021.06.011.

Taniguchi, C.M., Emanuelli, B., and Kahn, C.R. (2006). Critical nodes in signalling pathways: insights into insulin action. Nat. Rev. Mol. Cell Biol. *7*, 85–96. https://doi.org/10.1038/nrm1837.

Taubenschmid-Stowers, J., Rostovskaya, M., Santos, F., Ljung, S., Argelaguet, R., Krueger, F., Nichols, J., and Reik, W. (2022). 8C-like cells capture the human zygotic genome activation program in vitro. Cell Stem Cell *29*, 449–459.e6. https://doi.org/10.1016/j.stem.2022.01.014.

Team TBD (2021). BSgenome.Hsapiens.UCSC.hg38: Full genome sequences for Homo sapiens (UCSC version hg38, based on GRCh38.p13).

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*, 524–526. https://doi.org/10.1016/j.stem.2014.09.003.

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. https://doi.org/10.1016/j.stem.2016.06.011.

Tietze, E., Barbosa, A.R., Euclydes, V., Cho, H.J., Lee, Y.K., Feltrin, A., van de Leemput, J., Di Carlo, P., Sawada, T., Benjamin, K.J., et al. (2020). Single-cell analysis of human trophoblast stem cell specification reveals activation of fetal cytotrophoblast expression programs including coronavirus associated host factors and human endogenous retroviruses. Preprint at bioRxiv. https://doi.org/10.1101/2020.08.29.273425.

Tyser, R.C.V., Mahammadov, E., Nakanoh, S., Vallier, L., Scialdone, A., and Srinivas, S. (2021). Single-cell transcriptomic characterization of a gastrulating human embryo. Nature *600*, 285–289. https://doi.org/10.1038/s41586-021-04158-y.

Van de Sande, B., Flerin, C., Davie, K., De Waegeneer, M., Hulselmans, G., Aibar, S., Seurinck, R., Saelens, W., Cannoodt, R., Rouchon, Q., et al. (2020). A scalable SCENIC workflow for single-cell gene regulatory network analysis. Nat. Protoc. *15*, 2247–2276. https://doi.org/10.1038/s41596-020-0336-2.

Veenvliet, J.V., Bolondi, A., Kretzmer, H., Haut, L., Scholze-Wittler, M., Schifferl, D., Koch, F., Guignard, L., Kumar, A.S., Pustet, M., et al. (2020). Mouse embryonic stem cells self-organize into trunk-like structures with neural tube and somites. Science *370*, eaba4937. https://doi.org/10.1126/science.aba4937.

Vrij, E.J., Espinoza, S., Heilig, M., Kolew, A., Schneider, M., van Blitterswijk, C.A., Truckenmüller, R.K., and Rivron, N.C. (2016). 3D high throughput screening and profiling of embryoid bodies in thermoformed microwell plates. Lab Chip *16*, 734–742. https://doi.org/10.1039/c5lc01499a.

Wakefield, L.M., and Hill, C.S. (2013). Beyond TGFβ: roles of other TGFβ superfamily members in cancer. Nat. Rev. Cancer *13*, 328–341. https://doi.org/10.1038/nrc3500.

Wamaitha, S.E., Grybel, K.J., Alanis-Lobato, G., Gerri, C., Ogushi, S., McCarthy, A., Mahadevaiah, S.K., Healy, L., Lea, R.A., Molina-Arcas, M., et al. (2020). IGF1-mediated human embryonic stem cell self-renewal recapitulates the embryonic niche. Nat. Commun. *11*, 764. https://doi.org/10.1038/s41467-020-14629-x.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., et al. (2019). Welcome to the Tidyverse. J. Open Source Softw. *4*, 1686. https://doi.org/10.21105/joss.01686.

Wojdyla, K., Collier, A.J., Fabian, C., Nisi, P.S., Biggins, L., Oxley, D., and Rugg-Gunn, P.J. (2020). Cell-Surface Proteomics Identifies Differences in Signaling and Adhesion Protein Expression between Naive and Primed Human Pluripotent Stem Cells. Stem Cell Rep. *14*, 972–988. https://doi.org/10.1016/j.stemcr.2020.03.017.

Xiang, L., Yin, Y., Zheng, Y., Ma, Y., Li, Y., Zhao, Z., Guo, J., Ai, Z., Niu, Y., Duan, K., et al. (2020). A developmental landscape of 3D-cultured human pre-gastrulation embryos. Nature *577*, 537–542. https://doi.org/10.1038/s41586-019-1875-y.

Yanagida, A., Spindlow, D., Nichols, J., Dattani, A., Smith, A., and Guo, G. (2021). Naive stem cell blastocyst model captures human embryo lineage segregation. Cell Stem Cell *28*, 1016–1022.e4. https://doi.org/10.1016/j.stem.2021.04.031.

Yang, R., Goedel, A., Kang, Y., Si, C., Chu, C., Zheng, Y., Chen, Z., Gruber, P.J., Xiao, Y., Zhou, C., et al. (2021). Amnion signals are essential for mesoderm formation in primates. Nat. Commun. *12*, 5126. https://doi.org/10.1038/s41467-021-25186-2.

Yu, L., Wei, Y., Duan, J., Schmitz, D.A., Sakurai, M., Wang, L., Wang, K., Zhao, S., Hon, G.C., and Wu, J. (2021). Blastocyst-like structures generated from human pluripotent stem cells. Nature *591*, 620–626. https://doi.org/10.1038/s41586-021-03356-y.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137. https://doi.org/10.1186/gb-2008-9-9-r137.

Zhao, C., Reyes, A.P., Schell, J.P., Weltner, J., Ortega, N., Zheng, Y., Björklund, Å.K., Rossant, J., Fu, J., Petropoulos, S., et al. (2021). Reprogrammed iBlastoids contain amnion-like cells but not trophectoderm. Preprint at bioRxiv. https://doi.org/10.1101/2021.05.07.442980.

Zheng, G.X.Y., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., et al. (2017). Massively parallel digital transcriptional profiling of single cells. Nat. Commun. *8*, 14049. https://doi.org/10.1038/ncomms14049.

Zheng, Y., Xue, X., Shao, Y., Wang, S., Esfahani, S.N., Li, Z., Muncie, J.M., Lakins, J.N., Weaver, V.M., Gumucio, D.L., and Fu, J. (2019). Controlled modelling of human epiblast and amnion development using stem cells. Nature *573*, 421–425. https://doi.org/10.1038/s41586-019-1535-2.

Zhou, F., Wang, R., Yuan, P., Ren, Y., Mao, Y., Li, R., Lian, Y., Li, J., Wen, L., Yan, L., et al. (2019). Reconstituting the transcriptome and DNA methylome landscapes of human implantation. Nature *572*, 660–664. https://doi.org/10.1038/s41586-019-1500-0.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit monoclonal IgG anti-human BST2	Abcam	Cat#ab243230; RRID: AB_2915925
Mouse monoclonal IgG1 anti-human VIM (Vimentin)	Abcam	Cat#ab8978; RRID: AB_306907
Rat monoclonal IgG2b anti-human GATA3	Thermo	Cat#14-9966-82; RRID: AB_1210519
Rabbit monoclonal IgG anti-human NR2F2	Abcam	Cat#ab211776; RRID: AB_2893028
Rabbit monoclonal IgG anti-human E-Cadherin (CDH-1) (24E10) rabbit mAb	Cell signaling	Cat#3195; RRID: AB_2291471
Mouse monoclonal IgG1 anti-human APC CD324 (E-Cadherin)	Biolegend	Cat#324108; RRID: AB_756069
Mouse monoclonal IgG1 anti-human VIM (Vimentin)	Dako	Cat#M072501-2; RRID: AB_10013485
Rat monoclonal IgG2a anti-human SOX2	Invitrogen	Cat#14-9811-80; RRID: AB_11219070
Rabbit monoclonal IgG anti-human GATA2	Abcam	Cat#ab109241; RRID: AB_10865130
Rat monoclonal IgG2a anti-human GATA4	Invitrogen	Cat#14-9980-82; RRID: AB_763541
Rabbit polyclonal IgG anti-human DCN	Abcam	Cat#ab151988; RRID: AB_2915927
Rabbit monoclonal IgG anti-human LUM	Thermo	Cat#MA5-29402; RRID: AB_2785270
Goat polyclonal IgG anti-human SOX17	R&D System biotechne	Cat#AF1924; RRID: AB_355060
Rabbit monoclonal IgG anti-human SMAD5 (phospho S463 + S465) antibody [EP728(2)AY]	Abcam	Cat#ab76296; RRID: AB_1524420
Rabbit monoclonal IgG anti-human Anti-4E-BP1, phospho (Thr37/Thr46)	Cell signaling	Cat#2855 S; RRID: AB_560835
Rabbit monoclonal IgG anti-human Anti-Periostin (POSTN)	Abcam	Cat#ab14041; RRID: AB_2299859
Mouse monoclonal IgG2A Anti-Human Anti-Trop-2	R&D system	Cat#MAB650; RRID: AB_2205665
Mouse monoclonal IgG Anti-Human Nanog	BD Biosciences	Cat#560482; RRID: AB_1645598
Rabbit IgG monoclonal Anti-human Phospho-Smad2 (Ser465/467) (138D4)	Cell signaling technology	Cat#3108; RRID: AB_490941
Mouse IgG2b monoclonal Anti-human Anti-Smad2	Abcam	Cat#ab71109; RRID: AB_1281120
Rabbit monoclonal IgG Anti-human Anti-SMAD5	Abcam	Cat#ab40771; RRID: AB_777981
Chicken IgG monoclonal Anti-human β-Actin Antibody	Santa Cruz Biotechnology	Cat#sc-47778; RRID: AB_626632
Mouse monoclonal IgG1 Anti-human APC-SUSD2	Miltenyi Biotec	Cat#130-121-134; RRID: AB_2752220
Donkey IgG anti-rabbit (H+L) polyclonal secondary antibody, Alexa Fluor 647 conjugated	Thermo	Cat#A-31573 c; RRID: AB_2536183
Donkey Anti-Mouse IgG (H+L) polyclonal secondary Antibody, Alexa Fluor 488 Conjugated	Thermo Fisher Scientific	Cat# A21202; RRID: AB_141607
Donkey Anti-Rat IgG H&L (Alexa Fluor® 647) Polyclonal preadsorbed secondary Antibody	Abcam	Cat# ab150155; RRID: AB_2813835

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Donkey anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Polyclonal secondary antibody, Alexa Fluor 568	Thermo Fisher Scientific	Cat#A10042; RRID: AB_2534017
Donkey anti-Rabbit IgG (H+L) Highly Cross-Adsorbed polyclonal secondary Antibody, Alexa Fluor 488	Thermo Fisher Scientific	Cat#A21206; RRID: AB_2535792
Donkey anti-Rat IgG (H+L) Highly Cross-Adsorbed polyclonal secondary Antibody, Alexa Fluor 488	Thermo Fisher Scientific	Cat#A21208; RRID: AB_2535794
Donkey anti-Rat IgG (H+L) Cross-Adsorbed polyclonal secondary Antibody, DyLight 550	Thermo Fisher Scientific	Cat#SA5-10027; RRID: AB_2556607
Donkey anti-Mouse IgG (H+L) Highly Cross-Adsorbed polyclonal secondary Antibody, Alexa Fluor Plus 647	Thermo Fisher Scientific	Cat#A32787; RRID: AB_2762830
Chicken anti-Rat IgG (H+L) Cross-Adsorbed polyclonal secondary Antibody, Alexa Fluor Plus 647	Thermo Fisher Scientific	Cat#A21472; RRID :AB_2535875
Chemicals, peptides, and recombinant proteins		
MAPK inhibitor PD0325901	Axon Medchem	Cat#Axon 1408
Tankyrase inhibitor XAV939	Sigma-Aldrich	Cat#X3004
aPKC inhibitor Gö6983	Tocris	Cat#2285
Recombinant human leukemia inhibitory factor LIF	Peprotech	Cat#300-05
GSK3 inhibitor IM-12	Enzo Life Sciences	Cat# BML-WN102
B-Raf inhibitor SB590885	R&D systems	Cat# 2650/10
Src inhibitor WH-4-023	Advanced ChemTech	Cat# H620061
Insulin-Transferrin-Selenium-Ethanolamine (ITS -X) (100X)	Thermo Fisher Scientific	Cat#51500056
ALK4/5/7 Activin/Nodal/TDGF-β inhibitor A83-01	Peprotech	Cat#9094360
1-Oleoyl Lysophosphatidic Acid	R&D systems	Cat#3854
ALK-5 Activin/Nodal/TDGF-β inhibitor SB431542	Axon medchem	Cat#1661
Valproic acid	Sigma-Aldrich	Cat#V0033000
hEGF	Miltenyi Biotec	Cat#130-097-750
GSK-3 inhibitor CHIR99021	Axon medchem	Cat#Axon 1386
L-ascorbic acid	Sigma-Aldrich	Cat#A8960
Bovine serum albumin	Sigma-Aldrich	Cat#A3059
Bovine serum albumin	Sigma-Aldrich	Cat#A9418
DMEM/F12, HEPES	Thermo Fisher Scientific	Cat#31330038
Fetal Bovine serum	Thermo Fisher Scientific	Cat#10270-106
KSR	Thermo Fisher Scientific	Cat#10828028
FGF2	Peprotech	Cat#100-18C
Penicillin-streptomycin	Thermo Fisher Scientific	Cat#15140-122
ROCK inhibitor Y-27632 dihydrochloride	Tocris	Cat#1254
2-Mercaptoethanol	Thermo Fisher Scientific	Cat#31350-010
Neurobasal™ Medium	Thermo Fisher Scientific	Cat#21103-049
L-Glutamine	Thermo Fisher Scientific	Cat#25030081
MEM Non-Essential Amino Acids Solution (100X)	Thermo Fisher Scientific	Cat#11140050
N-2 Supplement (100X)	Thermo Fisher Scientific	Cat#17502048
B-27™ Supplement (50X), serum free	Thermo Fisher Scientific	Cat#17504044

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Accutase solution	Sigma-Aldrich	Cat#A6964
TrypLE™	Thermo Fisher Scientific	Cat#12605010
Collagen IV	Corning	Cat#354233
PBS, pH 7.4	Thermo Fisher Scientific	Cat#10010015
Essential 8™ Flex Medium Kit	Thermo Fisher Scientific	Cat#A2858501
Versene Solution	Thermo Fisher Scientific	Cat#15040066
Geltrex™ LDEV-Free	Thermo Fisher Scientific	Cat#A1413302
StemMACS™ Klf4 mRNA, human	Miltenyi Biotech	Cat#130-101-115
Lipofectamine™ RNAiMAX Transfection Reagent	Thermo Fisher Scientific	Cat#13778075
Opti-MEM™ I Reduced Serum Medium	Thermo Fisher Scientific	Cat#31985070
RPMI 1640 Medium	Thermo Fisher Scientific	Cat#21875034
GlutaMAX™ Supplement	Thermo Fisher Scientific	Cat#35050061
B-27™ Supplement, minus insulin	Thermo Fisher Scientific	Cat#A1895601
Recombinant Activin A	Peprotech	Cat#GMP120-14E
DMSO	Sigma Aldrich	Cat#D5879
10% Tween 20	Bio-Rad	Cat#1662404
SPRIselect-5mL	Beckman Coulter	Cat#B23317
Buffer EB	Qiagen	Cat#19086
TE Buffer	Thermo Fisher Scientific	Cat#12090015
Nuclease Free water	Thermo Fisher Scientific	Cat#AM9938
High Sensitivity D5000 ScreenTape	Agilent	Cat#5067-5592
High Sensitivity D5000 Reagents	Agilent	Cat#5067-5593
Pierce™ 16% Formaldehyde (w/v), Methanol-free	Thermo Fisher Scientific	Cat#28908
Donkey Serum	Sigma Aldrich	Cat#S30-M
1.0 M Tris pH 8.0	Amresco	Cat#E199
Gelatin	Sigma Aldrich	Cat#G2500
ProLong™ Gold Antifade Mountant with DAPI	Thermo Fisher Scientific	Cat#P-36931
Triton X-100	Sigma Aldrich	Cat#X-100
Phosphatase inhibitor Cocktail 2	Sigma Aldrich	Cat# P5726
Phosphatase inhibitor Cocktail 3	Sigma Aldrich	Cat# P0044
Protease inhibitor Cocktail	Sigma Aldrich	Cat# P8340
Gelatin from cold water fish skin	Sigma Aldrich	Cat#G7041
Critical commercial assays		
Chromium Next GEM Single Cell 3′ GEM, Library & Gel Bead Kit v3.1	10x Genomics	Cat#1000128
Chromium Next GEM Chip G Single Cell Kit	10x Genomics	Cat#1000127
Single Index Kit T Set A	10x Genomics	Cat#1000213
NextSeq 1000/2000 P2 Reagents (100 Cycles) v3	Illumina	Cat#20046811
NextSeq 1000/2000 P3 Reagents (100 Cycles) v3	Illumina	Cat#20040559
NextSeq 500 MID output	Illumina	Cat#20024904
NovaSeq 6000 SP Reagent Kit v1.5 (100 cycles)	Illumina	Cat#20028401
Deposited data		
Raw and analyzed data	This paper	GEO: GSE191286
Single-cell RNAseq of human naive and primed ESCs	(Messmer et al., 2019)	Array Express: E-MTAB-6819

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Single-cell RNAseq of human preimplantation embryos	(Petropoulos et al., 2016)	Array Express: E-MTAB-3929
Single-cell RNAseq of human postimplantation embryos	(Zhou et al., 2019)	GEO: GSE109555
Single-cell RNAseq of human pregastrulation embryos	(Xiang et al., 2020)	GEO: GSE136447
Single-cell RNAseq of human gastrulation embryos	(Tyser et al., 2021)	Array Express: E-MTAB-9388
Single-cell RNAseq of macaque embryos	(Yang et al., 2021)	GEO: GSE148683
Blastocyst-like structures generated from human pluripotent stem cells	(Yang et al., 2021; Yu et al., 2021)	GEO: GSE150578
Single-cell RNAseq of iBlastoids	(Liu et al., 2021)	GEO: GSE156596
Single-cell RNAseq of Kagawa et al. Human blastoids	(Kagawa et al., 2021)	GEO: GSE177689
Single-cell RNAseq of *in vitro* amnion-like cells	(Y. Zheng et al., 2019)	GEO: GSE134571
Experimental models: Cell lines		
Human: H9 (WA09)	WiCell	Primed hESC line H9
Human: ICSIG-1 IPSC0028 hiPSCs	Sigma	hiPSC line IPSC0028
Male mouse embryonic fibroblasts (MEFs) isolated from wild-type mouse B6 embryos.	In house generated	feeders
WIBR2 29M-GP26-TN9 hESCs	(Theunissen et al., 2014)	Primed hESCs
WIBR2-MGT 5iLA	In house generated	Naive hiPSCs line Converted from WIBR2 29M-GP26-TN9 hESCs using the 5iLA protocol
WIBR2-MGT PXGL	In house generated	Naive hiPSCs line Converted from WIBR2 29M-GP26-TN9 hESCs using the 5iLA protocol and maintained in PXGL.
Software and algorithms		
Codes used for the omics data analyses	This study	https://github.com/pasquelab/EXMCs
RStudio	(RStudio Team 2020)	N/A
STAR 2.7.3a	(Dobin et al., 2013)	N/A
Samtools v1.9	(Danecek et al., 2021)	N/A
Cellranger v4.0.0	(G. X. Y. Zheng et al., 2017)	N/A
Cellranger-atac v1.2.0	(Satpathy et al., 2019)	N/A
Seurat v4.0.1/v4.0.2	(Hao et al., 2021)	N/A
Ggplot	(Wickham et al., 2019)	N/A
Destiny	(Angerer et al., 2016)	N/A
ImageJ	(Schneider et al. 2012)	N/A
NIS-Elements	Nikon	N/A
Flowjo v10.7.2	Flowjo	N/A
Axio Vision v4.9.1.0	Axio microscope	N/A
Tidyverse	(Wickham et al., 2019)	N/A
pySCENIC	(Aibar et al., 2017; Van de Sande et al., 2020)	N/A
Signac v1.2.1	(Stuart et al., 2021)	N/A
MACS2 v2.2.7.1	(Zhang et al., 2008)	N/A
GenomeInfoDb v1.26.4	(Arora et al., 2022)	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
org.Hs.eg.db v3.12.0	(Carlson 2019)	N/A
BSgenome.Hsapiens.UCSC.hg38 v1.4.3	(Team TBD 2021)	N/A
Viridis v0.6.1	(Garnier et al., 2021)	N/A
Purrr v0.3.4	(Henry and Wickham 2022)	N/A
Tibble v3.1.0	(Müller and Wickham 2022)	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Vincent Pasque (vincent.pasque@kuleuven.be).

Materials availability

All stable reagents generated in this study are available from the lead contact without restriction except for human embryo derived cell lines and human induced pluripotent stem cell lines and their derivative for which permission must be requested from WiCell, Sigma or Dr. Rudolf Jaenisch and a material transfer agreement must be completed.

Data and code availability

- Raw and processed sequencing data (scRNA-seq, scATAC-seq) have been submitted to the NCBI GEO (http://www.ncbi.nlm.nih.gov/geo/) under accession number GEO: GSE191286.
- This paper analyzes existing, publicly available data. The accession numbers for the datasets are listed in the key resources table. All analysis code is available at: https://github.com/pasquelab/EXMCs.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Ethics statement

Work with human embryonic and induced pluripotent stem cells to model early human development was approved by the UZ/KU Leuven ethics committee (S52426, S64962, S66185, S66184 and S66375) and also by the Flemish Government (SBB 219 2020/0435).

Experiments on human embryos were performed in France, under the supervision of L. David. The use of human embryos donated to research as surplus of IVF treatment was allowed by the French embryo research oversight committee: Agence de la Biomédecine (RE18-010). All human preimplantation embryos used in this study were obtained from and cultured at the Assisted Reproductive Technology unit of the University Hospital of Nantes, France, which are authorized to collect embryos for research under approval of the Agence de la Biomédecine (AG110126AMP). Embryos used were initially created in the context of an assisted reproductive cycle with a clear reproductive aim and then voluntarily donated for research once the patients have fulfilled their reproductive needs or tested positive for the presence of monogenic diseases. Informed written consent was obtained from both parents of all couples that donated spare embryos following IVF treatment. Before giving consent, people donating embryos were provided with all of the necessary information about the research project and opportunity to receive counseling. No financial inducements are offered for donation. Molecular analysis of the embryos was performed in compliance with the embryo research oversight committee and The International Society for Stem Cell Research (ISSCR) guidelines (Kimmelman et al., 2016).

All blastoid experiments were performed at the IMBA, Austria. Blastoid generation was approved by the Commission for Science Ethics of the Austrian Academy of Sciences. This work did not exceed a developmental stage normally associated with 14 consecutive days in culture after fertilization. All experiments complied with all relevant guidelines and ethical regulations.

Human embryos extended culture

Embryos cultured for 5 days were thawed following the manufacturer's instructions (Cook Medical: Sydney IVF Thawing kit for slow freezing and Vitrolife: RapidWarmCleave or RapidWarmBlast for vitrification). The zona pellucida of each embryo was removed by brief laser impulse followed by manual extrusion of the embryo. Zona pellucida free embryos were washed in GTL medium and immediately transferred in 8-well IbiTreat m-plates (IB-80826; Ibidi GmbH) with 300 mL supplemented CMRL medium (Ma et al., 2019) and cultured at 37°C, in 21% O2/5% CO2. Half of the medium was replaced 48 h after thawing then every 24 h until day 10 of *in vitro* development.

Human blastoids

Human blastoids induction and culture was performed as described previously (Kagawa et al., 2021) with little modification. For the induction of blastoids, naive H9 hPSCs cultured with PXGL medium on MMC-MEF feeders were harvested with Accutase (Biozym).

Cells were resuspended in PXGL medium, supplemented with 10 μM Y-27632 (MedChemExpress) and seeded onto gelatin-coated plates and incubated at 37°C for 70 min to deplete feeders. Unattached cells were then collected, centrifuged to pellet, and resuspended in N2B27 medium containing 10 μM Y-27632, after which 30,000 cells were seeded onto a well of a 96 well plate containing 200 μm microwell array. Note that microwell arrays comprising microwells were imprinted into 96-well plates as previously described (Rivron et al., 2012; Vrij et al., 2016). After 24 h, the aggregation medium was replaced with N2B27 medium supplemented with PD0325901 (1 μM), A 83-01 (1 μM, MedChemExpress, HY-10432), 1-Oleoyl lysophosphatidic acid sodium salt (LPA) (500 nM, Tocris, 3854), human LIF (10 ng/mL), and Y-27632 (Tocris, 10 μM). The medium was refreshed every 24 h. 48 h after blastoid induction the medium was replaced with N2B27 supplemented with LPA (500 nM) and Y-27632 (10 μM). Blastoids were used for the downstream analysis 96 h after induction.

Cell lines

Human primed pluripotent stem cell culture

Human primed pluripotent stem cells (H9 hESCs (WiCell#WA09), Sigma hiPSCs (Sigma#iPSC EPITHELIAL-1-IPSC0028, ICSIG-1) and WIBR2 29M-GP26-TN9 hESCs (Theunissen et al., 2016) were grown with or without feeder conditions as per the cell lines, at normoxia conditions (5% CO2) and under humidified conditions at 37°C. In feeder-free conditions cells were cultured in pre-coated geltrex tissue culture treated plates in complete E8Flex medium (Stem cell technology). Cells were dissociated into smaller clumps every 5–6 days by incubating 5 min at room temperature in Versene. In feeder-dependent conditions primed hPSCs were grown on gelatin coated mitomycin-treated mouse embryonic fibroblast (MMC-MEF) feeders in human knockout serum replacement (KSR) primed medium containing 77.5% of DMEM/F12 (Gibco, 31330-038), 15% FBS (Gibco, 10270106), 5% KSR (Gibco, 10828028), non-essential amino acid (Gibco, 11140050), 2 mM L-glutamine (Gibco, 25030081), Penicillin-streptomycin (Gibco, 15140-122), β-mercapto-EtOH (Gibco, 31350-010) and adding 10 ng/ml FGF2 (Peprotech) freshly everyday. Cells were passaged every 6–7 days using a 20 min incubation in Collagenase, (ThermoFisher 17104019). Media was changed every day.

Human naive pluripotent stem cell culture

Naive hPSCs (H9 hESCs, WIBR2-MGT (converted from WIBR2 29M-GP26-TN9 hESCs) and Sigma hiPSCs) were cultured on MMC-MEF feeders in 5% O2 and 5% CO2 incubator under humidified conditions at 37°C. All naive hPSCs were cultured in PXGL medium (Bredenkamp et al., 2019b) which consists of 1:1 DMEM/F12 and Neurobasal, 0.5% N2-supplement, 1% B27-Supplement, 2 mM L-Glutamine, 0.1 mM β-mercaptoethanol, 1x penicillin-streptomycin, 1 μM PD0325901 (Axon Medchem), 2 μM XAV939 (Sigma-Aldrich, X3004), 2 μM Gö6983 (Tocris, 2285), 20 pg/mL human LIF (PeproTech, 300-05) and 10 μM Y-27632 (Tocris, 1254). Naive hPSCs were passaged every 4–5 days in a ratio 1:2 or 1:3 by single-cell dissociation with Accutase (Sigma-Aldrich, A6964-100 ML) followed by filtering through a 40 μm cell strainer (Corning, 352340).

Human primed to naive conversions

KLF4 mRNA conversion. Starting from day 1 or day 2 after seeding primed hPSCs in E8 onto Geltrex, cells were lipofected daily with KLF4 mRNA (Miltenyi, 130-101-115) for 9 days (Liu et al., 2017). Per well of a 6-well plate, 2 μL KLF4 mRNA were diluted in 250 μL Opti-MEM (Gibco, 31985- 047) and 6 μL of lipofectamine RNAimax (Invitrogen, 13778075) in another 250 μL of Opti-MEM. Then the diluted mRNA was added to the diluted lipofectamine and incubated for 20 min at room temperature. After medium was exchanged to 1.5 mL fresh E8 flex, the mixture was added dropwise. After 10 days of lipofection, cells were passaged with Versene onto MMC-MEF feeders in E8 flex and transfected again. Starting from the following day, the medium was switched to t2iLGö supplemented with Y-27632 (Tocris, 1254) and cells were transferred to hypoxia, while the lipofection was repeated every day for another 5 days. The naive t2iLGö medium contains a 50:50 mixture of DMEM/F12 (Gibco, 31330-038) and Neurobasal medium (Gibco, 21103-049), supplemented with 2 mM L-glutamine (Gibco), 0.1 mM β-mercaptoethanol (Gibco, 31350-010), 0.5% N2 supplement (Gibco, 17502-048), 1% B27 supplement (Gibco, 17504-044), 1% Penicillin-streptomycin (Gibco, 15140-122), 10 pg/mL human LIF (PeproTech, 300-05), 250 μM L-ascorbic acid (Sigma-Aldrich, A4544-100G), 10 μg/mL recombinant human insulin (Sigma, I9278-5 ML), 1 μM PD0325901 (Axon Medchem, 1408), 1 μM CHIR99021 (Axon Medchem, 1386), 2.5 μM Gö6983 (Tocris, 2285) (Liu et al., 2017). After a total number of 15 days of transfection, cells were passaged with Accutase (Sigma-Aldrich, A6964-100 ML) on the following day onto fresh feeders and continued to be cultivated in PXGL.

5iLA conversion. To convert primed hPSCs to naive hPSCs, trypsinized primed hPSCs were seeded onto a gelatin coated MMC-MEF feeders tissue culture treated 6 well plates and cultured with human KSR primed medium along with 10 μM of Y-27632 (Tocris, 1254) in a humidified normoxia (5% CO2) for 2 days. On the 3rd day and after giving a wash with Phosphate-buffered Saline (PBS (Gibco, 10010-015)) the medium was changed to 5iLA medium composed of 1:1 DMEM/F12 (Gibco, 31330-038) and Neurobasal (Gibco, 21103-049), 1% N2-supplement (Gibco, 17502-048), 2% B27 supplement (Gibco, 17504-044), 20 g/mL recombinant human LIF (PeproTech, 300-05), 2 mM L-glutamine (Gibco, 25030-081), 1% non-essential amino acid, 0.1 mM β-mercaptoethanol (Gibco, 31350-010), 1x Penicillin-streptomycin (Gibco, 15140-122), 50 μg/mL BSA (Sigma-Aldrich, A3059) and supplemented with 5 inhibitors: PD0325901 (Stemgent, 1 μM), IM-12 (Enzo, 1 μM), SB590885 (R&D systems, 0.5 μM), WH-4-023 (A Chemtek, 1 μM), Y-27632 (Tocris, 10 μM), and Activin A (Peprotech, 20 ng/mL) and grown in a humidified incubator in hypoxia condition (5% CO2 and 5% O2) at 37°C. After an initial wave of cell death around day 10–13, dome-shaped naive colonies started appearing. These cells were passaged into single cells every 4–5 days using 5 min incubation in Accutase (Sigma-Aldrich, A6964-100 ML) at 37°C.

For some experiments, H9 and WIBR2-MGT naive hPSCs which were derived and cultured in 5iLA conditions were switched at passage 12 into PXGL naive medium for stable maintenance and expansion.

Mouse feeders

MEFs were isolated from E14.5 pregnant WT C57/Black 6 mice. Male embryos were selected based on sex genotyping PCR and immortalized with Mitomycin C (Bioconnect). MEFs were cultured and harvested in a humidified incubator at 37°C and in 5% CO2 by using filter sterilized MEF medium consisting of 90% of DMEM supplemented with 10% FBS, 1% Glutamax, 1x Penicillin-streptomycin, 1x non-essential amino acid and 0.1 mM β-mercaptoethanol and on 0.1% gelatin-coated tissue culture treated plates.

Naive human pluripotent stem cells to trophoblast and EXMC fate conversion

Human naive to trophoblast and EXMC conversions were done using the following previously described protocols for hTSCs (Cinkornpumin et al., 2020; Dong et al., 2020; Guo et al., 2021; Io et al., 2021). Naive hPSCs were seeded in such a way that expecting at least 90% confluency by Day 2 on MMC-MEFs after dissociating in single-cell using TryplE (for 15 min at 37°C) in their respective naive culture conditions supplemented with 10 μM Y-27632 (Tocris, 1254). The very next day, after a wash with PBS (Gibco, 10010-015), the media was switched from naive to ASECRiAV medium (Okae et al., 2018) consisting of DMEM/F12 (Gibco, 31330038) supplemented with 0.3% BSA (Sigma, A3059), 0.2% FBS (Gibco, 10270-106), 1% Penicillin-streptomycin (Gibco, 15140-122), 1% insulin-transferrin-selenium-ethanolamine-X 100 supplement (Gibco, 51500056), 1.5 μg/mL L-ascorbic acid (Sigma, A8960), 0.5 μM A83-01 (Peprotech, 9094360), 1 μM SB431542 (Axon Medchem, 1661), 50 ng/ml hEGF (Miltenyi Biotec, 130-097-750), 2uM CHIR99021 (Axon Medchem, 1386), 0.8 mM Valproic acid (Sigma, V0033000), 0.1 mM β-mercapto-EtOH (Gibco, 31350-010) and 5 μM Y-27632 (Tocris, 1254). The medium was changed every two days and supplemented with 5 μM Y-27632. From passage 1 onwards both hTSCs and EXMCs were cultured and maintained on 5 μg/mL Collagen IV coated tissue culture treated plates in hypoxia conditions (5% O2 and 5% CO2) and passaged every 5 days at 1:3 or 1:6 splitting ratios. Collagen IV coated cell culture plates were coated overnight at 37°C. EXMCs and hTSCs used in all experiments were always cultured and maintained in ASECRiAV medium unless otherwise specified.

Naive human pluripotent stem cells to PrE and nXEN fate conversion

Human naive to PrE and nXEN conversions were done using a previously described protocol with minor adaptations (Linneberg-Agerholm et al., 2019). Naïve human PSCs grown in PXGL on feeders were seeded at a high density (split ratio of 1:1 or 2:1) directly in RACL medium, which is made of Roswell Park Memorial Institute 1640 medium (RPMI; Gibco, 218750-34) supplemented with 1 X GlutaMAX, 1 X B27 minus insulin (Gibco, A18956- 01), 1% (v/v) Pen/Strep, 100 ng/mL Activin A (PeproTech, 120-14E), 3 μM CHIR99021 and 10 pg/mL recombinant human LIF onto fresh MMC-MEF feeders in hypoxia. Cells were kept in RACL medium for at least 7–8 days and the medium was refreshed every day.

After PrE cells reached confluency, they were dissociated with TrypLE or Accutase and re-plated onto fresh MMC-MEF feeders at a ratio of 1:2 to 1:4 in NACL medium, which is composed of 1:1 DMEM/F12 and Neurobasal, 0.5% N2-supplement, 1% B27-Supplement, 2 mM L-Glutamine, 0.1 mM β-mercaptoethanol, 1x penicillin-streptomycin, supplemented with 100 ng/mL Activin A, 3 μM CHIR99021, 10 pg/mL recombinant human LIF and 10 μM Y-27632 (Linneberg-Agerholm et al., 2019). These nXEN cells were subcultured every 4–7 days with TrypLE or Accutase.

METHOD DETAILS

Immunofluorescence and microscopy

Immunofluorescence staining was performed as described previously (Pasque et al., 2014) with little modification. Cells were grown on 0.1% gelatinized 18 mm round coverslips with or without feeders. The next day cells were fixed in 4% paraformaldehyde-PBS for 10 min at room temperature in the dark and permeabilized with 0.5% Triton X-100 in PBS for 10 min and washed twice with 0.2% Tween 20 in PBS (PBST) for 5 min each before proceeding to the staining. After this step, cells were either stored at 4°C or directly subjected to staining. Primary and secondary antibodies were diluted in a blocking buffer containing mainly PBST with 5% normal donkey serum and 0.2% fish skin gelatin. Cells on coverslips were incubated at 4°C with the specific primary antibodies in blocking solutions (1:100 dilution for most antibodies, 1:50 dilution for NANOG, 1:40 dilution for FOXA2), after that washed three times with PBST each 5 min. After that it was incubated with the appropriate corresponding fluorophore conjugated secondary antibodies in blocking buffer (1:500 dilution) for 1 h in the dark, washed again 3 times with PBST 5 min each, washed with 0.002% DAPI (Sigma-Aldrich, D9542) solution in PBST. The coverslips were mounted in Prolong Gold reagent with DAPI after a final wash in PBST. Mounted coverslips were kept at room temperature in the dark overnight before imaging. All immunofluorescence images were taken in a Zeiss Axioimager A1 inverted microscope with an AxioCam MRc5 camera and processed in ImageJ. Bright field images were taken using a Nikon Eclipse Ti2 microscope and analyzed using ImageJ software.

Immunohistochemistry and microscopy on blastoids

Immunohistochemistry on human blastoids was performed as described previously (Kagawa et al., 2021) with little modification. The samples were fixed with 4% formaldehyde for 30 min at room temperature. Post fixation, formaldehyde solution was removed and the samples were washed at least three times with PBS. The samples were then permeabilized and blocked using 0.3% Triton X-100 and 10% normal donkey serum in PBS for at least 60 min. The samples were then incubated overnight at 4 °C with primary antibodies diluted in fresh blocking/permeabilization solution (1:200 dilution for VIM, 1:300 dilution for GATA2, 1:400 dilution for SOX2). The samples were washed with PBS containing 0.1% Triton X-100 (PBST) at least three times for 10 min each. The washing buffer was then replaced with Alexafluor tagged secondary antibodies (1:300 dilution, Abcam or Thermofisher scientific) along with a nuclear dye

Hoechst-33342 (1:300 Life Technologies, H3570) diluted in PBST for 30 min in dark at room temperature. The samples were then washed with PBST three times for 10 min each. The blastoid fluorescent images were acquired using Olympus IX83 microscope with a Yokogawa W1 spinning disk (Software: CellSense 2.3; camera: Hamamatsu Orca Flash 4.0). Confocal images were analyzed and display images were exported using FIJI.

Immunohistochemistry and microscopy on embryos

For human embryos, the samples were fixed with 4% paraformaldehyde for 15 min at room temperature and washed in PBS/BSA. Embryos were permeabilized and blocked in PBS containing 0.2% Triton X-100 and 10% FBS at room temperature for 60 min. Samples were incubated with primary antibodies (1:200 dilution for NR2F2, 1:400 dilution for GATA4, 1:100 for VIM) overnight at 4°C. Incubation with secondary antibodies (1:300 dilution) was performed for 2 h at room temperature along with 0.33% DAPI counterstaining. Confocal immunofluorescence images of human embryos were acquired with a Nikon confocal microscope and a 20× Mim or 25x Silicon objective. Optical sections of 1 μm-thick were collected. The images were processed using Fiji (http://fiji.sc) and Volocity 6.3 visualization softwares. Volocity software was used to detect and count nuclei. Day 10 human embryos were photobleached in order to restain it (Figures 2D and S2G, Videos S1 and S2). Under Nikon A1 SIM confocal, lasers at 647 nm were set to 100% power for around 10 min until fluorescence fainted. The NANOG (Goat/647 nm) stained embryos were restained with VIMENTIN (Mouse/647 nm).

Flow cytometry

hPSCs were dissociated using Accutase (Sigma-Aldrich, A6964-100 ML) into single cells by incubating 5 min at 37°C. Before proceeding the antibody staining, cells were washed 2 times with FACS buffer containing 1% BSA in PBS (Gibco, 10010-015). Fluorophore conjugated antibodies were diluted at a ratio of 1:50 which is 1 μL of antibody in 50 μL of FACS buffer for around 50000 to 100000 cells, and incubated at 4°C in the dark at least for 30 min. Cells were washed again with FACS buffer and passed through a 40 μM cell strainer (Corning, 352340) and analyzed using a BD influx. Single stained controls were used for compensation and setting up the precise and stringent gate in the flow cytometer.

Western Blot

Cells were collected and subsequently lysed using RIPA cell lysis buffer (Sigma) containing 1:100 protease inhibitor, phosphatase inhibitor cocktail 2 and phosphatase inhibitor cocktail 3 (Sigma). Lysates were rotated for 2 h at 4°C and supernatants were collected after centrifugation at max speed for 10 min at 4°C. Protein concentrations were determined using the Bradford Protein Assay (Bio-Rad). Equal quantities of protein per sample were combined with the Laemmli loading buffer, followed by denaturation for 5 min at 95°C and were subsequently loaded for SDS/PAGE separation on a 10% polyacrylamide gel, transferred to Nitrocellulose using the iBlot 2 Transfer system (ThermoFisher), followed by immunoblotting. Resulting blots were subsequently blocked with 5% BSA at room temperature for 1 h, followed by overnight incubation with the primary antibody (diluted 1:1000) at 4°C. After three times washing with a Tris-based saline buffer supplemented with Triton (TBS-T) for 5 min each, the blots were incubated with corresponding secondary antibodies (1:5000 dilution) for 1 h at room temperature. Blots were developed by bioluminescence using Pierce™ ECL Western Blotting Substrate (ThermoFisher, 32209). β-ACTIN was used as a loading control and protein quantifications were performed with ImageJ software. Cells were treated with or without the inhibitor for 10 days before collected for western blot.

Single-cell RNA sequencing

Cell preparation

Day 30 cells were washed with PBS (Gibco, 10010-015), dissociated from culture dishes using Accutase (Sigma-Aldrich, A6964-100 ML) (7 min at 37°C) in hypoxic condition for naive hPSCs and hTSCs and in normoxic condition for primed hPSCs, and finally diluted with DMEM/F12 (Gibco, 31330-038).Single-cell suspensions were filtered through a 40 μm cell strainer (Corning, 352340), centrifuged at 200 rcf for 5 min, resuspended in PBS with 0.04% Bovine Serum Albumin (BSA) (Sigma Aldrich, A9418-50G) (1000 cells/μl).

Cells were collected at day 0, 1, 2, 4, 8, 13 and 18 during the time course ASECRiAV conversion. During the conversion cells were passaged at day 5, 10 and 15 (related to Figure 5). Day 70 EXMCs which were generated with an independent conversion and isolated by FACS were also included. Cells were washed with PBS, treated with Accutase (10 min. At 37°C, hypoxia), and diluted with DMEM/F12. Single cell suspensions were filtered through a 40 μm cell strainer, and centrifuged (200 rcf, 500 min). Cells before the first passage (day 0, 1, 2 and 4) were depleted from feeders, while cells after the first passage (day 8, 13 and 18) and at day 70 were immediately resuspended in resuspension buffer (PBS, 1% BSA). For feeder depletion, cells were resuspended and plated on a gelatine-coated plate for 35 min in PXGL medium for cells at day 0 or in ASECRiAV medium for cells at day 1, 2 and 4, collected and centrifuged (200 rcf, 500 min), and resuspended in resuspension buffer. Single cells in resuspension buffer were centrifuged (200 rcf, 5 min) and resuspended again in resuspension buffer (1000 cells/μl). Finally, cells were filtered with the Flowmi 40 μm tip strainer (Bel-Art, H13680-0040).

Cells collected at day 6 of the RACL conversion and at day 24 of the NACL conversion, were washed with PBS, treated with Accutase (15 min at 37°, hypoxia), and diluted with DMEM/F12. Single cell suspensions were filtered through a 40 μm cell strainer, centrifuged at 200 rcf for 5 min, resuspended in resuspension buffer (PBS, 1%BSA), centrifuged (200rcf, 500 min), and finally resuspended in resuspension buffer (1000 cells/μl) and filtered with a Flowmi 40 μm tip strainer. All cells were counted with the Luna-FL automated Fluorescence Cell Counter (Logos Biosystems).

Library preparation and sequencing

Cells were loaded onto the 10X Chromium Single Cell Platform (10X Genomics) targeting 2000, 4000 or 5000 cells (Next GEM Single Cell 3′ library and Gel Bead Kit v3.1) according to the manufacturer's protocol (10x User Guide; CG000204, Revision D). Generation of gel beads in emulsion (GEMs), barcoding, GEM-RT cleanup, complementary DNA amplification and library construction were all performed according to the manufacturer's protocol. Individual sample quality was assessed using a Tapestation (Agilent). Qubit 2.0 (ThermoFisher Scientific) and KAPA Library Quantification Kit for Illumina Platform (KAPA Biosystems) were used for library quantification before pooling. The final library pool was sequenced on a NovaSeq6000 (Illumina) or NextSeq2000 (Illumina) instrument using NovaSeq SP reagent kit v1.5 (Illumina, 20028401) or NextSeq 1000/2000 P3 kit v3 for 2 lanes of 100-base-pair paired-end reads, or NextSeq 1000/2000 P2 kit v3 for 1 lane of 100-base-pair paired end reads.

Single-cell RNA-seq analysis

Raw sequence reads were quality-checked using the FastQC software. The CellRanger version 4.0.0 was used to process, align and summarize unique molecular identifier (UMI) counts against the 10X Genomics pre-built human GRCh38 and mouse mm10 reference genome datasets (2020-A, July 7, 2020). Downstream analyses were performed in R using Seurat (v4.0.1) (Satija et al., 2015). Human cells were retained and mouse cells were filtered out by adjusting the number of counts per cell (nCount_RNA) and the number of mapped genes per cell (nFeature_RNA) to only keep cells that were mostly mapped to the human GRCh38 (hg38) genome (nFeature_RNA > 100 and nFeature_RNA < 15000). Cells with more than 25% of mitochondrial counts were filtered out. The count matrix was normalized with Seurat global-scaling normalization method "LogNormalize" that normalizes the feature expression measurements for each cell by the total expression, multiplies this by a 10.000 scale factor, and log-transforms the result. Differential expression testing was performed with the FindMarkers function in Seurat based on the non-parametric Wilcoxon rank sum test applying the logFC threshold of averaged $\log_2$ FC > 0.25. A graph-based cell clustering approach was used to cluster cells with FindClusters function in Seurat.

Single-cell gene expression analysis of merged datasets

Single-cell RNA-seq datasets (Messmer et al., 2019; Petropoulos et al., 2016; Tyser et al., 2021; Zhou et al., 2019) (Related to Figure 1F) were integrated with the data generated within this study performed using Seurat v3 integration standard workflow (Butler et al., 2018; Stuart et al., 2019). Datasets were normalized and scaled before selecting the 2000 most variable genes. The FeaturePlot function was used to project individual gene expression on UMAP. Differential expression analysis was performed with the FindMarkers function based on the non-parametric Wilcoxon rank sum test applying the logFC threshold of averaged $\log_2$ FC > 0.25. Similar integrations were performed with blastoid data (Kagawa et al., 2021) and with monkey data (Tan et al., 2021; Yang et al., 2021).

Correlation coefficients were calculated on the basis of the top 2000 most highly variable genes across all data sets using corrplot v.0.92.

The data generated within this study was integrated with published scRNA-seq dataset as follow: Figure 1F (Messmer et al., 2019; Petropoulos et al., 2016; Tyser et al., 2021; Zhou et al., 2019), Figures 1H and 1I (Tan et al., 2021; Yang et al., 2021), Figure 7B (Kagawa et al., 2021; Petropoulos et al., 2016; Tyser et al., 2021), Figures S5A and S6G (Messmer et al., 2019; Petropoulos et al., 2016; Tyser et al., 2021; Xiang et al., 2020; Zhou et al., 2019).

Cell type annotations are adopted as follows: for Figure 1F (Castel et al., 2020; Tyser et al., 2021), Figure 1H (Tan et al., 2021; Yang et al., 2021), Figures 2A and 2B (Castel et al., 2020; Chhabra and Warmflash, 2021; Tyser et al., 2021; Zheng et al., 2019).

Gene regulatory network interference

GRNs were inferred using pySCENIC (Aibar et al., 2017; Van de Sande et al., 2020). First raw expression data were normalized by dividing feature counts of each cell by the total counts for that cell and multiplying by a factor of 10.000, followed by $\log_{1p}$ transformation. The normalized counts were used to generate the co-expression modules using GRNboost. Next, the RcisTarget package was used to assess target binding motif enrichment and create regulons with only genes containing a binding motif, where a regulon is a transcription factor and its target genes (Aibar et al., 2017). Subsequently, AUCell was used to measure regulon activity. Here, AUCell used the area under the curve to calculate the enrichment of the regulon across the ranking of all genes in a particular cell, resulting in a matrix of the activity of each regulon in each cell. Downstream analyses were done using the Seurat package (Satija et al., 2015).

Single-cell ATAC-seq cell preparation and sequencing

Cell lines samples collection and nuclei isolation

Cells were washed with PBS (Gibco, 10010-015), dissociated from culture dishes by Accutase (Sigma-Aldrich, A6964-100 ML) (7 min at 37°C) in hypoxic condition for naive hPSCs and hTSCs and in normoxic condition for primed hPSCs, and finally diluted with DMEM/F12. Single-cell suspensions were filtered through a 40 μm cell strainer (Corning, 352340), centrifuged at 200 rcf for 5 min, resuspended in PBS (Gibco, 10010-015) with 0.04% Bovine Serum Albumin (BSA), and counted with a NucleoCounter NC-100 (Chemometec). 100.000 to 1.000.000 cells were added to a 2-mL microcentrifuge tube and were centrifuged (300 rcf, 5 min at 4°C). The supernatant was removed without disrupting the cell pellet and 100μL chilled lysis buffer (10 mM Tris-HCl (pH 7.4), 10 mM NaCl, 3 mM MgCl2, 0.1% Tween-20, 0.1% Nonidet P40 Substitute, 0.01% digitonin and 1% BSA) was added and mixed by pipetting 10 times.

The microcentrifuge tube was incubated on ice for 4 min and then 1 mL chilled wash buffer (10 mM Tris-HCl (pH 7.4), 10 mM NaCl, 3 mM MgCl2, 0.1% Tween 20 and 1% BSA) was added and mixed by pipetting 5 times. Nuclei were centrifuged (500 rcf, 5 min at 4°C) and the supernatant was removed without disrupting the nuclei pellet. Nuclei were resuspended in a chilled Diluted Nuclei Buffer (10x Genomics, 2000153) at 610–1540 nuclei per μL based on the starting number of cells. The nuclei concentration was determined using a NucleoCounter NC-100 (Chemometec). Nuclei were pooled before loading onto the 10X Chromium using the following ratio (H9 primed hPSCs: Sigma primed hPSCs: H9 naive hPSCs: Sigma naive hPSCs: Sigma day 30 conversion = 1:1:1:1:2) and were immediately used to generate scATAC-seq libraries as described in the STAR Methods below.

Library preparation and sequencing

scATAC-seq libraries were prepared according to the Chromium Single Cell ATAC Reagent Kits User Guide (10x user guide; CG000168, Revision D). Nuclei were loaded onto the 10x Chromium Single Cell Platform (10x Genomics) at a concentration targeting 2000 nuclei, according to the manufacturer's protocol. Nuclei transposition, generation of gel beads in emulsion (GEMs), barcoding, GEM cleanup, and library construction were performed with the Chromium Single Cell ATAC Reagent Kits (v1 Chemistry) according to the manufacturer's protocol. Library Quality control and quantification was assessed using the Tapestation (Agilent) and Qubit 2.0 (ThermoFisher Scientific). The library was sequenced on a Nextseq500 (Illumina) instrument using the MID output kit (Illumina) (20024904).

Single-cell ATAC-seq analysis

Preprocessing of single-cell chromatin accessibility data was performed using the cellranger-atac version 1.2.0 pipeline (10X Genomics). Read filtering, alignment, cell and peak calling, as well as cell-by-peak count matrix generation were performed using the ''count'' option (cellranger-atac count) with default parameters against the 10X Genomics pre-built human GRCh38 and mouse mm10 reference genome datasets (GRCh38_and_mm10 Reference-1.2.0, November 21, 2019). Downstream analysis was performed using Signac (Stuart et al., 2021). The count matrix was filtered for cells where at least 15% of all fragments fell within peaks, with less than 5% of fragments falling within blacklist regions, with less than 4% nucleosome signal, with at least 2% enrichment for transcription start sites, and peaks with a minimum of 2500 fragments and a maximum of 20000 fragments. Dimensionality reduction was performed with the RunUMAP function from the Seurat package. Cell clusters were identified using the FindNeighbors function with parameters ''reduction = 'lsi', dims = 2:30'' and FindClusters function with parameters ''algorithm = 3, resolution = 0.5''. Upon initial clustering, cluster-specific peaks were called with MACS2 using the function CallPeaks with parameters ''group.by = 'idents'''. These peaks were used to generate a final Signac object which was used for all downstream analyses. Fragment counts were mapped to peaks using FeatureMatrix with parameters ''process_n = 2000''. The new count matrix was filtered for cells where at least 20% of fragments fell within peaks, with less than 0.8% of fragments falling within blacklist regions, with less than 2% nucleosome signal, with at least 2% enrichment for transcription start sites, and for peaks with a minimum of 1500 and a maximum of 10000 fragments. Clustering was performed on the new object as described above. Peaks were then filtered for mean accessibility, keeping all peaks with greater than 0.1 mean accessibility across cells. The Upset plot was generated using the UpSetR package (Lex et al., 2014). Motifs were mapped using the motifmatchr package and the JASPAR2018 database. Motifs were added using the AddMotifs function with default parameters. Motifs were mapped to all peaks uniquely accessible per cluster. FindMotifs was used to test enriched motifs, and peaks were controlled for length and GC content.

Accessibility scores were calculated per cell type by averaging the log normalized number of fragments between 2 kb upstream of the transcription start site and the transcription termination site for each gene across each cluster (Related to Figure 3E). Motifs were enriched in the cluster-specific peaks of each cluster (Related to Figure 3F).

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical tests and data processing were performed in R (v4.0.3). Information on each statistical test and multi-testing correction used can be found in the result section and figure legends. For statistical tests on single cell RNA-seq experiments, one replicate was included per time point and the number of individual cells of each cell type is indicated in Table S5. Cells were filtered to keep only cells with between 100 and 15000 human genes expressed, and cells with <25% mitochondrial reads were kept. Mouse cells were also excluded as described in the scRNA-seq analysis section. Differential gene expression analysis was performed using the Seurat function FindMarkers based on the non-parametric Wilcoxon rank sum test, using cutoffs of adjusted p-value<0.05 and log2 fold change>0.25. Regulon comparisons used a Wilcoxon rank sum test, with Bonferroni adjusted p-values.

For this work more than 35 naive to ASECRiAV conversions were performed. All experiments have been repeated at least three or two times, with exceptions. Experiments were repeated three times for Figures 2G, 2H, 3D, and S2B and two times for all other figures except for Figures 6E and 6G and S6B which were performed once. Experiments were not blinded. No data were excluded with the exception of cell filtering described in the scRNA-seq analysis section.