HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2025 May 16.

Published in final edited form as:

Cell Rep. 2025 April 22; 44(4): 115568. doi:10.1016/j.celrep.2025.115568.

HAND1, partially mediated through ape-specific LTR binding, is essential for human extra-embryonic mesenchyme derivation from iPSCs

Zukai Liu[1,2,5], **Yuliana Tan**[1,2,6], **William F. Flynn**[1], **Lili Sun**[1], **Ponthip Pratumkaew**[1,3], **Juliana Alcoforado Diniz**[1], **Nelio A.J. Oliveira**[1], **Justin A. McDonough**[1], **William C. Skarnes**[1,2], **Paul Robson**[1,2,4,7,*]

[1]The Jackson Laboratory for Genomic Medicine, Farmington, CT 06032, USA

[2]Department of Genetics and Genome Sciences, University of Connecticut Health Center, Farmington, CT 06030, USA

[3]Siriraj Center of Excellence for Stem Cell Research, Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand

[4]Institute for Systems Genomics, University of Connecticut, Farmington, CT 06030, USA

[5]Present address: Department of Genome Sciences, University of Washington, 3720 15th Ave. NE, Seattle, WA 98195, USA

[6]Present address: Boehringer Ingelheim Pharmaceuticals, Inc., 900 Ridgebury Road, Ridgefield, CT 06877, USA

[7]Lead contact

SUMMARY

The specification of extra-embryonic mesenchyme (ExMC) is a prime example of developmental divergence between mouse and human. Derived from definitive mesoderm during mouse gastrulation, the human ExMC first appears at peri-implantation prior to gastrulation and therefore its human cellular origin, still unknown, must differ. In a human pluripotent stem cell model, we report that ExMC shares progenitor cells with trophoblast, suggesting a trophectoderm origin. This ability to form ExMC appears to extend to human trophoblast stem cell lines. We define HAND1 as an essential regulator of ExMC specification, with null cells remaining in the trophoblast lineage. Bound by HAND1, ape-specific, endogenous retrovirus-derived LTR2B contributes to unique features of ExMC. Additionally, ExMC supports the maintenance of

*Correspondence: paul.robson@jax.org.

AUTHOR CONTRIBUTIONS

Conceptualization, Z.L. and P.R.; methodology, Z.L., Y.T., L.S., P.P., W.F.F., J.A.D., N.A.J.O., and J.A.M.; investigation, Z.L., L.S., P.P., and J.A.D.; visualization, Z.L., Y.T., and W.F.F.; funding acquisition, W.C.S. and P.R.; project administration, P.R.; supervision, P.R.; writing – original draft, Z.L. and P.R.; writing – review & editing, Z.L., Y.T., W.F.F., P.P., W.C.S., and P.R.

DECLARATION OF INTERESTS

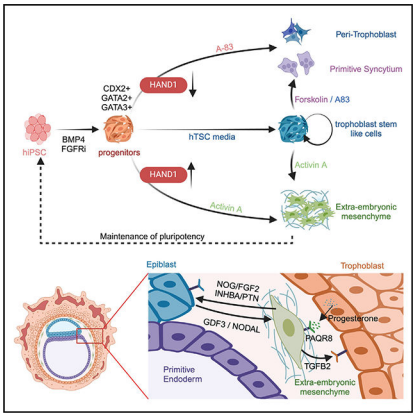
The authors declare no competing interests.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2025.115568.

pluripotent stem cells, possibly reflecting a role in maintaining epiblast pluripotency through peri-implantation development. Our data emphasize the nascent evolutionary innovation in human early development and provide a cellular system to study this.

Graphical Abstract



In brief

Liu et al. show that activin A promotes BMP4-treated human iPSC differentiation into ExMC and uncover the essential role of HAND1 in this process. In turn, ExMC secretes extracellular matrix and growth factors sufficient to support the maintenance of pluripotency, suggestive of one role for ExMC in human early embryogenesis.

INTRODUCTION

The extra-embryonic membranes of the mammalian embryo include several cell types within the trophoblast lineage, the extra-embryonic endoderm-derived yolk sac, and the extra-embryonic mesenchyme (ExMC) cells. These membranes are known to diverge extensively between species[1] as emphasized in the differing cellular origins of ExMC between mouse and human. Mouse ExMC originates from the presumptive mesoderm within the primitive streak during gastrulation,[2,3] whereas human ExMC appears much earlier in development. While its origin is unclear, human ExMC has been morphologically defined as early as the blastocyst[4] and is abundant in the peri-implantation to early post-implantation conceptus prior to gastrulation.[5]

The ability to molecularly define individual cells through single-cell transcriptomics (single-cell RNA sequencing [scRNA-seq]) has been a boon to understanding cellular processes in early mammalian development.[6,7] This has been applied to human pre-implantation embryos,[8,9] blastocysts pushed through extended *in vitro* cultures,[10] and a gastrulating embryo at approximately 9–12 days post implantation (Carnegie stage 7),[11] the latter of which captured some ExMC. Due to its inaccessibility, there is no molecular data of any kind from the human conceptus within the first week post implantation, a stage at which there is extensive ExMC.[12] Old World monkey (OWM) data generated from this stage fills some of this gap and captures ExMC emerging concurrently with the primary yolk sac.[13,14]

Caution, however, is warranted in extrapolating OWM data to the human, as there are known differences, for example, with the extent of conceptus invasion into the endometrium.[15] With their ability to give rise to the extra-embryonic lineages,[16–21] human induced pluripotent stem cells (iPSCs) provide a powerful model to investigate the molecular and cellular formation of the ExMC.

Here we probe, using scRNA-seq, the developmental dynamics within a previously described human iPSCs-to-trophoblast differentiation scheme and uncover trophoblast and mesenchymal populations within 6 days of differentiation, with both populations apparently emerging from a single progenitor. We subsequently optimize two differentiation protocols to enrich for either trophoblast or ExMC and utilize the latter to define HAND1 as an essential regulator of ExMC formation—in its absence, cells progress to primitive syncytium (PrSyn). In addition, we define HAND1-bound genomic loci within the developing ExMC that are clearly ape specific, derived from recent endogenous retroviral long terminal repeat (LTR) insertion events. These drive HAND1-mediated, ape-specific ExMC transcription in addition to creating ape-specific additional isoforms of pre-existing protein-coding genes. This study provides insight into the molecular and developmental innovations of the human ExMC and provides cell-culture systems, both trophoblast and ExMC, to enable continued investigations into this little-understood period of human development.

RESULTS

Identification of ExMC from BMP4- and FGFR-inhibitor-treated iPSCs

We generated a time-course of iPSC-derived trophoblast to make a comparison to recently published data from appropriate stages of human embryos. In addition to BMP4 (B), a known inducer of trophoblast lineage transcription factors (TFs),[18] our differentiation cocktail also included the FGFR inhibitor SU5402 (S), which is known to prevent the expression of BRACHYURY[22] and thereby block embryonic mesoderm lineage formation. We then sampled nine time points across days 1 through 6 of differentiation in this BS medium by scRNA-seq (Figure 1A). The dataset comprised 15,839 cells with >2,000 genes/cell with a median of 15,396 transcripts/cell. We excluded day 0 from our visualization of these data, as there was a distinct jump from day 0 to day 1, resulting from the rapid BMP4 induction of trophoblast lineage genes rather than the subsequent continuum seen days 1 through 6. Relatively uniform across days 1 and 2, a bifurcation point emerged through day 3 and day 4, leading to two distinct cell populations on days 5 and 6 (Figure 1B).

Cells from day 1 co-expressed pluripotency (*OCT4* and *SOX2*) and extra-embryonic cell (*GATA3*, *GATA2*, *TFAP2A*, and *TFAP2C*) lineage markers and showed epithelial morphology, thus resembling early blastocystic trophoblast, which we define as early progenitors (EPs). On day 2, cells had decreased expression of pluripotency markers and further increased expression of extra-embryonic cell-lineage markers.[18] Therefore, we define these cells as progenitor cells (Progenitor). From day 3 to day 4, the cells transitioned to a state similar to later-stage peri-implantation trophoblasts (Peri-TB, marked by *VGLL1* and *PHLDA2*) with a distinct population of cells expressing PrSyn markers (*CGA*, *PRR9*, *SP6*, *SDC1*, *ERVW-1*, *GCM1*, and *CYP19A1*) (Figures 1C and 1D), similar to the polar trophectoderm from blastocysts at day 7.[8] A second distinct population of cells at day 6

contained a mesenchymal cell signature (e.g., *COL3A1* and *COL6A3*) with genes (*SNAI2*, *HAND1*, and *LUM*) shared with ExMC defined in the OWM dataset. Two subpopulations in this ExMC cluster can be distinguished by cell-cycle markers, indicating a proliferating subpopulation (ExMC 2) (Figures 1C and 1D).

While *GATA3*/*TFAP2A*/*TFAP2C* are shared by trophectoderm and amniotic cells, *GATA2*/*SEMA6D*/*FABP3*/CDX2 are more restricted and transiently expressed in the trophectoderm lineage of the early-stage human blastocyst,[8,23,24] and in our model this same developmental dynamic is seen (Figure 1E). *NR2F2* marks the polar trophectoderm of the later-stage human blastocyst and is followed by increased expression of *CCR7* and *CGA* in PrSyn.[8,9] *NR2F2* follows *CDX2* expression in our model, with *CCR7* expression initiating at day 5 and remaining restricted to the trophoblast (Figure 1E), revealing trophectoderm-specific features in our progenitor cells. We next defined markers that distinguish blastocystic trophoblast from first-trimester placental trophoblast (Figures S1A and S1B)[8,25] and found that our iPSC-derived trophoblast (both Peri-TB and PrSyn) expresses unique markers of blastocystic trophoblast (*PRR9*, *CCKBR*, and *MRGPRX1*) and not those of the placenta (*FSTL3*, *ASCL2*, *DIO2*, *CSH1*, and *PSG4*) (Figure 1F), indicating that our trophoblast model better aligns with the earliest trophoblast progression during the peri-implantation stage of development. *HAND1*, a known marker of the ExMC lineage in OWM,[14] initiates expression in our iPSC-derived cells at the late progenitor stage and becomes restricted to the ExMC (Figure 1D), in line with the transient trophoblast expression of HAND1 in the day-6 to −7 human blastocyst.[10] RNA velocity analysis confirmed the trajectories through measured developmental time and linked both PrSyn and ExMC from day 6 to a shared precursor population (LP) at day 4 (Figure 1G).

We next integrated our dataset with seven published scRNA-seq datasets, four derived from OWM and three from human studies.[8,10,11,14,26–28] Our putative ExMC overlapped with ExMC lineages from both human and OWM (Figure 1H). Similarly, our trophoblast (Peri-TB and PrSyn) mapped with the published trophoblast subtypes, including those defined as cytotrophoblast (CTB), syncytiotrophoblast (STB), and extravillous trophoblast (EVT) (Figure 1I). Amniotic cells defined in these other studies formed a cluster distinct from any of our iPSC-derived cell derivatives (Figure 1J). To preclude caveats of cross-species analysis, we also integrated solely with the human data, which provided better separation of trophoblast developmental stages. Our iPSC-derived progenitor cells (cells transitioning between days 2 and 3 of differentiation, Figure S1J) mapped to the same cluster (cL.6) as trophectoderm from day-5 and −6 blastocysts (Figures S2E–S2I). Our defined ExMC population again co-clustered with ExMC identified in the post-implantation human embryo from days 16–19 (Figures S1C and S1D). Taken together, these results show that ExMC and PrSyn share the same progenitor population in the BS-treated iPSC differentiation scheme, and this progenitor population displays gene-expression signatures found in the trophectoderm of the human blastocyst.

Activin A promotes ExMC specification from progenitor cells

Since the BMP4/FGFR-inhibitor-based model gives rise to a mixed culture of ExMC and trophoblast subtypes, we next aimed to optimize media composition for induction of pure

ExMC. We used differentially expressed genes (DEGs) between our day-5 to −6 ExMC and trophoblast populations (454 genes up in trophoblast and 245 genes up in ExMC; Figure 2A) as input for Ingenuity Pathway Analysis (IPA), and found evidence of transforming growth factor β (TGF-β) signaling activity in ExMC development with multiple components in this pathway rising to significance (TGFB1, ZEB1, SMAD3, and SNAI1/2; Figure 2B). We next tested the TGF-β ligand activin A in our experimental system, comparing addition on day 0 (BSA-D0) to that on day 2 (BSA-D2). The BSA-D0 cultures maintained expression of pluripotency markers, consistent with a known role for activin A in pluripotency[29]; however, the BSA-D2 cultures significantly enhanced ExMC differentiation at the expense of the trophoblast lineage (Figure 2C). We next included an opposing differentiation scheme by replacing the TGF-β pathway activator with an inhibitor of this pathway (A-83) at day 2 (BSA-83-D2). Transient CDX2 expression, peaking at day 2.5, was similar in all three (BS, BSA-D2, and BSA-83-D2) conditions. SNAI2, an ExMC-specific TF, began to increase at day 3, was highest in the BSA-D2 conditions, and preceded the upregulation of the ExMC-specific *COL3A1*. Inhibiting TGF-β signaling (BSA-83-D2) induced elevated expression of the later trophoblast marker HLA-G and inhibited the expression of ExMC markers (Figure S2A), suggesting that the manipulation of the TGF-β signaling pathway within the initial trophoblast progenitor stage determines the developmental outcome.

Bulk transcriptome data of the day-6 time point in all three of these conditions, referred to now as BS, BSA, and BSA-83, were then compared to the human-derived scRNA-seq datasets. While the first component of a principal component analysis (PCA) represented technology differences (bulk vs. single cell), the second and third principal components were able to separate ExMC from trophoblast. Cells derived from BSA-treated iPSCs clustered with ExMC from our scRNA-seq dataset and the ExMC from the Carnegie stage 7 embryos. Both bulk RNA-seq BS and BSA-83 samples clustered together with other trophoblast lineage cells (Figure 2D). Extensive gene-expression differences (983 downregulated and 2,619 upregulated genes in the bulk RNA-seq) were evident in BSA vs. BS treatments, inclusive of cell-lineage markers initially identified in our scRNA-seq analysis (Figure 2E). We complemented some of these RNA markers with their corresponding protein epitopes, detecting both trophoblast ($PRR9^+$ and $HLA\text{-}G^+$) and ExMC ($HAND1^+$) cells in the initial differentiation scheme (BS condition) and virtually only $HAND1^+$ cells in the BSA cultures with no evidence for $PRR9^+/HLA\text{-}G^+$ cells (Figure 2F).

The purity of our ExMC, and its apparent emergence from an epithelial progenitor, prompted us to explore splicing dynamics (Figures S2B and S2C). Enriched in the ExMC population was an *MPRIP* isoform that is favored in metastatic pancreatic ductal adenocarcinoma (PDAC) compared to primary PDAC[30] (Figure S2D). In PDAC this isoform is regulated by splicing factor RBFOX2, and *RBFOX2* may also be responsible for this splicing in ExMC, as it is differentially expressed between ExMC and trophoblast (Figure 2E). Additional splicing switches include the extracellular matrix (ECM) protein *COL6A3* and adhesion junction protein *DST* (Figure S2D). Mutations in *DST* are associated with epidermolysis bullosa,[31] a disease manifesting in disorders of the skin mesenchyme.[32] The splicing architecture of the ExMC provides an additional mesenchymal feature attributed to these cells, and the dynamic nature of the cell model, deriving from precursor epithelium

representative of the blastocystic trophoblast, provides opportunities to further investigate splicing features of epithelial-mesenchymal transition (EMT).

Considering the unexpected emergence of an ExMC population from a trophoblast progenitor cell, we next sought to confirm the robustness of this differentiation scheme across additional human iPSC lines to ensure it was not an iPSC-line-specific oddity. Consistent with KOLF2.C1, two additional iPSC lines, WTC11, a well-characterized male line of Japanese ancestry[33] and WIBJ2, a female Northern European line, were capable of differentiating to mature trophoblast and ExMC lineages when induced with the BSA-83 and BSA conditions, respectively (Figure S3A). This indicates that our defined differentiation conditions are robust across genetic background and sex.

Thus, BSA culture conditions efficiently direct iPSCs to the ExMC lineage through an apparent trophoblast intermediate, and the activity of the TGF-β pathway after initiation into the trophoblast lineage but prior to day 4 of differentiation plays a key role in the decision to become, through an EMT process, ExMC or continue along the trophoblast lineage.

Trophoblast bona fides of cells derived from primed iPSCs

Although there has been some debate about the authenticity of trophoblast derived from "primed" human PSCs,[22,34] substantial evidence reveals highly specific trophoblast features of such cells,[17,19,35] for instance chorionic gonadotropin and the genes (*CGA* and the *CGB*s) encoding this pregnancy hormone. In our bulk RNA-seq data of iPSC-derived PrSyn, *CGA* is the most highly expressed protein coding transcript of all, with over 30,000 FPKM (fragments per kilobase of transcript per million mapped reads). An additional trophoblast bona fide of our iPSC-derived PrSyn is the expression of *CYP19A1* from its trophoblast-specific promoter (Figure S3B). This LTR-derived promoter was co-opted into the human lineage approximately 40 million years ago[36] and, though not present in the mouse genome, is capable of driving trophoblast-specific transgenic reporter expression in the mouse.[37–39]

Recent reports also indicate the feasibility of deriving human trophoblast stem-like cells (hTSLCs) from primed iPSCs.[40,41] We thus decided to replicate this work with our KOLF2.C1 iPSC line to provide further evidence of trophoblast bona fides. We were successful in this, specifically deriving hTSLCs from KOLF2.C1 iPSCs exposed to BMP4 for 2 days. Our hTSLCs displayed morphology similar to and expression level of pantrophoblast markers comparable to that of the original human trophoblast stem cell (hTSC) lines[42] CT29 and bTS11 (Figures 3A and 3B). Our hTSLCs remained proliferative over at least 60 days of culture (Figure 3D) and had approximately 50% DNA methylation status of the C19MC locus, indicative of the trophoblast lineage (Figure 3E). Using markers representative of early EVT (HLA-G, DIO2, FSTL3, and ISM2) and STB/PrSyn (SDC1, ERVW-1, and PRR9) found within extended-culture human blastocysts[10] (Figure 3F) indicated that our hTSLC lines were capable of differentiating into both EVT and STB, a defining feature of hTSCs (Figures 3G–3L).

As the ExMC appears to emerge from a trophoblast progenitor population in our initial scRNA-seq, further supported by the fate of such progenitors in the HAND1 null line (see below), we were curious to determine whether hTSC/hTSLC lines themselves harbored this

potential to give rise to ExMC. To placenta-derived (CT29) and blastocyst-derived (bTSC11) hTSCs, we added activin A and analyzed gene expression. These data suggest that indeed ExMC can arise from hTSCs (Figure 2G). We performed a similar experiment with our hTSLC cultures, which also provided evidence of the emergence of ExMC (Figure 3M). Together, these data provide evidence that KOLF2.C1 gives rise to a bona fide trophoblast, that an ExMC population arises from this lineage, and that this ability to generate ExMC is maintained in the mature TSC.

HAND1 is essential for ExMC specification

To identify potential regulators of ExMC specification, we analyzed the dynamic expression of 211 TFs expressed in our iPSC-to-PrSyn/ExMC model. *HAND1* emerged as a strong candidate, with the highest fold change and expression level compared to other ExMC-specific TFs (Figure 4A), and because of its known roles in both the trophoblast and ExMC lineages in the mouse.[43,44] In our differentiation system, HAND1 is the earliest upregulated ExMC TF, preceding the bifurcation point in ExMC emergence (Figure 4B) and co-localizing with trophoblast TFs (CDX2 and GATA3) in the progenitor cells (Figure 4C), placing it in position to drive ExMC formation.

Thus, we created homozygous null HAND1 iPSCs, differentiated these along with wild-type (WT) cells in ExMC medium, and collected samples from day 2 to day 6 for transcriptome analysis (Figure 4D). No significant differences were observed between WT and *HAND1*$^{-/-}$ cells on day 2, consistent with its lack of expression at this time point. After day 3, WT and *HAND1*$^{-/-}$ cells displayed distinct transcriptomic trajectories, with differences increasing over time. Three major gene clusters were identified by hierarchical clustering representing (1) genes shared by WT and *HAND1*$^{-/-}$ cells and downregulated from the day-2 time point (e.g. *SOX2*, *NANOG*, and *OCT4*), (2) genes upregulated in WT cells, and (3) genes upregulated in *HAND1*$^{-/-}$ cells. Subclusters of these latter two major clusters separated based on the timing of upregulation of their respective genes. Subclusters II, III, and IV were WT-cell specific and enriched in ExMC genes upregulated at days 3, 4, and 6 of differentiation, respectively. *HAND1* was in the first wave. The second wave (day 4) included *JUN* and *SNAI2*. *HAND2* and *ZEB2* TFs were upregulated on day 6 of differentiation as was *RUNX1*, which is known to maintain mesenchymal stem cells.[45] In addition, a number of ECM components show significantly increasing expression on days 4 and 6, recapturing the pattern with their *in vivo* counterparts (Figure S4). Subclusters V, VI, and VII consisted predominantly of trophoblast genes upregulated in the *HAND1*$^{-/-}$ cells on days 3, 4, and 6, respectively. For instance, there was a transient upregulation of *TP63* on day 3 and upregulation of *GCM1* and *CCR7* by day 4, and *PGF*, *CCKBR*, *PRR9*, *CYP19A1*, and chorionic gonadotropins (*CGA* and *CGB1*) on day 6 (Figure 4E). *CGA* expression dynamics exemplified this apparent continuation along the trophoblast lineage upon loss of *HAND1* (Figure 4F). Immunofluorescence confirmed, at the protein level, the absence of *SNAI2* expression and the increase in trophoblast markers (*HLA-G* and *PRR9*) within these *HAND1*$^{-/-}$ cell cultures at day 6 (Figures 4G and 4H). These results demonstrate that HAND1 is an essential regulator of the emergence of ExMC from a trophoblast progenitor.

HAND1 binding is enriched at proximal promoters of ExMC lineage-specific genes

To identify genomic features through which HAND1 may mediate ExMC differentiation, we next identified HAND1-bound regions of the genome through chromatin immunoprecipitation and sequencing (ChIP-seq) on day 4 of differentiation, generating H3K4me3 ChIP-seq in parallel and carrying this out on both WT and *HAND1*$^{-/-}$ cell lines. We chose day 4, as this timing was within 24 h of the first appearance of *HAND1* mRNA, reasoning that this would provide sufficient time for HAND1 protein synthesis and subsequent accumulation at genomic loci but early enough to capture the primary developmental events mediated by HAND1. A total of 5,127 high-confidence (present in all replicates of WT and in none of the knockout [KO]) HAND1 binding sites were identified. Of the defined HAND1 peaks, nearly half were located in promoter regions (within 2 kb of a transcription start site), 22.1% were within introns, and 18.1% were in intergenic regions (Figure 5A). Motif discovery on these peaks identified the consensus sequence AAACCAGA (Figure 5B) containing a previously defined HAND1-associated element, CCAGA.[46,47]The relatively weak and short *cis* element on what otherwise are high-confidence HAND1-bound regions likely results from the known degeneracy of the HAND1 half-site.[46]

Although the overall landscape of H3K4me3 binding sites were similar between WT and the null with only 144 genotype-unique peaks out of the 10,369 total peaks (Figure 5C), differences were observed when we classified these H3K4me3-bound loci as co-bound or not bound by HAND1. A higher H3K4me3 signal was evident in the HAND1-bound subset compared between WT and *HAND1*$^{-/-}$ cells but not in the HAND1-unbound peaks (Figure 5D), indicating that chromatin-bound HAND1 is associated with increased levels of H3K4me3 modifications, a histone mark of active promoter regions (Figures 5A and 5D). Linking expression data to HAND1 ChIP-seq, we identified 40% of downregulated genes and 15.1% of upregulated genes on day 3 in *HAND1*$^{-/-}$ cells that were directly bound by HAND1, suggesting direct activation and repression by HAND1 binding, respectively (Figure 5E). The few genes directly repressed by HAND1 included the human trophoblast lineage marker (*CYP19A1*) and a marker of polar trophectoderm (*CCR7*) (Figure 5F). In support of our data, it is notable that *CYP19A1* was recently defined as repressed by HAND1 in a human trophoblast cell line.[48] Numerous ExMC markers directly activated by HAND1 included *TGFBR1*, *NID2*, *ZEB2*, and *ACTC1*. These data suggest that HAND1 acts predominantly as a transcriptional activator in its role in ExMC formation.

HAND1 binds to ape-specific LTR2B

We also noted 16% of HAND1 peak summits mapped to transposable elements, out of which 43% were LTRs and therefore over-represented compared to LINEs and SINEs (long and short interspersed nuclear elements) with respect to their relative abundance in the human genome.[49] Considering the known contribution of retrotransposons to the human trophoblast lineage,[50] we next sought to determine the subgroup of LTRs mediating the recruitment of HAND1. We found that LTR2B was significantly over-represented, with 10% (35 out of the 353) of all known instances of this LTR subgroup bound by HAND1 (Figure 6A). The identification of LTR2B is intriguing, considering the rapid evolution of ExMC and the recent integration timing of LTR2B at the hominoid divergence from OWM.[51]

LTRs can function as enhancers or promoters and/or create additional first exons with their self-contained features of *cis*-regulatory elements, transcription start site, and a splicing donor site. To elucidate the function of HAND1-bound LTR2B, we manually annotated all 35 instances in the context of the surrounding genomic and transcriptomic features (Table S5). Examples of apparent enhancer activity of HAND1-bound LTR2Bs, as they are in close proximity to genes significantly downregulated in the *HAND1*$^{-/-}$ cells, include those associated with *ACTC1*, *ERVH48–1*, *CNTNAP2*, and *FUT8* and have fold downregulation in HAND1 null cells of 517.15, 1.55, 15.81, and 2.21, respectively. *ERVH48–1*, also known as *Suppressyn* (*SUPYN*), is reported to restrict infection by retroviruses and suppression of syncytial cell fusion of trophoblast in early human embryos and developing placenta.[52,53] FUT8, a glycolytic enzyme responsible for core fucosylation, is implicated in tumor metastasis across many cancer types, including melanoma,[54] thus again linking ExMC biology to features identified as functional in cancer metastasis. Additionally, we further identified some unannotated transcripts with LTR2B as the promoter/first exon, occurring in nearly a third (10 out of the 35) of HAND1-bound LTR2B (Table S5). Among these genes *PTN* is a notable example, since this isoform was previously described as a unique feature of human trophoblast[55] and, as defined here, the ExMC (Figures S5A, S5B, and S5G) but not in amnion epithelial cells from human post-implantation embryos, thus providing a useful marker to distinguish amnion from trophoblast and ExMC (Figures S5E and S5F).[11] Another intriguing example is *DCAF6*. Its dominant transcript in ExMC starts from the LTR2B located in an intron in *ADCY10* (Figures S5C and S5D), splices ~88 kb downstream into exon 2 of *DCAF6*, and produces a different N-terminal end. This 88-kb length spans 14 exons of *ADCY10*, the entire 5-exon gene *MPC2*, and the first exon of *DCAF6*. This isoform of *DCAF6* is downregulated in the *HAND1*$^{-/-}$ cells and is also expressed in the human post-implantation ExMC (Figures S5C and S5H). Such modification/truncation of a parental protein resulting from an insertion of LTR2B into an internal exon is particularly emphasized in the case of *ANKS1A*.

For *ANKS1A*, transcription is initiated from a HAND1-bound LTR2B sequence located within intron 7 of the parental gene; this LTR2B sequence splices into the immediate downstream exon. The expression of this transcript tightly correlates with what would be expected of a HAND1 ExMC-specific target gene (Figures 6B–6E) and is absent in OWM ExMC (Figure 6C), fitting with the ape-specific insertion event of this LTR. The resulting 15-exon transcript encodes a 593-amino-acid protein detectable by western blot (Figure 6F) and is much smaller than the parental protein, which contains 541 additional N-terminal amino acids and is detectable in undifferentiated iPSCs (Figure 6F). This shortened ExMC isoform does not contain the ankyrin repeat domain (ANK) or the transmembrane domain, although it does retain the two tandem sterile alpha motif (SAM) domains and one phosphotyrosine-binding (PTB) cytoplasmic domain (Figure 6F).

To further characterize the function of this ape-specific truncated isoform of *ANKS1A*, we deleted this LTR2B sequence via CRISPR-Cas9 and differentiated this *ANKS1A*$^{LTR2B-/-}$ line using the ExMC induction protocol (i.e., BSA). From RNA-seq analysis, the loss of this LTR2B alone is sufficient to abolish the expression of the short isoform of ANKS1A without any effect on its full-length transcript (Figure 6G). Although ANKS1A is not known to be a direct regulator of transcription, we still identified 188 DEGs between

WT and *ANKS1A*$^{LTR2B-/-}$lines in RNA-seq analysis (Figure 6I). The majority of the 97 downregulated genes are day-6 ExMC markers and are also absent in similarly differentiated *HAND1*$^{-/-}$ cells as would be expected of a transcript dependent on HAND1 function (Figure 6H). Notably, the most downregulated gene (MEG3, a long non-coding RNA) has been implicated in EMT processes within the HTR8/SVneo trophoblast cell line.[56] Half of the 91 upregulated genes showed much lower expression levels in ExMC compared to day 2 of differentiation, and the remaining upregulated genes did not show a specific pattern (Figure 6J). These results indicate a functional role for the HAND1-bound LTR2B-derived ANKS1A transcript in the biology of the differentiated ExMC and, together with a number of the other HAND1-bound LTR2B elements, they have contributed to primate/hominoid-specific features of ExMC biology.

Integration of ExMC biology within the peri-implantation conceptus

Mesenchymal cells have long been known as a source of, and receiver for, signals that induce and pattern neighboring cell types in developing tissues.[57] Thus, we next sought out potential signaling axes by identifying ligand-receptor pairs in relevant cell types to contextualize the potential role of the ExMC. *NODAL* and *GDF3*, activin-like molecules expressed within human epiblast/primitive endoderm of the peri-implantation embryo (Figure 7A), are potential endogenous inductive signals for the formation of ExMC *in vivo*, similar to activin A-induced *in vitro* formation of ExMC. Another potential axis of signaling is between the trophoblast and ExMC where in the former, progesterone is synthesized from cholesterol through the activity of enzymes encoded by *CYP11A1* and *HSD3B1*. The progesterone receptor *PAQR8*[58] is highly expressed in ExMC cells (Table S4). In addition, ExMC, together with the trophoblast, synthesize *PTN* (Figure 7B), which enhances the survival and stemness in human PSCs through its receptor PTPRZ1.[59] The expression of *PTPRZ1* is seen in the epiblast of the late-stage human blastocyst and in gastrulating embryos after the emergence of ExMC but is absent in the earlier-stage inner cell mass (Figure 7A), suggesting a role for ExMC in providing a growth-promoting signal to the neighboring epiblast. In addition to *PTN*, a variety of growth factors (e.g., *FGF2*, *NOG*, and *INHBA*) along with ECM known to contribute to the maintenance of human pluripotency[29,60,61] were abundantly expressed by the ExMC (Figure 7C). This suggested that the ExMC may provide a microenvironmental niche to the epiblast through peri-implantation development. To test this, we used a co-culture model with iPSC-derived ExMC as the soil in which to seed undifferentiated iPSCs, in basal medium without any additional growth factor supplementation. Unlike iPSCs cultured on a layer of iPSC-derived trophoblast, which died within a day of seeding (Figure 7D), iPSCs on ExMC expanded into sizable colonies over 4 days of culture (Figure 7D) and maintained high expression of pluripotency markers OCT4 and SOX2, with colonies surrounded by COL3A1-positive ExMC (Figure 7E). These data thus suggest roles for the ExMC within the microenvironment of the peri-implantation human embryo, providing a supportive ECM and a source of growth factors in addition to integrating signals itself from neighboring cell types.

DISCUSSION

Human iPSC differentiation schemes to model the trophoblast lineage are becoming more widely accepted by the community,[40,41,62–65] but there remains some debate on whether these models, particularly derived from primed iPSC cultures, might be more representative of amnion cells.[66–69] Here, we distinguish three extra-embryonic lineages (trophoblast, amnion, and ExMC) by integrating multiple existing scRNA-seq datasets of true and blastocyst-derived cell types of the early primate conceptus. Mapping iPSC-derived scRNA-seq data onto this provides strong evidence of trophoblast identity as does the highly specific trophoblast features we characterize here. In line with previous reports,[40,41] the ability to derive hTSLCs from the primed iPSC line used here also argues that primed iPSCs are capable of giving rise to the trophoblast lineage. In addition, primed OWM embryonic stem cells transferred into OWM blastocysts were identified in the trophectoderm and persisted in the trophoblast lineage to day 17.[70] This indicates the *in vivo* trophoblast potential of simian primed PSCs, a characteristic not shared with their mouse counterparts.

While studies have suggested that “primed” hPSC-derived trophoblasts pass through an amnion-like state,[71] it remains challenging to draw definitive conclusions without data on true perito early post-implantation human conceptuses that capture the developmental dynamics at play in the formation of these cell types. *In vivo*, the amnion is not thought to give rise to the trophoblast lineage, whereas the epiblast within the late-stage human blastocyst, a stage in which some “primed” iPSC markers begin to be expressed, are still capable of giving rise to the trophoblast. Within our iPSC-based system we favor a model whereby BMP4 induction directly induces the trophoblast lineage through direct upregulation of trophoblast TFs (*GATA2*, *GATA3*, *TFAP2A*, and *TFAP2C*), leading undifferentiated iPSCs to progressively become more like cells within the early trophoblast lineage. Although the initial activation of trophoblast TFs via BMP4 used *in vitro* may differ from *in vivo* trophoblast formation, features, such as transient *FABP3*/*CDX2* expression preceding an *NR2F2*$^{+}$/*CCR7*$^{+}$ cell population that then further progresses to primitive syncytium, parallel the expression dynamics seen within the blastocystic trophoblast. The eventual primitive syncytium signature is very reminiscent of the polar trophectoderm within the late-stage blastocyst.

The addition to this paradigm is the emergence of the ExMC from the FABP3/CDX2/HAND1^{+} progenitor. Trophectoderm from the day-6 and day-7 human blastocyst are FABP3/CDX2/HAND1^{+}. We showed that by manipulating TGF-β signaling, these progenitor cells in our iPSC system are toggled between commitment to an ExMC or trophoblast. In addition, the loss of HAND1 leads this progenitor population to remain in the trophoblast lineage even under ExMC induction medium. We also show that such progenitor cells can further be stabilized as hTSLCs with full differentiation potential to both mature trophoblast (EVT, STB/PrSyn) and ExMC. The apparent ExMC potential is likely maintained in established trophoblast cell lines like the hTSCs shown in this study and HTR8/SVneo and JEG-3.[72] While more evidence is required to fully confirm, taken together, our findings suggest the human trophoblast lineage may be the source of the first ExMC lineage to form *in vivo*, in cell ontology terms this likely represents the fibroblast of villous mesenchyme (CL:0002558). We do not exclude the existence of other sources of

ExMC such as epiblast,[73] hypoblast, and amnion, multiple waves of ExMC specification throughout embryonic development are likely. Considering the rapid evolution of ExMC among species and the lack of feasibility of tracing cell lineage during natural human implantation, *ex vivo* implantation models of the human blastocyst will be a potential future solution to further uncover this biology.

We identify HAND1 as an essential TF driving iPSC-derived ExMC formation. While this cellular fate decision is unique from that in the mouse, HAND1 does have roles in both the mouse trophoblast and ExMC,[43,44] and thus there has likely been modification from an ancestral role that HAND1 has played in extra-embryonic lineages. In the human trophoblast, HAND1 appears to establish the ability of trophoblast to respond to TGF-β by, in part, directly controlling the expression of TGFBR1. TGF-β signaling is a well-known inducer of EMT, and it is an EMT mechanism from which the ExMC emerges from the precursor trophoblast, as indicated by HAND1 regulation of a suite of well-known EMT TFs (*SNAI1*, *SNAI2*, and *ZEB1*). While HAND1 is transiently detected in the blastocyst trophectoderm prior to day 6 but not in their derivatives like CTB, EVT, and STB,[10,74] current single-cell datasets of the human blastocyst do not identify an ExMC population. This may be a result of the PrSyn pathway being favored over ExMC differentiation, as PrSyn has the initial primary role of mediating attachment to, and invasion of, the endometrium. ExMC may arise shortly thereafter, potentially dependent on endometrial signals or signals from the maturing epiblast (e.g., GDF3 and NODAL)—unquestionably, there is extensive ExMC within 2 days after implantation.[5] We provide data supporting a role for this ExMC in the maintenance and expansion of the epiblast. This specialized niche ensures the proper development of the early conceptus through peri-implantation.

There is a remarkable diversity in the extra-embryonic lineages and the implantation process across primates. For instance, the depth of invasion of the ape blastocyst into the endometrium is much greater than in the OWM.[15] Considering these ape-specific features, it is noteworthy that we identify LTR2B as at least part of the human genomic features that mediate HAND1 function in the ExMC, as this class of endogenous retroviral element is largely ape specific.[51] Whether the ExMC contributes to the highly invasive nature of the human blastocyst awaits further investigation, although it is intriguing that a number of the HAND1-bound LTR2B-regulated genes we identify here have also been implicated in cancer metastasis.[54,75] Regardless, our study adds to the repertoire of LTR classes[76] that recruit cell-type-specific TFs to contribute to the primate-specific features of early development and points to the creativity of these LTR co-option events in not only altering transcription but also creating additional protein isoforms, of which examples here include *ANKS1A* and *DCAF6*.

Cellular models representative of early human development are essential for understanding embryo implantation and the significant loss of viable embryos that are associated with these early stages of embryonic development.[77] It is essential to have human-based models for this, as there is such great variation in early development between species. A directed differentiation scheme we provide here to achieve relatively uniform ExMC cultures enables the genetic dissection of mechanistic programs. Adaptation of recent 3D culture systems of early human development[78–82] to include the ExMC will assist in further understanding

the developmental roles played by the ExMC. The iPSC-to-ExMC-directed differentiation system also has utility in dissecting the molecular dynamics in EMT, such as the alternative splicing pattern we have shown. Finally, considering the recent evolutionary innovation within the extra-embryonic lineage, human-based cellular systems capable of modeling these features will be an essential component toward a functional understanding of aspects of the human genome[83] that may not be recapitulated in other valuable model systems such as the mouse.

Limitations of the study

Currently, the early emergence of ExMC is not captured in either *in vivo* implanting human embryos or the extended cultured human blastocysts *in vitro*. Thus, it is challenging to distinguish ExMC populations from different embryonic stages, and the potential multiple origins of ExMC remain to be confirmed by *in vivo* evidence. In addition, ExMC can form a specialized niche to support the pluripotency of undifferentiated human iPSCs in our co-culture system. Further validation of this function of ExMC in natural embryos and its relevance to early pregnancy loss is still needed.

STAR★METHODS

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Human iPSC lines (KOLF2.C1, WTC11 and WIBJ2) were obtained from the Cellular Engineering Scientific Service Core at The Jackson Laboratory where SNP array analysis is routinely carried out for authentication purposes to confirm normal karyotype status and line identification. KOLF2.C1 is derived from skin fibroblast of a white British male in the 55–59 year age range. WIBJ2 is derived from skin fibroblast of a white British female in the 55–59 year age range. WTC11 is derived from skin fibroblast of an Asian male in the 30–34 year age range. All cell lines were grown in StemFlex media (Gibco, #A3349401) on SynthemaxII (Corning, #3535) coated plates at 5% CO_2, 20% O_2 and under humidified conditions at 37°C. Briefly, StemFlex media was balanced to room temperature (RT) for 30 min before using and plates were coated by the SynthemaxII (2 mL per well for 6-well plate at final concentration of 25 μg/mL) at RT for 2 h. KOLF2.C1 cells were thawed at 37°C, washed once by StemFlex media supplemented with RevitaCell (Gibco, #A2644501), and then seeded in coated plates. Media were changed every two days and cells were passaged every 4–5 days when they reached 80% confluency. Cell can be passaged in clumps by treatment with ReLeSR (STEMCELL, #05872) at RT for 5 min or in single cell by treatment with Accutase (ThermoFisher, #00-4555-56) at 37°C for 7 min. RevitaCell or Rock inhibitor Y-27632 (10 μM, TOCRIS, #1254) was added in the StemFlex media when the cells were suspended as single cell. Test of mycoplasma contamination was performed every 2 months.

METHOD DETAILS

Generation of HAND1$^{-/-}$ cell line: To generate the specific *HAND1* KO iPSC line we followed established human stem cell CRISPR/Cas9 editing protocols with minor modifications. Briefly, specific CRISPR/Cas9 sites 5′-GGTGTGCGTAGCTGCCCACG/AGG and 5′- GTGAGCCGAGGCCCGCGCCG/AGG that flank exons 1 and 2 of *HAND1* were identified in the human genome using the online

tool WGE (https://wge.jax.org/wge/). The resulting design was aimed at deleting the entire coding region of this gene, through CRISPR-mediated double-stranded breaks and non-homologous end-joining (NHEJ) repair. Cas9 ribonucleoprotein (RNP) nucleofection was carried out using a single cell suspension of KOLF2.C1 human iPSCs, single guide RNAs specific to the identified CRISPR sites (Synthego, modified sgRNA), recombinant Cas9 protein (IDT, HiFi v3), and Primary Cell P3 buffer (Lonza) using the Amaxa Nucleofector 4D (Lonza), Program CA-137. Cells were immediately recovered in StemFlex medium (Gibco) containing 1x RevitaCell (Thermo) for 1 day. The cells remained in culture in StemFlex for six additional days until they reached 80% confluency. Single cell-derived clones were recovered as described (Skarnes et al. 2019) and were individually genotyped by target site PCR and Sanger sequencing using the primers listed in Table S1. Clones of interest were further expanded for two passages to generate replicate stock vials.

Trophoblast and ExMC differentiation from iPSC: KOLF2.C1 cells were passaged as single cells with Accutase and seeded onto SynthemaxII coated 6-well plates at 25–50k cells per well. 24-h after cell passage, StemFlex media was changed to trophectoderm induction media (TE media). Basal TE media consists of DMEM/F12, supplemented with 20% KnockOut Serum (ThermoFisher, #10828028), 2 mM L-Glutamine (Gibco, #25030081), 1x MEM non-essential amino acid (Gibco, #11140050), 0.1 mM β-Mercaptoethanol (Sigma, #3148). In addition, 100 ng/mL BMP4 (R&D, #314-BP-050/CF), and 20 μM SU5402 (Millipore Sigma, # 57263) were added at day 0 and onwards to induce trophoblast lineage (BS condition). For ExMC induction, 100 ng/μL BMP4 and 20 μM SU5402 were added at day 0 and onwards, 100 ng/mL Activin A (PeproTech, #120–14E) was added at the end of day 2 (BSA condition). In BSA-83 condition, BMP4 and SU5402 were added from day 0 and 1 μM A-83 (TOCRIS, #2939) was added at the end of day 2. Ligands were added as indicated in the main text in other conditions. Media was refreshed every 2 days until the end of differentiation (day 6).

Derivation of human trophoblast stem cell like cell from iPSC: KOLF2.C1 cells were passaged with Accutase and seeded into SynthemaxII coated 6-well plate at 200 k cells per well in complete Stemflex media. 24 h after cell passage, stemflex media was changed to hTSC media [DMEM/F12 supplemented with 0.1 mM β-mercaptoethanol, 0.2% fetal bovine serum, 0.3% BSA, 1% ITS-X, 1.5 μg/mL L-ascorbic acid, 50 ng/mL EGF, 2 μM CHIR99021, 0.5 μM A83–01, 1 μM SB431542, 0.8 mM VPA, 5 μM Y26732] with 10 ng/ml BMP4 for induction. The induction medium was replaced daily while treating BMP4 for 2 days. On day 3, cells were passaged by TypLE (37°C for 5–7 min) when they reached 80–90% confluency at the split ratio 1:5 for 3 passages to Matrigel coated plates. Do not passage the cell at too low confluency and adjust the split ratio accordingly. After 3 passages, hTSLC colonies can be observed and these cells were passaged with TypLE into a non-coated plate in hTSC maintaining media [DMEM/F12 supplemented with 1% KSR, 0.15% BSA, 1% ITS-X, 35 μg/mL L-ascorbic acid, 25 ng/mL EGF, 2 μM CHIR99021, 5 μM A83–01, 2.5uM Y27632, 0.8 mM VPA] with 1ul iMatrix511. Within a few more passages (passage 6–7), stabled hTSLC can be cryopreserved in KSR with 10% DMSO or be used for further experiments. Usually, hTSLC grew faster in hTSC maintaining media than in

hTSC media. For each passage, cells can be treated by TypLE up to 15 min to ensure a full dissociation. Unlike iPSC, no cell lifter is needed during the passage of hTSLC cells.

Culture of hTSC: hTSC lines (CT29, bTS11) were obtained from RIKEN. Cells were originally thawed in hTSC medium and plated on the SynthemaxII coated plates. Medium was changed every two days and cells were passaged by TrypLE (1mL per well, 37°C for no longer than 15 min) at a split ratio between 1:3 to 1:10 depending on cell confluency.

Differentiation of hTSC and hTSLC to EVT: hTSC and hTSLC were differentiated into EVT as described previously with a few modifications.[41] Briefly, 6-well plates were coated by 1% Matrigel in DMEM/F12 media at 37°C for at least 1 h. Then hTSC and hTSLC cells were passaged on the coated plate at different density (70k per well for hTSC lines and 40k for hTSLC) in the EVT1 medium [DMEM/F12 supplemented with 0.1 mM 2-mercaptoethanol, 4%KSR, 0.3% BSA, 1% ITS-X, 0.5uM Thiazovivin, 7.5uM A83–01, 1uM XAV939, 0.5uM PD0325901, 0.2uM WNT-C59, 20uM Y-27632] with 2% Matrigel and 100 ng/ml NRG-1. After 3 days, EVT1 medium was replaced by the EVT2 medium [DMEM/F12 supplemented with 0.1mM 2-mercaptoethanol, 4% KSR, 0.3% BSA, 1% ITS-X, 7.5uM A83–01, 2.5uM Y-27632] with 0.5% Matrigel. EVT like cells appeared at around day 3 and proliferated rapidly after the medium change. All cells were collected for RNA extraction and RT-qPCR on day 4 and 6.

Differentiation of hTSC and hTSLC to STB: hTSC and hTSLC were differentiated into STB as described previously with a few modifications.[42] Briefly, 6-well plates were coated by SynthemaxII at room temperature for 2 h. Then hTSC cells were passaged on the coated plate at 100k cell per well in STB medium [DMEM/F12 supplemented with 0.1mM 2-mercaptoethanol, 0.3% BSA, 4% KSR, 1% ITS-X, 2.5uM Y-27632, 2uM Forskolin]. And hTSLC cells were passaged on the uncoated plate at 100k cell per well with 1ul iMatrix511 in STB medium. The medium was changed on day 3. All cells were collected for RNA extraction and RT-qPCR on day 4 and 6.

Differentiation of hTSC and hTSLC to ExMC: Briefly, 6-well plates were coated by SynthemaxII at room temperature for 2 h. Then hTSC cells were passaged on the SynthemaxII coated plate at 300k cell per well in hTSC basal medium [DMEM/F12 supplemented with 0.1mM 2-mercaptoethanol, 0.3% BSA, 4% KSR, 1% ITS-X, 2.5uM Y-27632] with 100 ng/ml Activin A. And hTSLC cells were passaged on the uncoated plate at 300k cell per well with 1ul iMatrix511 in ExMC medium. The medium was changed every two days. All cells were collected for RNA extraction and RT-qPCR on day 4 and 6.

Cell growth curve of hTSC and hTSLC cell: The hTSCs and hTSLC lines were maintained as indicated above on SynthemaxII coated plates in hTSC media or uncoated plate with iMatrix511 in hTSCM media respectively. Cells were passaged every 3–4 days at the split ratio between 1:3 to 1:8 as needed. Cells were counted in each passage by the Life Technologies Countess II Cell Counter and cumulative cell number was calculated by the cell number times the last split ratio.

Co-culture of iPSC and ExMC/Trophoblast: KOLF2.C1 were plated 2.5×10^4 cells per well in 6-well plates and then differentiated with BSA or BSA-83 protocol to ExMC and trophoblasts, respectively, as described above. On day 6, media was discarded, and cells washed with DPBS 5x to ensure the removal of all supplements (BMP4, SU5402 and Activin A or A-83) in the differentiation media. Undifferentiated KOLF2.C1 cells were passaged by ReLeSR in clumps and put on top of ExMC or trophoblast with hPSC basal media at the passage ratio of 1:10. hPSC basal media was prepared as following: DMEM/F12 (Gibco, # 11330032) supplemented with 20% KnockOut Serum, 2mM GlutaMax (Gibco, # 35050061), 1x MEM non-essential amnio acid (Gibco, #11140050), 0.1 mM β-Mercaptoethanol (Sigma, #3148). Media was then changed every day. All cells were fixed by 4% PFA for immunofluorescence at day 4 of co-culture.

Immunofluorescence (IF) analysis of TB and ExMC differentiation: The iPSCs were seeded at 4×10^3 cells per well in 96-well CellCarrier Ultra microplates (PerkinElmer, #6055302) and then differentiated in different conditions as indicated in the figures respectively. Differentiated culture was washed by DPBS (ThermoFisher, #14190250) twice, and then fixed with 4% paraformaldehyde (PFA) for 10 min at RT. At day 6 fixed cells were blocked and permeabilized with 10% donkey serum and 0.1% Triton X-100 in DPBS for 30–60 min at RT. Then cells were incubated with primary antibodies overnight at 4°C as following: 1:100 anti-PRR9 (Abcam, # ab121953), 1:100 anti-HAND1 (R&D, #AF3168), 1:100 anti-HLA-G (Abcam, #ab52455), 1:100 anti-SNAI2 (CST, #9585S), 1:100 anti-SOX2 (SantaCruz, #sc17320), 1:100 anti-COL3A1(Novus Biologicals, NB600–594), 1:100 anti-OCT3/4 (BD Biosciences, #560306). Then secondary antibodies donkey anti-mouse AF488 (Invitrogen, #A-21202), donkey anti-rabbit AF568 (Invitrogen, #A10042), and donkey anti-goat AF647 (Invitrogen, #A32849) were conjugated accordingly at RT for 1 h. Nuclei were counterstained with DAPI at 1 μg/mL for 10 min. Image was taken with an Opera PhenixTM High-Content Screening System (PerkinElmer) for confocal 3D imaging using a 203 objective.

RNA extraction and RT-qPCR: For cell harvest, media were discarded and then DPBS were added in the 6-well plate at 1 mL/well. Cells were detached from the plate by cell lifer and then centrifuged at 300 × g for 3 min at RT. Supernatant was removed carefully and the cell pellet was used for the total RNA extraction by NucleoSpin RNA Plus (Takara, #740984.250) according to the instructions. About 500–1000 ng RNA was reversed transcribed into cDNA with High-Capacity RNA-to-cDNA Kit (ThermoFisher #4387406) depending on the original RNA concentration. Then cDNA was diluted 1:40 in water and 4 μL of diluted cDNA was used for qPCR together with 1 μL primer mix (500 nM) and 5 μL SYBR super mix. GAPDH was used as reference gene and all relative expression levels of targeting genes are normalized to GAPDH using $2^{\wedge}$(-deltaCt). Three technical replicates were performed in each condition of qPCR. All sequences used in qPCR are listed in Table S2.

Western blot: To harvest cells, directly detach cells from the 6-well plates on day 6 of differentiation by cell lifter, transfer them to 1.5 mL Eppendorf tubes, centrifuge at 300g for 3 min and then wash them with DPBS twice. Put all cell pellet on ice and add 100ul RIPA

buffer to each tube, incubate for 30 min and then centrifuge at 4°C, 12000rpm for 12 min. Move the supernatant to a new 1.5 mL Eppendorf tube as the whole cell lysates. Take 2 μl of cell lysate for each sample, dilute at 1:10 and then measure the protein concentration by Pierce BCA Protein Assay Kit (ThermoFisher, #23227) following the instruction. Protein samples were prepared at the concentration of 0.9 μg/μl by adding 4x Laemmli sample buffer (BIO-RAD, #1610747) and boiling at 98°C for 10 min.

To run the gel, 15 μg of proteins were loaded on the pre-make gel and run in the 1x running buffer at 120V for 1 h. Proteins were transferred from gel to PVDF membrane in the transfer buffer at 0.3A for 1.5 h. Membrane was blocked by 5% purified milk at RT for 30 min and then incubated with primary and secondary antibody for 1.5 h separately and washed three times (5 min for each) at the interval. All membranes were imaged by BIO-RAD ChemiDocTM MP system with Clarity Max Western ECL Substrate (BIO-RAD, #1705062).

The following antibodies were used: anti-ANKS1A (BETHYL, # A303–050A, 1:1000), anti-HAND1 (R&D, #AF3168, 1:1000), anti-GAPDH (CST, #2118S, 1:2000), HRP-conjugated Affinipure Donkey Anti-Goat IgG(H + L) (Proteintech, #SA00001–3, 1:2000), HRP-conjugated Affinipure Donkey Anti-Rabbit IgG(H + L) (Proteintech, # SA00001–9, 1:2000).

RNA extraction, library preparation and sequencing: Wild type KOLF2.C1 iPSCs were seeded in 6-well plate at 5×10^4 cells/well and then differentiated in different conditions (BS, BSA, BSA-83) as described above. Differentiated cells were harvested at day 6 for total RNA extraction using NucleoSpin RNA Plus (Takara, #740984.250) according to the instructions. RNA-seq libraries were prepared with KAPA Stranded Total RNA-Seq kit (Roche) according to manufacturer's instruction. The samples were normalized to 300 ng input and spiked with ERCC. First, ribosomal RNA was depleted using RiboErase followed by DNase treatment. Purified RNA was then fragmented at 85°C for 6 min, targeting fragments range 250–300 bp. Fragmented RNA is reverse transcribed with an incubation of 25°C for 10 min, 42°C for 15 min, and an inactivation step at 70°C for 15 min. This was followed by second strand synthesis at 16°C, 60 min. Double stranded cDNA fragments were purified using AMPure XP beads (Beckman) and were then A-tailed and ligated with Illumina unique adaptors (Illumina). Adaptor-ligated DNA was purified using AMPure XP beads. This is followed by 10 cycles of PCR amplification. The final library was cleaned up using AMPure XP beads. Quantification of libraries were performed using real-time qPCR (Thermo Fisher). Sequencing was performed on Illumina NovaSeq 6000 S4 flow cell, targeting 20~30M 2×150 bp paired end reads per sample.

For time course bulk RNA sequencing. WT and HAND1−/− KOLF2.C1 were seeded in 6-well plate as 5×10^4 cells/well and then cultured in StemFlex media or differentiated at different time points to ensure collection of differentiated cells at the same day. Library preparation was done with the same protocol described above. Sequencing was performed on Illumina NovaSeq 6000 S4 flow cell, targeting 20~30M 2×100 bp paired end reads per time point.

Data pre-processing and differential expressed gene (DEG): Illumina basecall files were converted to FASTQ files using bcl2fastq v2.20.0.422 (Illumina). FASTQ reads were aligned to the GRCh38.p1 human reference assembly (GCF_000001405.27, annotations from Ensembl Release 93) by the nf-core/rnaseq (v3.0) pipeline using default parameters. Briefly, FASTQ files were processed by FastQC (0.11.9) for quality control and adapter sequences were trimmed by TrimGalore (0.6.6). After pre-processing of raw reads, alignment and quantification were done by STAR (2.6.1d) and Salmon (1.4.0), respectively. Entries in the resulting Salmon merged gene expression matrix scaled by both library size and gene length and this matrix was used as input for differential gene expression analysis by DESeq2 (1.28.0).

For comparison of different culture conditions (BS, BSA, BSA-83), BS condition was used as the control group. DEGs were identified by comparing BS and BSA and filtered by |Log2FoldChange| > 1.5 and padj <0.001.

For comparison of WT and HAND1−/−cell line at different time points, we first used the WT sample at day 2 as the control and compared it with WT samples at days 3, 4, and 6 to identify differentially expressed genes during ExMC differentiation. We then compared WT and *HAND1*$^{-/-}$ samples separately at each timepoint (day 3, 4, 6) to identify the effects of HAND1 deficiency on ExMC differentiation. Finally, we combined all DEGs and applied more stringent cutoff (|Log2FoldChange|>4 and padj<0.001) to identify the most significant DEGs. The heatmap in Figures 2E, 3A, and 3E was created using the clustermap function of the Seaborn (0.11.2) Python package, and gene modules were identified using unsupervised hierarchical clustering (average linkage).

RNA alternative splicing and Ingenuity Pathway Analysis (IPA): Sorted bam files from the nf-core/rna-seq pipeline were used as input for alternative splicing using rMATS (v4.0.1) with default parameters. Differential splicing peaks were visualized by Integrative Genomic Viewer (IGV, 2.12.3) and percent of spliced in (PSI) were quantified by rMATS. Signaling pathway and gene ontology analysis were done by Ingenuity Pathway Analysis (IPA; Content version: 52912811).

ChIP sequencing and analysis: ChIP was performed on ExMC at day 4 of differentiation. Briefly, 5×10^5 WT (three replicates) and HAND1−/− KOLF2.C1 (two replicates) cells were seeded on 10 cm dish and then differentiated in ExMC induction media (BSA) for 4 days. ChIP DNA samples were prepared with EZ-Magna ChIP HiSens Chromatin Immunoprecipitation Kit (Millipore Sigma, #17–10461), with parameter modification. Sonication was performed at 40% power, 20 s on and 50 s rest for 18 cycles, in 1 mL SCW buffer/sample included in the kit. Fragmented DNA was immunoprecipitated with anti-HAND1 (R&D, #AF3168) and anti-H3K4me3 antibody (included in EZ-Magna ChIP kit, #CS200580), anti-IgG (included in EZ-Magna ChIP kit, #CS200581) with overnight incubation at 4°C. Reverse crosslinking was performed per manufacture recommendations. Input DNA and immunoprecipitated DNA were cleaned up using Monarch PCR & DNA Clean up (NEB, #T1030S).

Samples were normalized to 10 ng, and for samples with insufficient yield, all the sample was taken for library preparation. Immunoprecipitated DNA was treated with end-repair, A-tailing, and ligation of Illumina compatible adapters (IDT) using the KAPA-Illumina library creation kit (KAPA biosystems). The ligated product was enriched with 8–15 cycles of PCR (Roche/KAPA Biosystems) followed by AMPure beads cleanup. Quantification of libraries were performed using real-time qPCR (Thermo Fisher). The final libraries were normalized and pooled. Quantification of library the pool was performed using real-time qPCR (KAPA and Thermo Fisher). The pool was then denatured and loaded on an Illumina NovaSeq 6000 S4 flow cell, targeting 10~15M 2×150bp read pairs per sample.

Reads were aligned to the same GRCh38 reference assembly described above, filtered, and quantified with nf-core/chipseq (v1.2.2) workflow using standard parameters and narrowPeak mode for MACS2 (v2.2.7.1) peak calling with the cutoff as padj<1e-3. Curated HAND1 binding peaks were identified from consensus peaks present in all three WT replicates but absent in both HAND1−/− replicates. Consensus binding motifs were identified by Homer (v4.11) with the following settings '-size 200 -mask -len 4,6,8 -p 4 -preparse'. For transposable element (TE) analysis, the genomic regions of HAND1 binding peaks were defined such that they contained the peaks from all three replicates to ensure that these regions were in the center of the peaks. The potential bound TEs were identified by overlapping TE regions from hg38 rmsk reference file (downloaded from UCSC genome browser) with all HAND1 binding regions. All peaks were visualized on IGV (v2.12.3). The LTR enrichment score was calculated by LTR_peak_percentage/LTR_frequency. LTR frequency means the ratio of LTR_copy_number/total_LTR_copy_ number across the whole genome. Similarly, LTR_peak_percentage was calculated as the number of HAND1 peaks overlapped with given LTR divided by number of all HAND1 peaks overlapped with any LTRs.

Single-cell RNA-seq analysis of iPSC-derived trophoblast: For differentiation, KOLF2.C1 cells were passaged by Accutase and seeded into 6-well plates in StemFlex media supplemented with RevitaCell. Differentiation started at 24 h after seeding cells by changing the StemFlex media to TE media with BMP4 and SU5402 and media were changed every two days as described above. Cells were seeded and differentiated at the same time and harvested at different time points. To ensure sufficient material for scRNA-seq, different numbers of cells were seeded per well for each time point as follows: 1 × 106 cells for day 1 and day 2, 4 × 105 cells for day 4, 2 × 105 cells for day 5 and day 6.

For preparation of 10x Genomics scRNA-seq, differentiated cells were washed by DPBS (2 mL/well) for 3 times and then treated with 1.2 mL Accutase at 37°C for 15 min. Cells from day 6 were treated by Accutase for 18 min. To stop the reaction, 1 mL filtered DMEM/F12 into was added to the well and a 1 mL pipette with wide bore tip was used to pipette up and down to get the cells off the plate bottom. The cell suspension was transferred to a 15 mL conical tube containing 9 mL filtered DMEM/F12. Cells were pipetted up and down 5x to further dissociate to single cells and spun down at 800rpm for 3 min. Supernatant was discarded and the cell pellet was resuspended in 2 mL filtered DMEM/F12, pipetted up and down 5 times, and filtered through a 40 μm cell strainer. Cells were spun down at 800 rpm for 3 min, supernatant removed, resuspended in DPBS with 0.4% BSA by weight, and

transferred into a 1.5 mL Eppendorf tube. Cell concentration and viability were assessed on a Countess II automated cell counter (ThermoFisher). Up to 12,000 cells were loaded onto one lane of a 10x Genomics Chip G. Single cell capture, barcoding and library preparation were performed using the 10x Genomics Chromium platform. cDNA and libraries were checked for quality on Agilent 4200 Tapestation, quantified by KAPA qPCR, and each library was sequenced independently on a separate lane of an Illumina HiSeq 4000 HO flow cell, targeting 6,000 barcoded cells with an average sequencing depth of 50,000 reads per cell.

Another batch of scRNA-seq dataset were generated with V3 chemistry, and with sampling cells at 12-h intervals of differentiation from day 2.5 to day 4.5. In brief, cells were seeded on a 6-well plate at 1 × 105 cells/well and differentiation initiated 24 h later. The initiation of differentiation was staggered so that all time points could be harvested at the same time. These samples were harvested, washed by PBS and then hash-tagged by lipid based oligos as described.[90] Briefly, anchor was conjugated with cholesterol, which can spontaneously bind into cell membrane and the barcode contained a 5′ complementary sequence to anchor and a unique sequence followed by oligoA at the 3′ end. For barcoding, approximately 5 × 10^5 or fewer cells were labeled by a unique barcode separately by co-incubation with the anchor and barcodes mixture (1:1, final concentration 2 mM) on ice for 5 min and another 5 min with co-anchor (2 μM) for stabilization of the hashtag. After barcoding, cells were washed twice by 1 mL of 1% cold BSA in PBS, centrifuged at 4°C. Before loaded to 10x chromium, all samples were counted.

Illumina base call files for all libraries were converted to FASTQs using bcl2fastq v2.16.0.10 (Illumina) and FASTQ files associated with the gene expression libraries were aligned to the GRCh38.93 reference genome (10x Genomics GRCh38 ref. 3.0.0) using the version 3.1.0 Cell Ranger count pipeline (10x Genomics). All libraries were then combined using the Cellranger aggr pipeline. Downstream analysis was performed with scanpy81 (v1.7.0). The raw feature-barcode matrix from the aggregation was filtered to remove cells with fewer than 2000 genes, library-size normalized, and log transformed. The 2000 most highly variable genes were selected for downstream analysis, including principal component analysis, batch correction with Harmony, nearest neighborhood graph generation, and UMAP embedding. RNA velocity analysis was performed by the scVelo (0.2.3) as previously described. Cell-cell interactions were identified by CellChat and the most significant ligand-receptor was shown as violin plot. All related packages are listed and cited above.

Integration of multiple scRNA seq datasets: Original cell lineage annotation was not available online in two studies and new annotations were generated by scanpy (v1.7.0) based on the cell markers found in original literature. For integration, all count matrixes were combined using the subset of approximately 10,000 overlapping genes and approximately 1,800 highly variable genes were used for downstream analysis: principal component analysis, batch correction with Harmony (correcting for effects induced by both study and sequencing platform), and UMAP embedding.

QUANTIFICATION AND STATISTICAL ANALYSIS

Quantification methods are described in individual figure legends. Results are indicated as means. Error bars represent either standard error of the mean (SEM) or 95% confidence interval as indicated in individual figure legends. N in the legends is the number of replicates. Statistical test was done by non-paired two sides t test in Prism8.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We gratefully acknowledge the Single Cell Biology, Genome Technologies, and Cellular Engineering services, and the Cyberinfrastructure high-performance computing resources at The Jackson Laboratory. These shared services are supported in part by the JAX Cancer Center (P30 CA034196). This study was funded by the National Institutes of Health grant 1UM1HG012651 (W.C.S. and P.R.), National Institutes of Health grant 1U54AG075941 (P.R.), National Institutes of Health grant 5P30CA034196 (P.R.), and Fulbright-Thailand Research Fund Junior Research Scholarship PHD/0123/2556 (P.P.).

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Paul Robson (paul.robson@jax.org).

Materials availability

The genetically modified human iPSC lines generated in this study are either available through https://www.jax.org/ipsc or available upon request from the lead contact with few restrictions, as outlined in *Terms and Conditions of Use* available at the listed website.

Data and code availability

- The datasets generated in this study have been deposited in the GEO database under accession number GEO: GSE220844. These data are publicly available as of the date of publication. Accession numbers are listed in the key resources table.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

REFERENCES

1. Mossman WH (1987). Vertebrate Fetal Membranes (Rutgers University Press).
2. Tam PP, and Beddington RS (1987). The formation of mesodermal tissues in the mouse embryo during gastrulation and early organogenesis. Development 99, 109–126. 10.1242/dev.99.1.109. [PubMed: 3652985]
3. Arnold SJ, and Robertson EJ (2009). Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. Nat. Rev. Mol. Cell Bio 10, 91–103. 10.1038/nrm2618. [PubMed: 19129791]

4. Eskes TKAB (1988). Developmental stages in human embryos. Eur. J. Obstet. Gyn. R. B 28, 356–357. 10.1016/0028-2243(88)90025-1.
5. Luckett WP (1978). Origin and differentiation of the yolk sac and extra-embryonic mesoderm in presomite human and rhesus monkey embryos. Am. J. Anat 152, 59–97. 10.1002/aja.1001520106. [PubMed: 98035]
6. Guo G, Huss M, Tong GQ, Wang C, Li Sun L, Clarke ND, and Robson P (2010). Resolution of Cell Fate Decisions Revealed by Single-Cell Gene Expression Analysis from Zygote to Blastocyst. Dev. Cell 18, 675–685. 10.1016/j.devcel.2010.02.012. [PubMed: 20412781]
7. Tang F, Barbacioru C, Bao S, Lee C, Nordman E, Wang X, Lao K, and Surani MA (2010). Tracing the Derivation of Embryonic Stem Cells from the Inner Cell Mass by Single-Cell RNA-Seq Analysis. Cell Stem Cell 6, 468–478. 10.1016/j.stem.2010.03.015. [PubMed: 20452321]
8. Petropoulos S, Edsgärd, Reinius B, Deng Q, Panula SP, Code-luppi 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. 10.1016/j.cell.2016.03.023. [PubMed: 27062923]
9. 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. 10.1016/j.stem.2021.04.027. [PubMed: 34004179]
10. 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. 10.1038/s41586-019-1875-y. [PubMed: 31830756]
11. Tyser RCV, 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. 10.1038/s41586-021-04158-y. [PubMed: 34789876]
12. Hertig AT, Rock J, and Adams EC (1956). A description of 34 human ova within the first 17 days of development. Am. J. Anat 98, 435–493. 10.1002/aja.1000980306. [PubMed: 13362122]
13. Enders AC, and King BF (1988). Formation and differentiation of extra-embryonic mesoderm in the rhesus monkey. Am. J. Anat 181, 327–340. 10.1002/aja.1001810402. [PubMed: 3389303]
14. 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. 10.1038/nature19096. [PubMed: 27556940]
15. Ross C, and Boroviak TE (2020). Origin and function of the yolk sac in primate embryogenesis. Nat. Commun 11, 3760. 10.1038/s41467-020-17575-w. [PubMed: 32724077]
16. Hislop J, Song Q, Kamyar Keshavarz F, Alavi A, Schoenberger R, LeGraw R, Velazquez JJ, Mokhtari T, Taheri MN, and Rytel M (2023). Modelling post-implantation human development to yolk sac blood emergence. Nature 626, 367–376. 10.1038/s41586-023-06914-8. [PubMed: 38092041]
17. Xu R-H, Chen X, Li DS, Li R, Addicks GC, Glennon C, Zwaka TP, and Thomson JA (2002). BMP4 initiates human embryonic stem cell differentiation to trophoblast. Nat. Biotechnol 20, 1261–1264. 10.1038/nbt761. [PubMed: 12426580]
18. Krendl C, Shaposhnikov D, Rishko V, Ori C, Ziegenhain C, Sass S, Simon L, Müller NS, Straub T, Brooks KE, et al. (2017). GATA2/3-TFAP2A/C transcription factor network couples human pluripotent stem cell differentiation to trophectoderm with repression of pluripotency. Proc. Natl. Acad. Sci. USA 114, E9579–E9588. 10.1073/pnas.1708341114. [PubMed: 29078328]
19. Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, and Jones JM (1998). Embryonic Stem Cell Lines Derived from Human Blastocysts. Science 282, 1145–1147. 10.1126/science.282.5391.1145. [PubMed: 9804556]
20. Jang YJ, Kim M, Lee B-K, and Kim J (2021). Induction of human trophoblast stem-like cells from primed pluripotent stem cells. Preprint at bioRxiv. 10.1101/2021.07.14.452371.
21. Amita M, Adachi K, Alexenko AP, Sinha S, Schust DJ, Schulz LC, Roberts RM, and Ezashi T (2013). Complete and unidirectional conversion of human embryonic stem cells to trophoblast by BMP4. Proc. Natl. Acad. Sci. USA 110, E1212–E1221. 10.1073/pnas.1303094110. [PubMed: 23493551]

22. Bernardo AS, Faial T, Gardner L, Niakan KK, Ortmann D, Senner CE, Callery EM, Trotter MW, Hemberger M, Smith JC, et al. (2011). BRACHYURY and CDX2 Mediate BMP-Induced Differentiation of Human and Mouse Pluripotent Stem Cells into Embryonic and Extraembryonic Lineages. Cell Stem Cell 9, 144–155. 10.1016/j.stem.2011.06.015. [PubMed: 21816365]

23. Blakeley P, Fogarty NME, del Valle I, Wamaitha SE, Hu TX, Elder K, Snell P, Christie L, Robson P, and Niakan KK (2015). Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. Development 142, 3151–3165. 10.1242/dev.123547. [PubMed: 26293300]

24. Niakan KK, and Eggan K (2013). Analysis of human embryos from zygote to blastocyst reveals distinct gene expression patterns relative to the mouse. Dev. Biol 375, 54–64. 10.1016/j.ydbio.2012.12.008. [PubMed: 23261930]

25. Vento-Tormo R, Efremova M, Botting RA, Turco MY, Vento-Tormo M, Meyer KB, Park J-E, Stephenson E, Polański K, Goncalves A, et al. (2018). Single-cell reconstruction of the early maternal–fetal interface in humans. Nature 563, 347–353. 10.1038/s41586-018-0698-6. [PubMed: 30429548]

26. 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. 10.1126/science.aax7890. [PubMed: 31672918]

27. 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. 10.1126/science.aaw5754. [PubMed: 31672917]

28. Yang R, Goedel A, Kang Y, Si C, Chu C, Zheng Y, Chen Z, Gruber PJ, Xiao Y, Zhou C, et al. (2021). Amnion signals are essential for mesoderm formation in primates. Nat. Commun 12, 5126. 10.1038/s41467-021-25186-2. [PubMed: 34446705]

29. Beattie GM, Lopez AD, Bucay N, Hinton A, Firpo MT, King CC, and Hayek A (2005). Activin A Maintains Pluripotency of Human Embryonic Stem Cells in the Absence of Feeder Layers. Stem Cell. 23, 489–495. 10.1634/stemcells.2004-0279.

30. Jbara A, Lin K-T, Stossel C, Siegfried Z, Shqerat H, Amar-Schwartz A, Elyada E, Mogilevsky M, Raitses-Gurevich M, Johnson JL, et al. (2023). RBFOX2 modulates a metastatic signature of alternative splicing in pancreatic cancer. Nature 617, 147–153. 10.1038/s41586-023-05820-3. [PubMed: 36949200]

31. Khalesi R, Harvey N, Garshasbi M, Kalamati E, Youssefian L, Vahidnezhad H, and Uitto J (2022). Pathogenic DST sequence variants result in either epidermolysis bullosa simplex (EBS) or hereditary sensory and autonomic neuropathy type 6 (HSAN-VI). Exp. Dermatol 31, 949–955. 10.1111/exd.14562. [PubMed: 35276021]

32. Vahidnezhad H, Youssefian L, Saeidian AH, and Uitto J (2019). Phenotypic Spectrum of Epidermolysis Bullosa: The Paradigm of Syndromic versus Non-Syndromic Skin Fragility Disorders. J. Invest. Dermatol 139, 522–527. 10.1016/j.jid.2018.10.017. [PubMed: 30393082]

33. Viana MP, Chen J, Knijnenburg TA, Vasan R, Yan C, Arakaki JE, Bailey M, Berry B, Borensztejn A, Brown EM, et al. (2023). Integrated intracellular organization and its variations in human iPS cells. Nature 613, 345–354. 10.1038/s41586-022-05563-7. [PubMed: 36599983]

34. Lee CQE, Gardner L, Turco M, Zhao N, Murray MJ, Coleman N, Rossant J, Hemberger M, and Moffett A (2016). What Is Trophoblast? A Combination of Criteria Define Human First-Trimester Trophoblast. Stem Cell Rep. 6, 257–272. 10.1016/j.stemcr.2016.01.006.

35. Yabe S, Alexenko AP, Amita M, Yang Y, Schust DJ, Sadovsky Y, Ezashi T, and Roberts RM (2016). Comparison of syncytiotrophoblast generated from human embryonic stem cells and from term placentas. Proc. Natl. Acad. Sci. USA 113, E2598–E2607. 10.1073/pnas.1601630113. [PubMed: 27051068]

36. van de Lagemaat LN, Landry J-R, Mager DL, and Medstrand P (2003). Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. Trends Genet. 19, 530–536. 10.1016/j.tig.2003.08.004. [PubMed: 14550626]

37. Kamat A, Graves KH, Smith ME, Richardson JA, and Mendelson CR (1999). A 500-bp region, ≈40 kb upstream of the human CYP19 (aromatase) gene, mediates placenta-specific expression in transgenic mice. Proc. Natl. Acad. Sci. USA 96, 4575–4580. 10.1073/pnas.96.8.4575. [PubMed: 10200304]

38. Kamat A, Hinshelwood MM, Murry BA, and Mendelson CR (2002). Mechanisms in tissue-specific regulation of estrogen biosynthesis in humans. Trends Endocrinol. Metab 13, 122–128. 10.1016/s1043-2760(02)00567-2. [PubMed: 11893526]
39. Wenzel PL, and Leone G (2007). Expression of Cre recombinase in early diploid trophoblast cells of the mouse placenta. Genesis 45, 129–134. 10.1002/dvg.20276. [PubMed: 17299749]
40. Jang YJ, Kim M, Lee B-K, and Kim J (2022). Induction of human trophoblast stem-like cells from primed pluripotent stem cells. Proc. Natl. Acad. Sci. USA 119, e2115709119. 10.1073/pnas.2115709119. [PubMed: 35537047]
41. Wei Y, Wang T, Ma L, Zhang Y, Zhao Y, Lye K, Xiao L, Chen C, Wang Z, Ma Y, et al. (2021). Efficient derivation of human trophoblast stem cells from primed pluripotent stem cells. Sci. Adv 7, eabf4416. 10.1126/sciadv.abf4416. [PubMed: 34380613]
42. 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. 10.1016/j.stem.2017.11.004. [PubMed: 29249463]
43. Riley P, Anson-Cartwright L, and Cross JC (1998). The Hand1 bHLH transcription factor is essential for placentation and cardiac morphogenesis. Nat. Genet 18, 271–275. 10.1038/ng0398-271. [PubMed: 9500551]
44. Firulli AB, McFadden DG, Lin Q, Srivastava D, and Olson EN (1998). Heart and extra-embryonic mesodermal defects in mouse embryos lacking the bHLH transcription factor Hand1. Nat. Genet 18, 266–270. 10.1038/ng0398-266. [PubMed: 9500550]
45. Kim W, Barron DA, Martin RS, Chan KS, Tran LL, Yang F, Ressler SJ, and Rowley DR (2014). RUNX1 is essential for mesenchymal stem cell proliferation and myofibroblast differentiation. Proc. Natl. Acad. Sci. USA 111, 16389–16394. 10.1073/pnas.1407097111. [PubMed: 25313057]
46. Hollenberg SM, Sternglanz R, Cheng PF, and Weintraub H (1995). Identification of a new family of tissue-specific basic helix-loop-helix proteins with a two-hybrid system. Mol. Cell Biol 15, 3813–3822. 10.1128/mcb.15.7.3813. [PubMed: 7791788]
47. Tsankov AM, Gu H, Akopian V, Ziller MJ, Donaghey J, Amit I, Gnirke A, and Meissner A (2015). Transcription factor binding dynamics during human ES cell differentiation. Nature 518, 344–349. 10.1038/nature14233. [PubMed: 25693565]
48. Zhu H, Ren Q, Yan Z, Xu S, Luo J, Wu X, and Tang C (2022). Human HAND1 inhibits the conversion of cholesterol to steroids in trophoblasts. J. Genet. Genomics 49, 350–363. 10.1016/j.jgg.2021.07.014. [PubMed: 34391879]
49. Nurk S, Koren S, Rhie A, Rautiainen M, Bzikadze AV, Mikheenko A, Vollger MR, Altemose N, Uralsky L, Gershman A, et al. (2022). The complete sequence of a human genome. Science 376, 44–53. 10.1126/science.abj6987. [PubMed: 35357919]
50. Senft AD, and Macfarlan TS (2021). Transposable elements shape the evolution of mammalian development. Nat. Rev. Genet 22, 691–711. 10.1038/s41576-021-00385-1. [PubMed: 34354263]
51. Le Dantec C, Vallet S, Brooks WH, and Renaudineau Y (2015). Human Endogenous Retrovirus Group E and Its Involvement in Diseases. Viruses 7, 1238–1257. 10.3390/v7031238. [PubMed: 25785516]
52. Frank JA, Singh M, Cullen HB, Kirou RA, Benkaddour-Boumzaouad M, Cortes JL, Garcia Pérez J, Coyne CB, and Feschotte C (2022). Evolution and antiviral activity of a human protein of retroviral origin. Science 378, 422–428. 10.1126/science.abq7871. [PubMed: 36302021]
53. Sugimoto J, Sugimoto M, Bernstein H, Jinno Y, and Schust D (2013). A novel human endogenous retroviral protein inhibits cell-cell fusion. Sci. Rep 3, 1462. 10.1038/srep01462. [PubMed: 23492904]
54. Agrawal P, Fontanals-Cirera B, Sokolova E, Jacob S, Vaiana CA, Argibay D, Davalos V, McDermott M, Nayak S, Darvishian F, et al. (2017). A Systems Biology Approach Identifies FUT8 as a Driver of Melanoma Metastasis. Cancer Cell 31, 804–819.e7. 10.1016/j.ccell.2017.05.007. [PubMed: 28609658]
55. Schulte AM, Lai S, Kurtz A, Czubayko F, Riegel AT, and Wellstein A (1996). Human trophoblast and choriocarcinoma expression of the growth factor pleiotrophin attributable to germ-line insertion of an endogenous retrovirus. Proc. Natl. Acad. Sci. USA 93, 14759–14764. 10.1073/pnas.93.25.14759. [PubMed: 8962128]

56. Wang R, Zou L, and Yang X (2021). microRNA-210/Long non-coding RNA MEG3 axis inhibits trophoblast cell migration and invasion by suppressing EMT process. Placenta 109, 64–71. 10.1016/j.placenta.2021.04.016. [PubMed: 33990028]
57. Grobstein C (1953). Inductive Epithelio-mesenchymal Interaction in Cultured Organ Rudiments of the Mouse. Science 118, 52–55. 10.1126/science.118.3054.52. [PubMed: 13076182]
58. Kasubuchi M, Watanabe K, Hirano K, Inoue D, Li X, Terasawa K, Konishi M, Itoh N, and Kimura I (2017). Membrane progesterone receptor beta (mPRb/Paqr8) promotes progesterone-dependent neurite outgrowth in PC12 neuronal cells via non-G protein-coupled receptor (GPCR) signaling. Sci. Rep 7, 5168. 10.1038/s41598-017-05423-9. [PubMed: 28701790]
59. Soh BS, Song CM, Vallier L, Li P, Choong C, Yeo BH, Lim EH, Pedersen RA, Yang HH, Rao M, and Lim B (2007). Pleiotrophin Enhances Clonal Growth and Long-Term Expansion of Human Embryonic Stem Cells. Stem Cell. 25, 3029–3037. 10.1634/stemcells.2007-0372.
60. Wang G, Zhang H, Zhao Y, Li J, Cai J, Wang P, Meng S, Feng J, Miao C, Ding M, et al. (2005). Noggin and bFGF cooperate to maintain the pluripotency of human embryonic stem cells in the absence of feeder layers. Biochem. Bioph. Res. Co 330, 934–942. 10.1016/j.bbrc.2005.03.058.
61. Rodin S, Antonsson L, Niaudet C, Simonson OE, Salmela E, Hansson EM, Domogatskaya A, Xiao Z, Damdimopoulou P, Sheikhi M, et al. (2014). Clonal culturing of human embryonic stem cells on laminin-521/E-cadherin matrix in defined and xeno-free environment. Nat. Commun 5, 3195. 10.1038/ncomms4195. [PubMed: 24463987]
62. Mischler A, Karakis V, Mahinthakumar J, Carberry CK, San Miguel A, Rager JE, Fry RC, and Rao BM (2021). Two distinct trophectoderm lineage stem cells from human pluripotent stem cells. J. Biol. Chem 296, 100386. 10.1016/j.jbc.2021.100386. [PubMed: 33556374]
63. Soncin F, Morey R, Bui T, Requena DF, Cheung VC, Kallol S, Kittle R, Jackson MG, Farah O, Chousal J, et al. (2022). Derivation of functional trophoblast stem cells from primed human pluripotent stem cells. Stem Cell Rep. 17, 1303–1317. 10.1016/j.stemcr.2022.04.013.
64. Viukov S, Shani T, Bayerl J, Aguilera-Castrejon A, Oldak B, Sheban D, Tarazi S, Stelzer Y, Hanna JH, and Novershtern N (2022). Human primed and naïve PSCs are both able to differentiate into trophoblast stem cells. Stem Cell Rep. 17, 2484–2500. 10.1016/j.stemcr.2022.09.008.
65. Ahern DT, Bansal P, Armillei MK, Faustino IV, Kondaveeti Y, Glatt-Deeley HR, Banda EC, and Pinter SF (2022). Monosomy X in isogenic human iPSC-derived trophoblast model impacts expression modules preserved in human placenta. Proc. Natl. Acad. Sci. USA 119, e2211073119. 10.1073/pnas.2211073119. [PubMed: 36161909]
66. Kobayashi N, Okae H, Hiura H, Kubota N, Kobayashi EH, Shibata S, Oike A, Hori T, Kikutake C, Hamada H, et al. (2022). The microRNA cluster C19MC confers differentiation potential into trophoblast lineages upon human pluripotent stem cells. Nat. Commun 13, 3071. 10.1038/s41467-022-30775-w. [PubMed: 35654791]
67. Zheng Y, Yan RZ, Sun S, Kobayashi M, Xiang L, Yang R, Goedel A, Kang Y, Xue X, Esfahani SN, et al. (2022). Single-cell analysis of embryoids reveals lineage diversification roadmaps of early human development. Cell Stem Cell 29, 1402–1419.e8. 10.1016/j.stem.2022.08.009. [PubMed: 36055194]
68. Guo G, Stirparo GG, Strawbridge SE, Spindlow D, Yang J, Clarke J, Dattani A, Yanagida A, Li MA, Myers S, et al. (2021). Human naive epiblast cells possess unrestricted lineage potential. Cell Stem Cell 28, 1040–1056.e6. 10.1016/j.stem.2021.02.025. [PubMed: 33831366]
69. 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. 10.1016/j.stem.2021.03.013. [PubMed: 33831365]
70. Cao J, Li W, Li J, Mazid MA, Li C, Jiang Y, Jia W, Wu L, Liao Z, Sun S, et al. (2023). Live birth of chimeric monkey with high contribution from embryonic stem cells. Cell 186, 4996–5014.e24. 10.1016/j.cell.2023.10.005. [PubMed: 37949056]
71. Ohgushi M, Taniyama N, Vandenbon A, and Eiraku M (2022). Delamination of trophoblast-like syncytia from the amniotic ectodermal analogue in human primed embryonic stem cell-based differentiation model. Cell Rep. 39, 110973. 10.1016/j.celrep.2022.110973. [PubMed: 35732132]

72. Abou-Kheir W, Barrak J, Hadadeh O, and Daoud G (2017). HTR-8/SVneo cell line contains a mixed population of cells. Placenta 50, 1–7. 10.1016/j.placenta.2016.12.007. [PubMed: 28161053]
73. Xiao Z, Cui L, Yuan Y, He N, Xie X, Lin S, Yang X, Zhang X, Shi P, Wei Z, et al. (2024). 3D reconstruction of a gastrulating human embryo. Cell 187, 2855–2874.e19. 10.1016/j.cell.2024.03.041. [PubMed: 38657603]
74. Liu D, Chen Y, Ren Y, Yuan P, Wang N, Liu Q, Yang C, Yan Z, Yang M, Wang J, et al. (2022). Primary specification of blastocyst trophectoderm by scRNA-seq: New insights into embryo implantation. Sci. Adv 8, eabj3725. 10.1126/sciadv.abj3725. [PubMed: 35947672]
75. Feng Z-J, Gao S-B, Wu Y, Xu X-F, Hua X, and Jin G-H (2010). Lung cancer cell migration is regulated via repressing growth factor PTN/RPTP β/ζ signaling by menin. Oncogene 29, 5416–5426. 10.1038/onc.2010.282. [PubMed: 20639902]
76. Frost JM, Amante SM, Okae H, Jones EM, Ashley B, Lewis RM, Cleal JK, Caley MP, Arima T, Maffucci T, and Branco MR (2023). Regulation of human trophoblast gene expression by endogenous retroviruses. Nat. Struct. Mol. Biol 30, 527–538. 10.1038/s41594-023-00960-6. [PubMed: 37012406]
77. Rossant J, and Tam PPL (2022). Early human embryonic development: Blastocyst formation to gastrulation. Dev. Cell 57, 152–165. 10.1016/j.devcel.2021.12.022. [PubMed: 35077679]
78. Yu L, Wei Y, Duan J, Schmitz DA, Sakurai M, Wang L, Wang K, Zhao S, Hon GC, and Wu J (2021). Blastocyst-like structures generated from human pluripotent stem cells. Nature 591, 620–626. 10.1038/s41586-021-03356-y. [PubMed: 33731924]
79. Kagawa H, Javali A, Khoei HH, Sommer TM, 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. 10.1038/s41586-021-04267-8. [PubMed: 34856602]
80. Liu X, Tan JP, Schröder J, Aberkane A, Ouyang JF, Mohenska M, Lim SM, Sun YBY, Chen J, Sun G, et al. (2021). Modelling human blastocysts by reprogramming fibroblasts into iBlastoids. Nature 1–6. 10.1038/s41586-021-03372-y.
81. 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. 10.1016/j.stem.2021.04.031. [PubMed: 33957081]
82. Karvas RM, Zemke JE, Ali SS, Upton E, Sane E, Fischer LA, Dong C, Park KM, Wang F, Park K, et al. (2023). 3D-cultured blastoids model human embryogenesis from pre-implantation to early gastrulation stages. Cell Stem Cell 30, 1148–1165.e7. 10.1016/j.stem.2023.08.005. [PubMed: 37683602]
83. Adli M, Przybyla L, Burdett T, Burridge PW, Cacheiro P, Chang HY, Engreitz JM, Gilbert LA, Greenleaf WJ, Hsu L, et al. (2025). MorPhiC Consortium: towards functional characterization of all human genes. Nature 638, 351–359. 10.1038/s41586-024-08243-w. [PubMed: 39939790]
84. Wolf FA, Angerer P, and Theis FJ (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biol 19, 15. 10.1186/s13059-017-1382-0. [PubMed: 29409532]
85. Love MI, Huber W, and Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DEseq2. Genome Biol 15, 550. 10.1186/s13059-014-0550-8. [PubMed: 25516281]
86. Shen S, Park JW, Lu Z, Lin L, Henry MD, Wu YN, Zhou Q, and Xing Y (2014). rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. Proc National Acad Sci 111, E5593–E5601. 10.1073/pnas.1419161111.
87. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, Baglaenko Y, Brenner M, Loh P, and Raychaudhuri S (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods 16, 1289–1296. 10.1038/s41592-019-0619-0. [PubMed: 31740819]
88. Bergen V, Lange M, Peidli S, Wolf FA, and Theis FJ (2020). Generalizing RNA vevocity to transient cell states through dynamical modeling. Nat. Biotechnol, 1–7. 10.1038/s41587-020-0591-3. [PubMed: 31919444]
89. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan C-H, Myung P, Plikus MV, and Nie Q (2021). Inference and analysis of cell-cell communication using. CellChat. Nat Commun 12, 1088. 10.1038/s41467-021-21246-9. [PubMed: 33597522]

90. McGinnis CS, Patterson DM, Winkler J, Conrad DN, Hein MY, Srivastava V, Hu JL, Murrow LM, Weissman JS, Werb Z, et al. (2019). MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indicies. Nat. Methods 16, 619–626. 10.1038/s41592-019-0433-8. [PubMed: 31209384]

Highlights

- Fast and efficient derivation of extra-embryonic mesenchyme from human iPSCs
- Defined HAND1 as the essential regulator for ExMC specification from human iPSCs
- LTR2B produces ape-specific isoforms of ExMC-related genes
- ExMC forms a specialized niche sufficient to support pluripotency

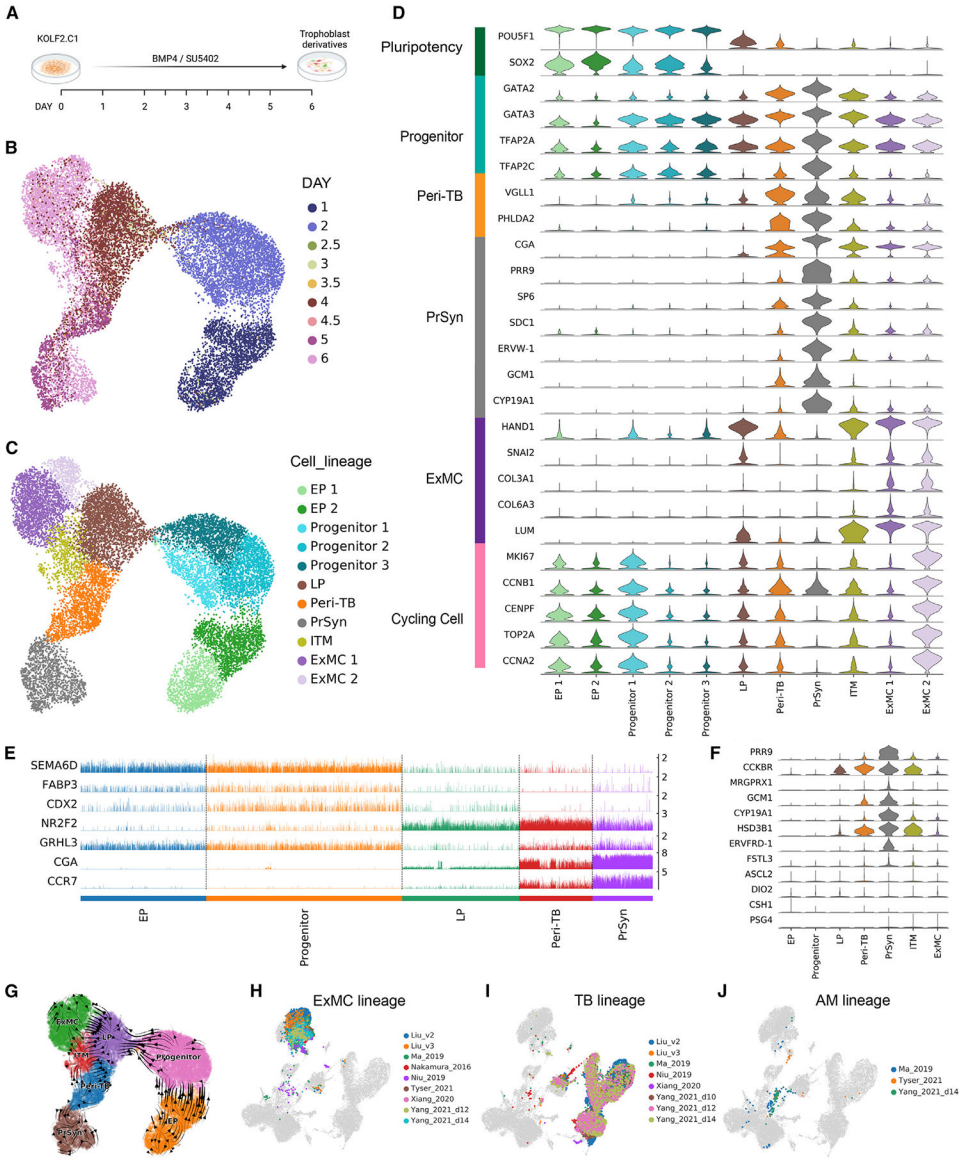


Figure 1. Identification of ExMC from BMP4- and FGFR-inhibitor-treated iPSCs

(A) Schematic of time-course scRNA-seq on the BMP4/FGFRi-induced trophoblast cell model.

(B) UMAP plot shown as different time points.

(C) UMAP plot on 11 cell clusters with manual annotation of cell lineages. Cells were identified as early progenitors (EP), progenitors, late progenitors (LP), peri-implantation trophoblast (Peri-TB), primitive syncytium (PrSyn), extra-embryonic mesenchymal cells (ExMC), and intermediates (ITM).

(D) Representative markers for each lineage.

(E) Sequential activation of trophoblast markers.

(F) Violin plot of the expression of stage-specific trophoblast markers identified in blastocyst and placenta.

(G) Differentiation trajectory of BMP4/FGFRi-treated iPSCs revealed by RNA velocity analysis.

(H–J) Integration of scRNA-seq datasets from published studies and distinct clustering of three extra-embryonic cell lineages (ExMC) (H), trophoblast lineage including Peri-TB and PrSyn (TB) (I), and amnion (AM) (J) in each study.

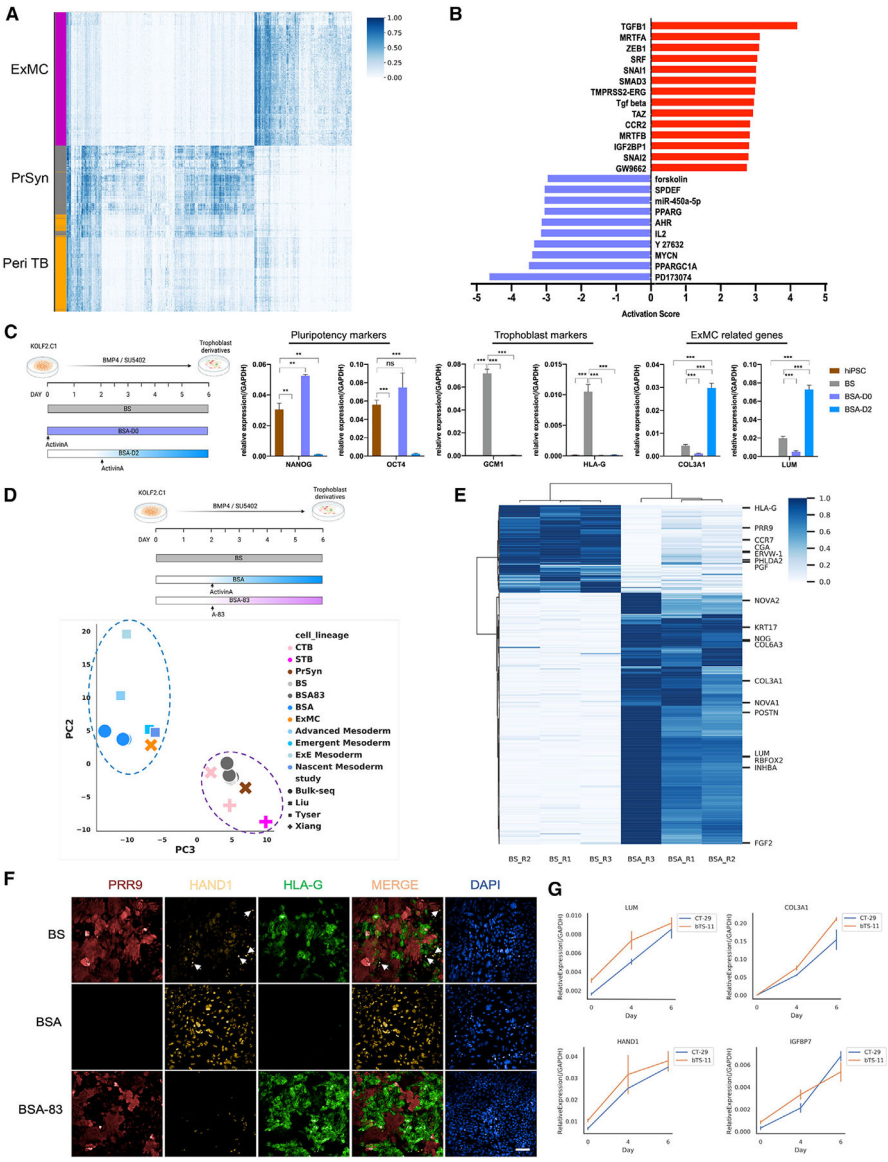


Figure 2. Activin A promotes ExMC specification from trophoblasts

(A) Heatmap of differentially expressed genes (DEGs, $p_{adj} < 0.05$, $|\log_2$ fold change (FC)$| > 1.5$) between ExMC and trophoblast subtypes (PrSyn and Peri-TB) identified in scRNA-seq.

(B) Upstream regulator analysis by IPA with DEGs identified in (A) as input.

(C) Expression of marker genes quantified by RT-qPCR in optimized ExMC induction protocols. Error bars from $n = 3$ for each condition calculated as standard error of the mean (SEM). A non-paired two-sided t test was used to determine significance. ns, not significant; $^{**}p < 0.01$, $^{***}p < 0.001$.

(D) PCA of RNA-seq data at day 6 of differentiation of KOLF2.C1 iPSC under the three conditions indicated in the schematic, together with ExMC and trophoblast cell clusters identified from *in vitro* cultured blastocyst (Xiang) and *in vivo* implanted human embryos (Tyser).

(E) Heatmap of DEGs between BS- and BSA-induced iPSC differentiation (as in D), with three biological replicates for each condition.
(F) Immunofluorescence of three lineage-specific markers (PRR9 and HLA-G for trophoblasts and HAND1 for ExMC) on day 6 of differentiation. White arrowheads indicate $HAND1^{+}$ ExMC in BS condition. Scale bar, 100 mm.
(G) Expression of ExMC markers on ExMC derived from hTSCs quantified by time-course RT-qPCR. Error bars from $n = 3$ for each condition calculated as 95% confidence interval.

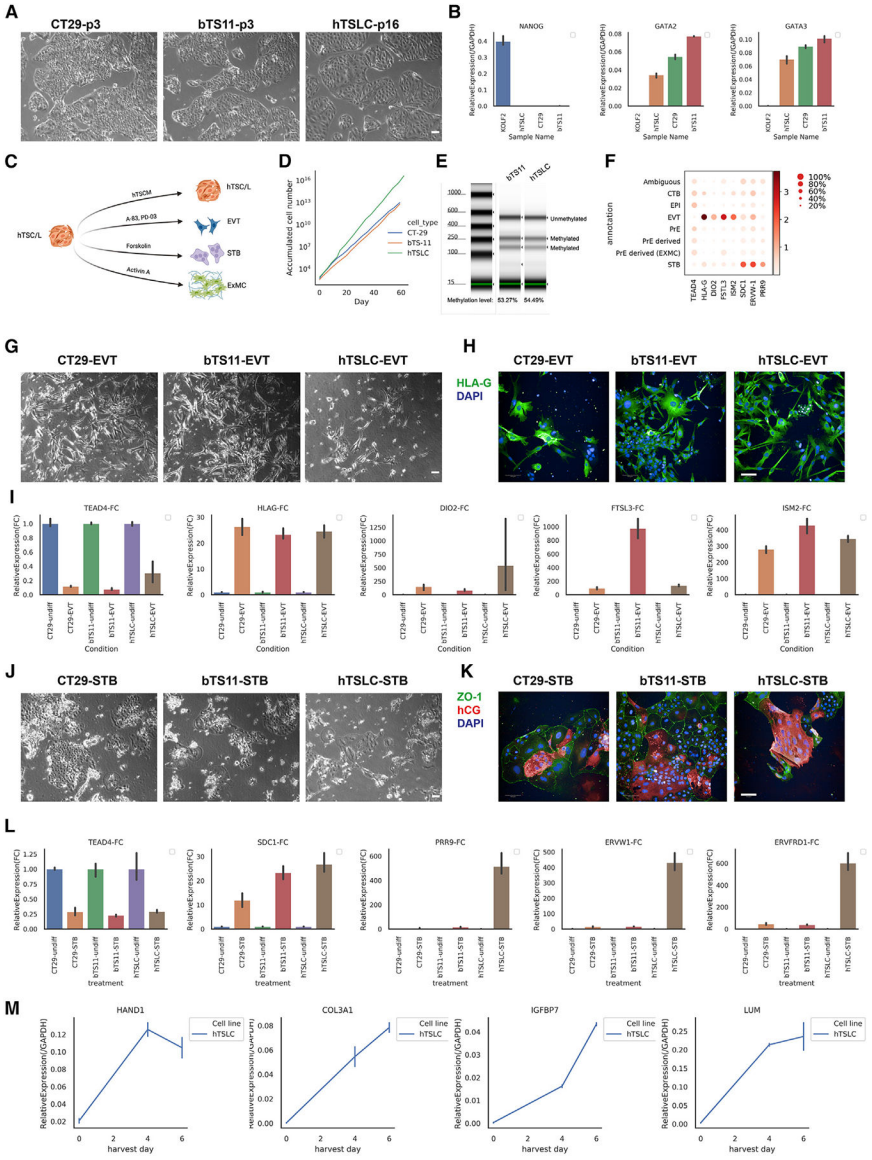


Figure 3. KOLF2.C1 iPSCs retain trophoblast potential

(A) Phase-contrast images of hTSC lines (CT29 and bTS11) and hTSLCs. Scale bar, 100 mm.

(B) RT-qPCR of pluripotency marker NANOG and hTSC markers GATA2/GATA3. Error bar denotes 95% confidence interval ($n = 3$).

(C) Schematic of differentiation schemes tested.

(D) Growth curve of hTSC lines and hTSLCs.

(E) Analysis of DNA methylation in C19MC DMR by COBRA in bTS11 and hTSLCs.

(F) Identification of trophoblast-specific genes within the extended *in vitro* cultured human blastocyst data in the study of Xiang et al.[10]

(G) Phase-contrast images of EVTs derived from hTSC/L. Scale bar, 100 mm.

(H) Immunofluorescence of EVTs. Scale bar, 100 mm.

(I) RT-qPCR analysis of TSC marker (TEAD4) and EVT markers in undifferentiated and EVT differentiation conditions. Relative expression level is normalized by their own undifferentiated hTSC/L lines. FC, fold change. Error bar denotes 95% confidence interval ($n = 3$).

(J) Phase-contrast images of STBs derived from hTSC/L. Scale bar, 100 mm.

(K) Immunofluorescent detection of hCG, ZO-1, and DAPI on STBs. Scale bar, 100 mm.

(L) RT-qPCR analysis of TSC marker (TEAD4) and STB markers in undifferentiated and STB differentiation conditions. Relative expression level is normalized by their own undifferentiated hTSC/L lines. Error bar denotes 95% confidence interval ($n = 3$).

(M) RT-qPCR time-course analysis of ExMC markers on hTSLCs differentiated in ExMC conditions. Error bar denotes 95% confidence interval ($n = 3$).

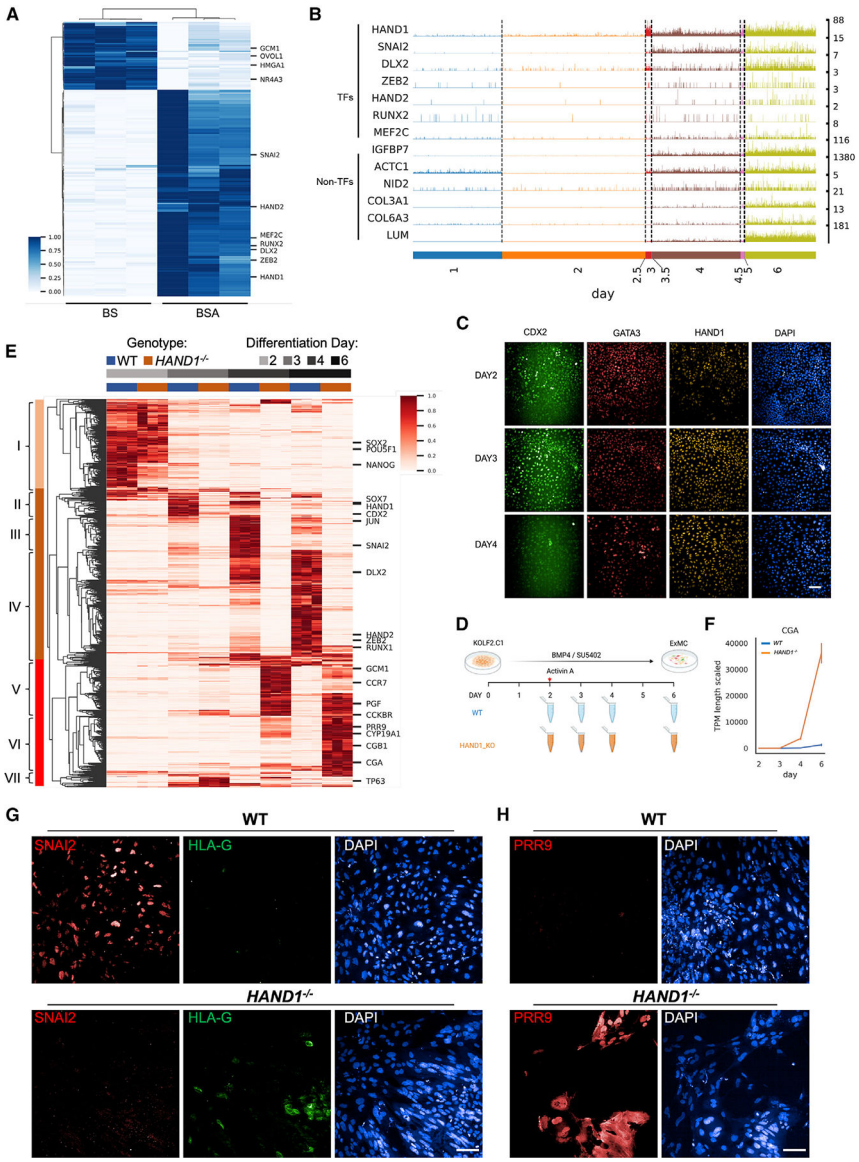


Figure 4. HAND1 is essential for ExMC specification

(A) Heatmap of top differentially expressed TFs between iPSC-derived trophoblasts (BS) and ExMC (BSA).

(B) Expression of key genes across differentiation from scRNA-seq (Figure 1C), with later trophoblast lineage (ITM, Peri-TB, and PrSyn) removed.

(C) Immunofluorescence detection of CDX2, GATA3, HAND1, and DAPI on days 2, 3, and 4 of ExMC differentiation. Scale bar, 100 mm.

(D) Schematic of samples taken for RNA-seq analysis.

(E) Heatmap of the top DEGs from samples in (D).

(F) CGA expression dynamics in this dataset. Error bar denotes 95% confidence interval (n = 3).

(G and H) Immunofluorescence at day 6 of indicated markers. Scale bar, 100 μm.

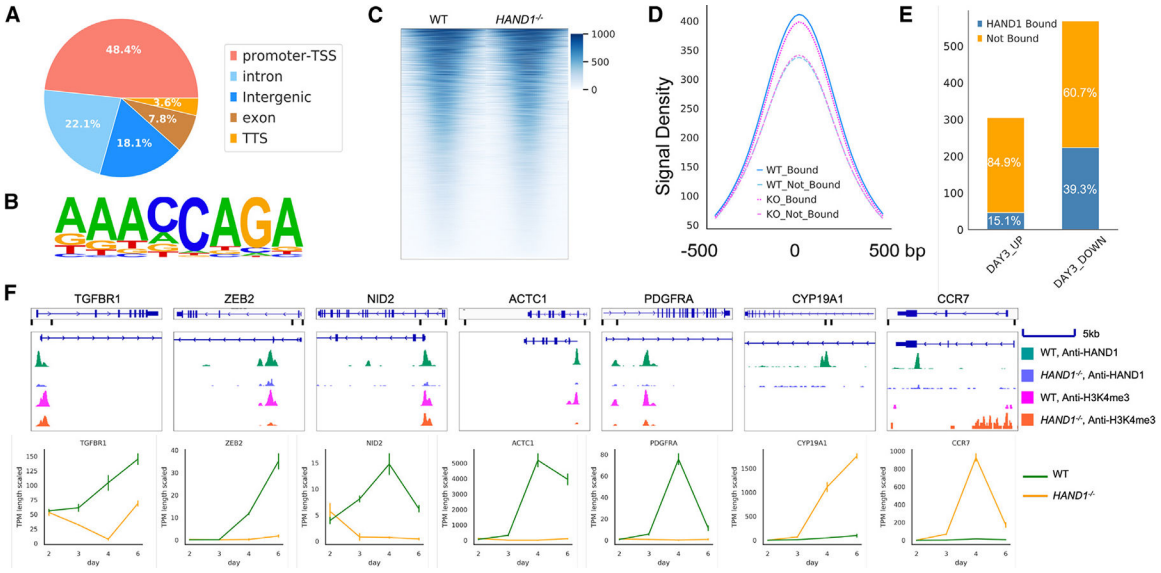


Figure 5. HAND1 directly regulates lineage-specific genes

(A) Distribution of HAND1-bound locations in iPSC under ExMC induction media on day 4.

(B) Determination of the enriched motif in HAND1-bound regions.

(C) Heatmap of H3K4me3 signaling in WT and *HAND1*$^{-/-}$ iPSC under ExMC induction media on day 4.

(D) H3K4me3 signal at regions co-bound and not co-bound by HAND1 in both WT and *HAND1*$^{-/-}$ cell lines.

(E) Overlap of earliest DEGs (*HAND1*$^{-/-}$ vs. WT on day3) and direct-targeting genes of HAND1 identified in ChIP-seq.

(F) Representative direct targets of HAND1. Top: HAND1 and H3K4me3 binding peaks viewed in IGV. Bottom: gene-expression level (TPM scaled by gene length) in both WT and *HAND1*$^{-/-}$ cells quantified by time-course RNA-seq. Error bar denotes 95% confidence interval (n = 3).

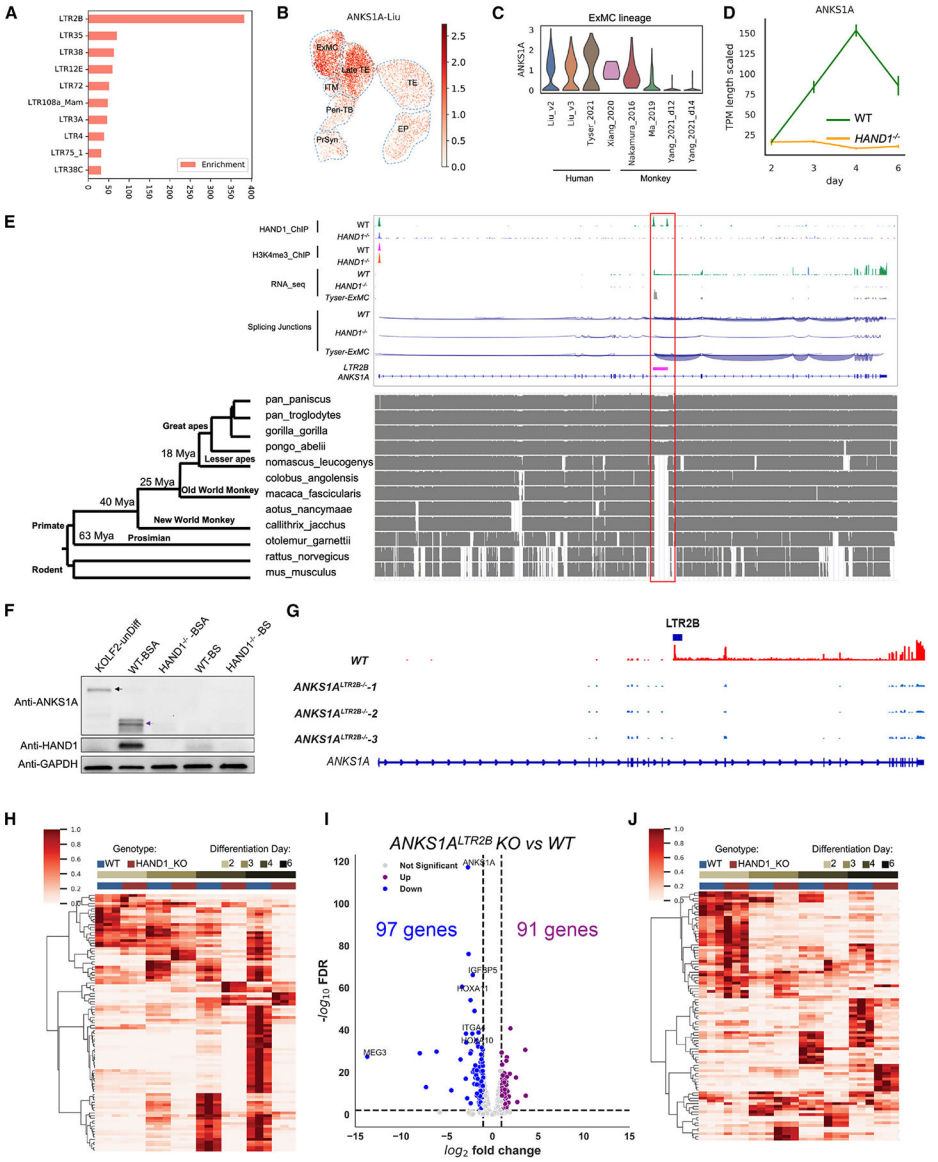


Figure 6. HAND1 binds to young transposable elements

(A) Over-representative LTR sequences found in HAND1 ChIP-seq peaks.

(B–D) Expression of *ANKS1A* in scRNA-seq data of iPSC differentiation (B), ExMC lineage from human and monkey studies (C), and time-course RNA-seq of WT and *HAND1*$^{-/-}$ iPSC-derived ExMCs (D). Error bar denotes 95% confidence interval (n = 3).

(E) Top: HAND1/H3K4me3 binding peaks, RNA-seq coverage and splicing junctions in *ANKS1A* genomic region. Bottom: comparative genomic view of the *ANKS1A* locus with rodent and primate sequences. Red rectangle: LTR2B insertion site.

(F) Protein detection of ANKS1A and HAND1 in iPSC-derived ExMC by western blot. Undifferentiated KOLF2 cell was used as a negative control (KOLF2-unDiff). Black arrow: full-length ANKS1A. Purple arrow: ExMC-specific short isoform of ANKS1A.

(G) RNA-seq coverage at the ANKS1A locus in WT and three homozygous clones of *ANKS1A*$^{LTR2B-/-}$iPSC-derived ExMC on day 6.

(H–J) DEGs of iPSC-derived ExMC from WT and *ANKS1A*$^{LTR2B-/-}$ lines. (I) Volcano plot of all significant DEGs identified by DESeq2 ($p_{adj} < 0.01$, $|\log_2 FC| > 1$, maximum scaled TPM > 10). Significant downregulated genes in *ANKS1A*$^{LTR2B-/-}$ line are indicated in solid blue, while upregulated genes are indicated by solid purple. Hierarchical clustering of the expression of these same 97 downregulated (H) and 91 upregulated (J) genes during ExMC differentiation time course of WT and *HAND1*$^{-/-}$ iPSCs.

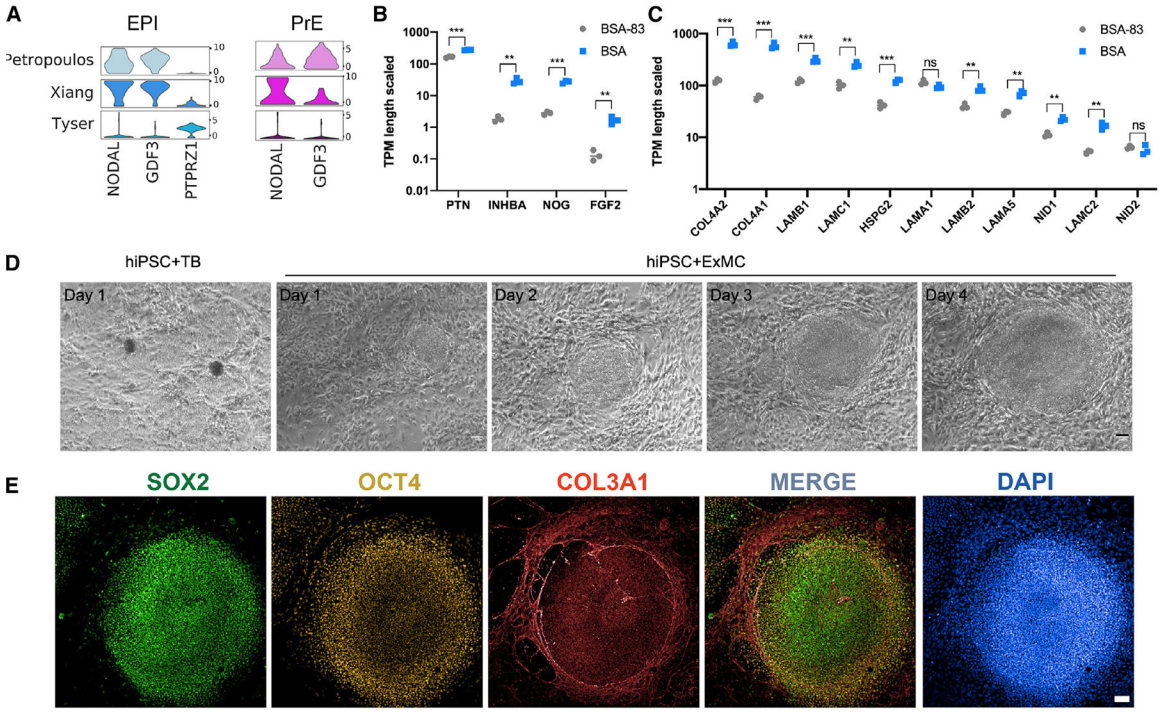


Figure 7. ExMC supports PSC maintenance

(A) Expression of different ligands in epiblast (EPI) and primitive endoderm (PrE) from both *in vitro* and *in vivo* human early embryo single-cell datasets.

(B and C) Expression of growth factors and components of extracellular matrix in iPSC-derived trophoblast (BSA-83) and ExMC (BSA) as detected from RNA-seq data. Each dot represents one observation with $n = 3$ per condition. A non-paired two-sided t test was used to determine significance. ns, not significant; $^{**}p < 0.01$, $^{***}p < 0.001$.

(D) Representative cell images from co-culture of undifferentiated iPSCs with iPSC-derived trophoblast (TB) or ExMC, as indicated. Scale bar, 100 μm.

(E) Expression of both pluripotency (SOX2 and OCT4) and ExMC (COL3A1) markers on day 4 of iPSC/ExMC co-cultures characterized by immunostaining. Scale bar, 100 μm.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-ANKS1A	BETHYL	Cat #: A303-050A; RRID: AB_10893960
anti-HAND1	R&D Systems	Cat #: AF3168; RRID: AB_2115853
anti-GAPDH	Cell Signaling Technology	Cat #: 2118S; RRID: AB_561053
HRP-conjugated Affinipure Donkey Anti-Goat IgG(H + L)	Proteintech	Cat #: SA00001-3; RRID: AB_2890882
HRP-conjugated Affinipure Donkey Anti-Rabbit IgG(H + L)	Proteintech	Cat #: SA00001-9; RRID: AB_2890888
anti-H3K4me3	Sigma-Aldrich	Cat #: CS200580
anti-IgG	Sigma-Aldrich	Cat #: CS200581
anti-PRR9	Abcam	Cat #: ab121953; RRID: AB_11130995
anti-HLA-G	Abcam	Cat #: ab52455; RRID: AB_880552
anti-SNAI2	Cell Signaling Technology	Cat #: 9585S; RRID: AB_2239535
anti-SOX2	Santa Cruz Biotechnology	Cat #: sc17320; RRID: AB2286684
anti-COL3A1	Novus Biologicals	Cat #: NB600-594; RRID: AB_10001330
anti-OCT3/4	BD Biosciences	Cat #: 560306; RRID: AB_1645312
donkey anti-mouse AF488	Invitrogen	Cat #: A-21202; RRID: AB_141607
donkey anti-rabbit AF568	Invitrogen	Cat #: A10042; RRID: AB_2534017
donkey anti-goat AF647	Invitrogen	Cat #: A32849; RRID: 2762840
donkey anti-goat AF488	Invitrogen	Cat #: A-11055; RRID: AB_2534102
donkey anti-rabbit AF647	Invitrogen	Cat #: A-31573; RRID: AB_2536183
donkey anti-mouse AF568	Invitrogen	Cat #: A10037; RRID: 11180865
Chemicals, peptides, and recombinant proteins		
StemFlex media	Gibco™	Cat #: A3349401
SynthemaxII	Corning	Cat #: 3535
RevitaCell™	Gibco™	Cat #: A2644501
ReLeSR™	STEMCELL	Cat #: 5872
Accutase	ThermoFisher	Cat #: 00-4555-56
Y-27632	TOCRIS	Cat #: 1254
KnockOut Serum	ThermoFisher	Cat #: 10828028
L-Glutamine	Gibco™	Cat #: 25030081
MEM non-essential amnio acid	Gibco™	Cat #: 11140050
β-Mercaptoethanol	Millipore Sigma	Cat #: M3148
BMP4	R&D Systems	Cat #: 314-BP- 050/CF
SU5402	Millipore Sigma	Cat #: 57263
Activin A	PeproTech	Cat #: 120-14E
A-83	TOCRIS	Cat #: 2939
GlutaMax	Gibco™	Cat #: 35050061
DMEM/F12	Gibco™	Cat #: 11330032
DMEM/F12	Gibco™	Cat #: 11320033

REAGENT or RESOURCE	SOURCE	IDENTIFIER
DPBS	ThermoFisher	Cat #: 14190250
4x Laemmli sample buffer	BIO-RAD	Cat #: 1610747
10xTBS	BIO-RAD	Cat #: 1706435
TGX™ Precast Protein Gels	BIO-RAD	Cat #: 4561093DC
Clarity Max Western ECL Substrate	BIO-RAD	Cat #: 1705062
10x Tris/Glycine/SDS buffer	BIO-RAD	Cat #: 1610732
10x Tris/Glycine Buffer	BIO-RAD	Cat #: 1610734
PVDF membranes	BIO-RAD	Cat #: 1620177
Tween® 20 Non-Ionic	Fisher Scientific	Cat #: BP337-100
Methanol	Millipore Sigma	Cat #: 179957-1L
SsoAdvanced Universal SYBR	BIO-RAD	Cat #: 1725272
DMSO	Millipore Sigma	Cat #: D1435
Thiazovivin	Selleck	Cat #: S1459
PD0325901	Selleck	Cat #: S1036
XAV939	Selleck	Cat #: S1180
WNT-C59	Selleck	Cat #: S7037
NRG-1	Cell Signaling Technology	Cat #: 5218SC
Forskolin	Selleck	Cat #: S2449
cAMP	Sigma	Cat #: D-0260
TrypLE	Gibco™	Cat #: 12604021
SB431542	TORCRIS	Cat #: 1614
ITS-X	Gibco™	Cat #: 51500056
L-Ascorbic acid	TORCRIS	Cat #: 4055
Recombinant Human EGF Protein	Bio-Techne	Cat #: 236-EG
CHIR 99021	TORCRIS	Cat #: 4423
Valproic acid	TORCRIS	Cat #: 2815
Critical commercial assays		
EZ-Magna ChIP™ HiSens Chromatin Immunoprecipitation Kit	Millipore Sigma	Cat #: 17-10461
Pierce™ BCA Protein Assay Kit	ThermoFisher	Cat #: 23227
High-Capacity RNA-to-cDNA™ Kit	ThermoFisher	Cat #: 4387406
NucleoSpin® RNA Plus	Takara	Cat #: 740984.250
Deposited data		
Raw single cell RNA-seq data	This paper	GEO: GSE220844
Raw bulk RNA-seq data	This paper	GEO: GSE220844
Raw ChIP-seq data	This paper	GEO: GSE220844
Experimental models: Cell lines		
KOLF2.C1 human iPSC line	JAX Cellular Engineering Service	Cellosaurus: HPSI0114i-kolf_2-C1
WIBJ2 human iPSC line	JAX Cellular Engineering Service	hPSC^{reg}: WTSIi046-A
WTC11 human iPSC line	JAX Cellular Engineering Service	hPSC^{reg}: UCSFi001-A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
KO of HAND1 CDS in KOLF2.C1	JAX Cellular Engineering Service	KOLF2.C1-HAND1$^{-/-}$
KO of LTR2B within ANKS1A in KOLF2.C1	JAX Cellular Engineering Service	KOLF2.C1-ANKS1A$^{LTR2B-/-}$
Software and algorithms		
CellRanger 3.1.0	10X Genomics	https://www.10xgenomics.com/support
Scanpy81 1.7.0	Wolf et al.[84]	https://scanpy.readthedocs.io/en/stable/tutorials.html
DESeq2 1.28.0	Love et al.[85]	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
rMATS 4.0.1	Shen et al.[86]	https://rnaseq-mats.sourceforge.net/
Harmony 1.0	Korsunsky et al.[87]	https://hbctraining.github.io/scRNA-seq_online/lessons/06a_integration_harmony.html
bcl2fastq 2.20.0.422	Illumina	https://support.illumina.com/downloads/bcl2fastq-conversion-software-v2-20.html
FastQC 0.11.9	Simon Andrews	https://github.com/s-andrews/FastQC
nf-core/rnaseq 3.0	AWS	https://nf-co.re/rnaseq
nf-co.re/chipseq 1.2.2	AWS	https://nf-co.re/chipseq
IPA 52912811	QIAGEN	https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/analysis-and-visualization/qiagen-ipa/
IGV 2.12.3	Broad Institute	https://igv.org/
scVelo 0.2.3	Bergen et al.[88]	https://scvelo.readthedocs.io/en/stable/
CellChat 1.6.0	Jin et al.[89]	https://github.com/sqjin/CellChat