





ARTICLE OPEN



Dissecting peri-implantation development using cultured human embryos and embryo-like assembloids

Zongyong Ai^{1,2,3,9} , Baohua Niu^{1,2,9}, Yu Yin^{1,2,3,9}, Lifeng Xiang^{4,9}, Gaohui Shi^{1,2,9}, Kui Duan^{1,2,3}, Sile Wang^{1,2,3}, Yingjie Hu^{1,2}, Chi Zhang^{1,2}, Chengting Zhang^{1,2,3}, Lujuan Rong^{1,2}, Ruize Kong⁴, Tingwei Chen^{1,2}, Yixin Guo⁵, Wanlu Liu⁵, Nan Li^{1,2}, Shumei Zhao^{1,2,3}, Xiaoqing Zhu^{1,2,3}, Xuancheng Mai⁴, Yonggang Li⁴, Ze Wu⁴, Yi Zheng⁶, Jianping Fu^{6,7,8} , Weizhi Ji^{1,2,3}  and Tianqing Li^{1,2,3} 

© The Author(s) 2023

Studies of cultured embryos have provided insights into human peri-implantation development. However, detailed knowledge of peri-implantation lineage development as well as underlying mechanisms remains obscure. Using 3D-cultured human embryos, herein we report a complete cell atlas of the early post-implantation lineages and decipher cellular composition and gene signatures of the epiblast and hypoblast derivatives. In addition, we develop an embryo-like assembloid (E-assembloid) by assembling naive hESCs and extraembryonic cells. Using human embryos and E-assembloids, we reveal that WNT, BMP and Nodal signaling pathways synergistically, but functionally differently, orchestrate human peri-implantation lineage development. Specially, we dissect mechanisms underlying extraembryonic mesoderm and extraembryonic endoderm specifications. Finally, an improved E-assembloid is developed to recapitulate the epiblast and hypoblast development and tissue architectures in the pre-gastrulation human embryo. Our findings provide insights into human peri-implantation development, and the E-assembloid offers a useful model to disentangle cellular behaviors and signaling interactions that drive human embryogenesis.

Cell Research (2023) 33:661–678; <https://doi.org/10.1038/s41422-023-00846-8>

INTRODUCTION

In humans, the blastocyst, a hollow round structure comprising ~200 cells, implants into the uterine wall to initiate post-implantation development at embryonic day 6–7 (E6–7). In the late blastocyst, the inner cell mass (ICM) segregates into the hypoblast (HB), which gives rise to the yolk sac, and the epiblast (EPI), which differentiates into the three definitive germ layers to form the embryo proper. Despite their critical roles in human development, we still have very limited knowledge of post-implantation development of the HB and EPI lineages. Recently, we and others have successfully developed in vitro human embryo culture systems, extending the growth of human blastocysts towards the pre-gastrulation stage.^{1–4} However, the precise cell atlas, lineage specification and developmental signals during the peri-implantation human development, especially those related to the HB and EPI derivatives, remain obscure.⁵


The use of human embryos for investigating peri-implantation development is fraught with ethical concerns and hampered by technical barriers, such as genetic manipulation in specific lineages and reproducible embryo development.⁵ Stem cell-based embryo models provide a more practical alternative in place of human embryos to decode early human development. Previous studies used primed human pluripotent stem cells

(hPSCs) to model some embryonic phenotypes associated with post-implantation EPI development such as embryoid body, micropatterned colony, asymmetric human EPI models (3D EPI model and post-attached embryoid), gastruloid, sequential somite-like structures (Axiloid or Segmentoid) and amniotic sac embryoid.^{6–16} However, these embryoids could not recapitulate the specification and development of the extraembryonic mesoderm (ExM) and HB lineages, only mimicked some aspects of human postimplantation EPI development such as anteroposterior symmetry breaking, anterior-posterior axial patterning, somitogenesis, dorsal-ventral patterning and amniogenesis. Although several recent reports showed that human blastocyst-like structures (blastoids) are generated from naive hPSCs or human extended pluripotent stem cells (hEPSCs) or primed-to-naive intermediates or by reprogramming fibroblasts, these blastoids showed a poor developmental potential through implantation.^{17–23} To our knowledge, some vital 3D tissue structures in human peri-implantation embryos, such as the amniotic cavity and yolk sac, are not well recapitulated in extended cultured human blastoids.

To fill the knowledge gap of human peri-implantation development, herein we provided a complete cell atlas of early post-implantation embryonic lineages in 3D-cultured human

¹State Key Laboratory of Primate Biomedical Research; Institute of Primate Translational Medicine, Kunming University of Science and Technology, Kunming, Yunnan, China.

²Yunnan Key Laboratory of Primate Biomedical Research, Kunming, Yunnan, China. ³Yunnan Provincial Academy of Science and Technology, Kunming, Yunnan, China.

⁴Department of Reproductive Medicine, The First People's Hospital of Yunnan Province, Kunming, Yunnan, China. ⁵Zhejiang University-University of Edinburgh Institute (ZJU-UoE Institute), Zhejiang University School of Medicine, International Campus, Zhejiang University, Haining, Zhejiang, China. ⁶Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI, USA. ⁷Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI, USA. ⁸Department of Cell & Developmental Biology, University of Michigan Medical School, Ann Arbor, MI, USA. ⁹These authors contributed equally: Zongyong Ai, Baohua Niu, Yu Yin, Lifeng Xiang, Gaohui Shi. email: aizy@lpbr.cn; wji@lpbr.cn; litq@lpbr.cn

Received: 17 March 2023 Accepted: 24 June 2023
Published online: 17 July 2023

blastocysts. Furthermore, we successfully created a human embryo-like assembloid (E-assembloid) that recapitulates developmental landmarks, 3D cellular organization and lineage composition of human peri-implantation embryo, thus establishing a novel embryo model for studying human early post-implantation development. Using human embryos and E-assembloids, our study unravels the transcriptional states of HB and EPI derivatives and functional roles of WNT, BMP and Nodal signaling pathways in mediating their lineage specifications.

RESULTS

Single-cell transcriptional profiling of 3D-cultured human embryos

Although some cell types in the peri-implantation human embryo have recently been characterized,^{3,4,24} a complete cell atlas of the post-implantation HB and EPI derivatives and their corresponding gene expression profiles, especially for amniotic epithelium (AME), ExM, visceral/yolk endoderm (VE/YE) and anterior visceral endoderm (AVE), are still lacking. To address this knowledge gap, we first generated a complete cell atlas for 3D-cultured human embryos⁴ from embryonic day (E10–14) (Supplementary information, Fig. S1a) using the 10× Genomics platform. Different cell types, including trophoblast (TrB), EPI, ExM and extraembryonic endoderm (XEN, an HB derivative), emerged successively between E10 and E14 in the single-cell RNA-sequencing (scRNA-seq) data (Fig. 1a, b; Supplementary information, Fig. S1b). To focus on EPI and HB lineage development, we removed TrB cells from the scRNA-seq data. The remaining cells were subclustered into 11 different populations, including post-implantation early and late EPI (PostE-EPI and PostL-EPI), undefined cells (UC), AME1 and 2, primitive streak1 and 2 (PS1 and 2), ExM1 and 2, VE/YE and AVE (Fig. 1c, d; Supplementary information, Fig. S1c).

We next analyzed gene expression signatures of these lineages and identified molecular markers useful for discerning different cell types in the peri-implantation human development. As expected, both PostE-EPI and PostL-EPI express core pluripotency genes *POU5F1*, *NANOG* and *SOX2*. PostE-EPI expresses greater levels of naive genes *UTF1*, *FGF4* and *NODAL*, whereas PostL-EPI shows higher expression of primed genes *IGFBP2*, *USP44*, *CD24* and *DNMT3B* (Fig. 1e, f). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis revealed that PostE-EPI and PostL-EPI are enriched with genes for ‘oxidative phosphorylation’ and ‘glycolysis’, respectively (Fig. 1f), suggesting a metabolic shift during EPI development. We also examined expression of genes associated with the formative pluripotency state;^{25,26} there is greater expression of *DPPA2* and *GDF3*, but lower or similar expression of *SOX11*, *OTX2* and *ZNF* genes in PostE-EPI, compared to those in PostL-EPI (Supplementary information, Fig. S1d), suggesting that there are differences between formative hPSCs and embryonic EPI in gene expression profiles. The UC cluster weakly expresses pluripotency genes and is enriched with genes for ‘apoptosis’, suggesting their apoptotic EPI identity (Fig. 1e, f). Compared with the AME2, the AME1 shows lower expression of AME markers but higher expression of pluripotency genes; compared with the EPI, the AME1 shows lower expression of pluripotency markers but higher expression of amnion genes, indicative of a transitional AME state (Fig. 1e, f). RNA velocity map and clustering further reveal the similar identities of AME1 and AME2 (Fig. 1g, h). PS1 and PS2 share early primitive streak genes *TBXT*, *SP5*, *HOXA1*, *EOMES* and *NODAL*, with PS2 highly expressing *CDX1*, *CDX2*, *MIXL1*, *MESP1*, *HAND1* and *GATA6* (Fig. 1e, f). ExM1 and ExM2 upregulate epithelial to mesenchymal transition genes *VIM* and *SNAI2*; in addition to *KDR*⁴ and *COL6A1*,²⁷ we identified *FLT1*, *MEIS2*, *COL3A1*, *NID2* and *FOXF1* as potential markers of ExM (Fig. 1e; Supplementary information, Fig. S1e). Unexpectedly, *POSTN*, a specific AVE marker in marmoset,²⁸ is exclusively expressed in the ExM2 (Fig. 1e; Supplementary information,

Fig. S1e). Expressions of *GATA4*, *GATA6*, *PDGFRA*, *FN1* and *COL4A1* are all upregulated in the ExM (ExM1 and ExM2) and XEN (VE/YE and AVE), in contrast to *APOA1*,²⁹ *IHH*,³⁰ *OTX2*,²⁹ *FOXA1*,²⁷ *FOXA2*, *HNFA4* and *SOX17*, which are specific to XEN (Fig. 1e; Supplementary information, Fig. S1e). *AFP*, *TTR*, *APOA4*, *APOC3* and *MTTP* show upregulated expression in VE/YE. Conversely, *LHX1*, *HHEX*, *EOMES*, *GSC*, *CER1*, *LEFTY1*, *LEFTY2* and *NOG* are mainly expressed in AVE (Fig. 1e; Supplementary information, Fig. S1e), suggesting that human AVE is conserved to secrete inhibitors of WNT, BMP and Nodal signaling pathways.

We further performed immunostaining for E13 human embryos and confirmed that XEN cells are positive for *GATA6*, *SOX17*, *FOXA1*, *PODXL1*, N-Cadherin and *OTX2*, and ExM cells are positive for *GATA6*, *KDR*, *SNAI1* and *COL3A1* but negative for *SOX17*, *FOXA1*, *PODXL1*, *OTX2* or pluripotency markers (Supplementary information, Fig. S1f, g), consistent with the scRNA-seq data (Fig. 1e; Supplementary information, Fig. S1e). We also observed a few *OCT4*⁺*KDR*⁺ ExM cells without *GATA6* expression (Supplementary information, Fig. S1f, g). Based on these data, we identified *OCT4*⁺*SOX2*⁺*KDR*⁺*SOX17*⁺*FOXA1*⁺*PODXL1*⁺ cells as XEN cells and *OCT4*⁺*SOX2*⁺*SOX17*⁺*FOXA1*⁺*KDR*⁺ as ExM cells.

We also performed KEGG pathway analysis to reveal signaling pathways involved in lineage development. The AME, PS, ExM, VE/YE and AVE clusters are enriched with genes associated with BMP and WNT signalings (Fig. 1f), suggesting the important roles of BMP and WNT signaling pathways in the post-implantation EPI and HB lineage specifications. Specifically, *BMP2* and *BMP4* are highly upregulated in the XEN and ExM, respectively, whereas BMP target genes *ID1*, *ID2*, *ID3* and *ID4* are mainly expressed in the AME and PS (Fig. 1f; Supplementary information, Fig. S1e), implying potential roles of XEN and ExM in inducing AME and PS specifications, which is further confirmed by intercellular communication analysis (Fig. 1i). *WNT6* is specifically expressed in AME, while *WNT3*, *WNT5A* and *WNT5B* are mainly expressed in PS (Fig. 1f; Supplementary information, Fig. S1e). Interaction analysis also showed that AME and PS are the major senders of WNT signaling pathway (Supplementary information, Fig. S1h). Interestingly, WNT inhibitors *DKK1*, *SFRP1* and *SFRP2* are mainly upregulated in ExM, AVE and EPI, respectively (Supplementary information, Fig. S1e); however, *DKK1* and *SFRP1* are mainly expressed in monkey AVE and EPI, respectively,^{28,31} suggesting different regulatory patterns of WNT signaling pathway in humans and monkeys.

Derivation of naive hESCs under normoxia condition

To model post-implantation HB and EPI lineage development, we sought to establish an embryoid using naive hPSCs, which are the in vitro counterparts of the pre-implantation EPI (Pre-EPI). Naive hPSCs have been derived using t2iLGö, 5i/LAF or HENSM medium conditions under 5% O₂.^{32–34} Such a low oxygen level, however, causes pluripotency loss and cell death in human embryos cultured through the implantation stage.¹ We thus first sought to derive naive hPSCs in normoxia (21% O₂). In the previous works, we established the AIC medium, which is composed of modified N2B27 medium supplemented with AIC (Activin-A, IWP2 and CHIR99021), and supports efficient derivation and long-term expansion of primed hPSCs through single-cell passage.³⁵ Naive hPSCs grown in the t2iLGö proliferate more rapidly and are karyotypically more stable than those cultured in 5i/LAF.^{36–38} We therefore singled out the key components of t2iLGö for the maintenance of naive pluripotency, including the MEK inhibitor PD0325901, the protein kinase C inhibitor Gö6983 and human LIF,³⁹ which were incorporated into the AIC medium to establish a new culture system, termed AIC-N medium (Fig. 2a). AIC-N allows isolation and expansion of naive hESCs from blastocysts or their conversion from conventional primed hESCs in normoxia (Fig. 2b). AIC-N hESCs could be expanded for at least 50 passages in AIC-N and retained pluripotency properties similar to naive hPSCs

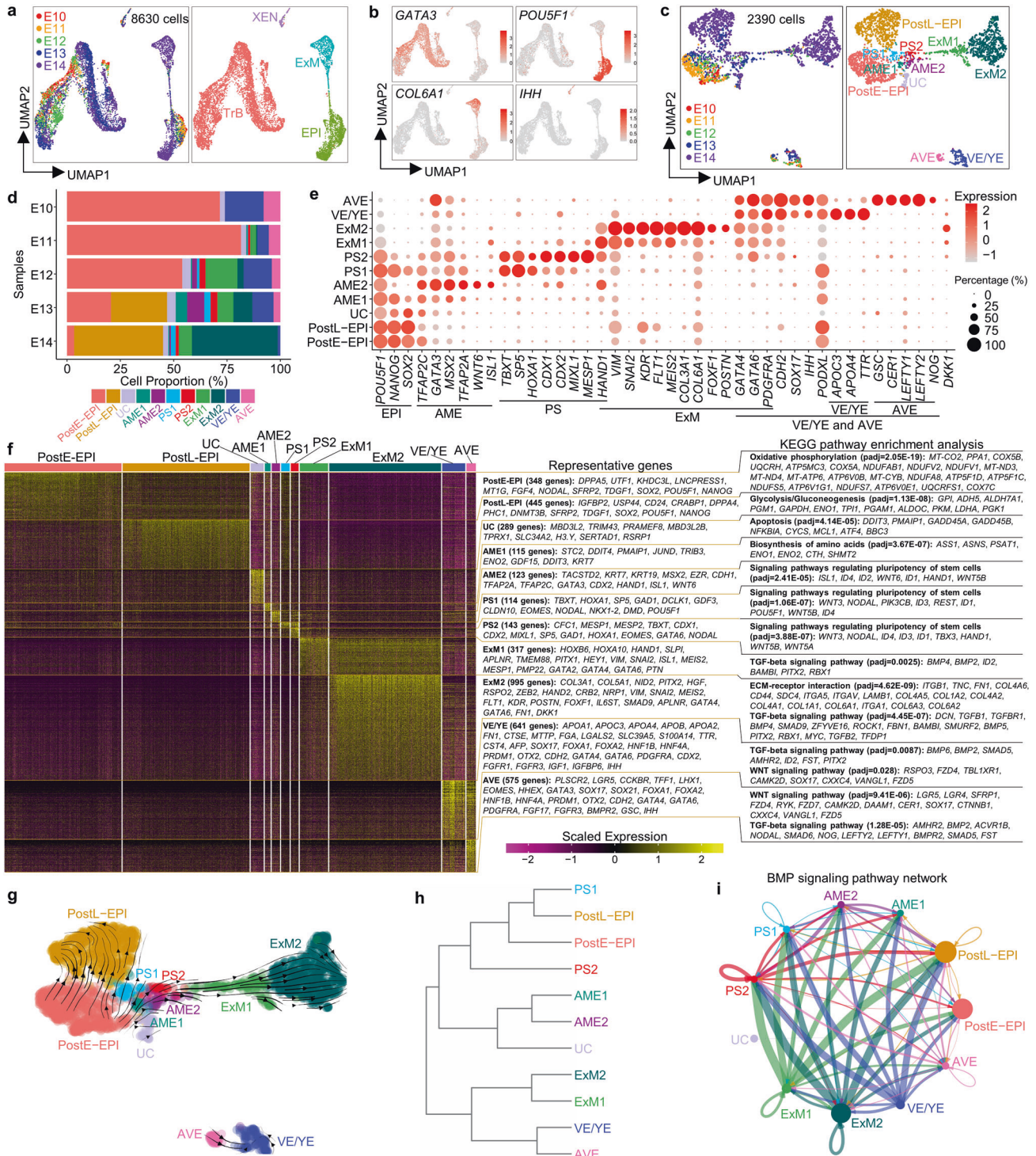
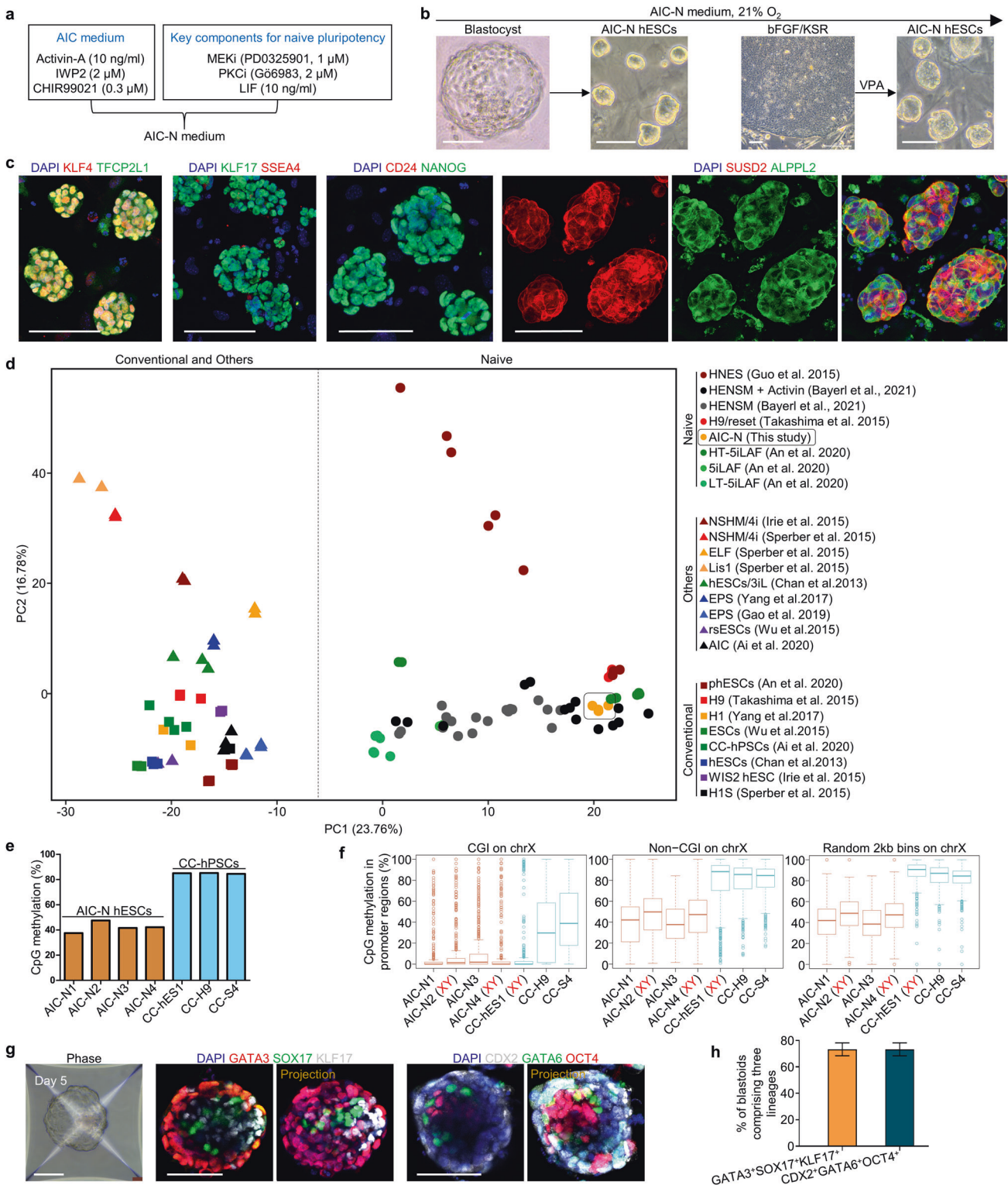


Fig. 1 Single-cell transcriptome analysis of 3D-cultured human post-implantation embryos. scRNA-seq data are from 10x Genomics platform. **a** Uniform Manifold Approximation and Projection (UMAP) visualization of single-cell transcriptome datasets from 3D-cultured human post-implantation embryos at different embryonic days. **b** UMAP plots of genes that were differentially expressed in four different populations. **c** UMAP visualization of all cell types after excluding trophoblast subpopulation. **d** Proportions of the indicated subtypes of cells in extended cultured human embryos at different developmental time points according to the results of scRNA-seq data. **e** Dot plots of candidate genes specific for cell subtypes. **f** Heatmap of differentially expressed genes (DEGs) among different cell subtypes displayed in **c**. Representative genes (left) and KEGG pathway enrichment analysis (right) are shown. Padj of Wilcoxon's rank-sum test < 0.05, $\log_2(\text{FC}) > 0.25$ and expressed in > 25% of cells of the given cluster. **g** RNA velocity vectors projected onto the UMAP-based embeddings of the single-cell transcriptome dataset from human embryos shown in **c**. **h** The hierarchical clustering tree of different cell subpopulations. **i** The inferred BMP signaling pathway network. Circle sizes are proportional to the number of cells in each subpopulation and line weight represents the communication probability. See also Supplementary information, Fig. S1.



established under 5% O₂, in terms of their expression of pluripotency markers, transcriptome, methylation and potential to differentiate into embryonic and extraembryonic lineages (teratomas, blastoids and trophoblast lineages) (Fig. 2c–h; Supplementary information, Fig. S2a–h).^{32–34,40–42}

Generation and development of human E-assembloids

We next attempted to aggregate AIC-N hESCs with either human trophoblast stem cells (hTSCs)⁴³ or naive hESC-derived

trophoblast (TE)-like cells (nTEs)^{41,42} in a basal medium diluted with half hTSC medium (termed M1) supplemented with 1% Matrigel, for the development of an integrated embryo model. However, poor (~3%) integrations of AIC-N hESCs with either hTSCs or nTEs were observed (Fig. 3a–c; Supplementary information, Fig. S3a–d). We then considered inducing trophoblast differentiation from primed AIC-hESCs³⁵ by sequential treatment of the cells with BMP4/SB431542 and the hTSC medium (Supplementary information, Fig. S3e).⁴³ The resulting cells, herein

Fig. 2 Derivation of naive hESCs under normoxia (21% O₂). **a** Schematic diagram of the AIC-N medium by incorporating the key components for the maintenance of naive pluripotency³⁹ into the AIC medium.³⁵ **b** Representative phase contrast images showing the generation of naive hESCs (AIC-N hESCs) from blastocysts and by conversion of primed hPSCs under normoxia. Four AIC-N hESC lines were established from seven blastocysts. The conversion of primed hPSCs required addition of valproic acid (VPA) for 4–6 d. **c** Immunostaining of naive, primed and general pluripotency markers for AIC-N hESCs. **d** Principle component analysis (PCA) of the gene expression profiles of hPSCs grown in various conditions. **e** Whole-genome CpG methylation levels of four AIC-N hESC lines and three primed hPSC (CC-hPSC) lines based on bisulfite sequencing (BS-seq) analysis. **f** CpG methylation levels at X-linked promoter CpG islands (CGIs) (left), non-CGI promoter regions (middle) and random 2 kb bins (right) in AIC-N hESCs and CC-hPSCs. The random 2 kb bins do not overlap any CGIs or non-CG promoters. Promoters are defined as $+/-1$ kb regions around transcription start sites. **g** Representative phase contrast and immunostaining images showing the generation of blastoids from AIC-N hESCs. **h** Quantification of the percentage of blastoids comprising three lineages (TE-, HB- and EPI-like cells), three independent experiments; more than 100 blastoids were quantified for each experiment. Data are presented as mean \pm SD. Scale bars, 100 μ m. See also Supplementary information, Fig. S2.

termed BMP4-induced cells (BICs), could be expanded over 10 passages and gradually acquire a hTSC-like identity over culture (Supplementary information, Fig. S3f, g). Gene expression patterns suggest that BICs undergo a fate transition to post-implantation trophoblast-like state over culture after exhibiting a transient amnion-like phenotype (Supplementary information, Fig. S3h–k). Interestingly, when assembling AIC-N hESCs and BICs different days after induction in M1 supplemented with 1% Matrigel, the wrapping capacity of BICs for AIC-N hESCs decreased with their extended culture, with D2 BICs showing the highest envelopment efficiency (>90%) concomitant with efficient cell survival (Fig. 3a–c; Supplementary information, Fig. S3l).

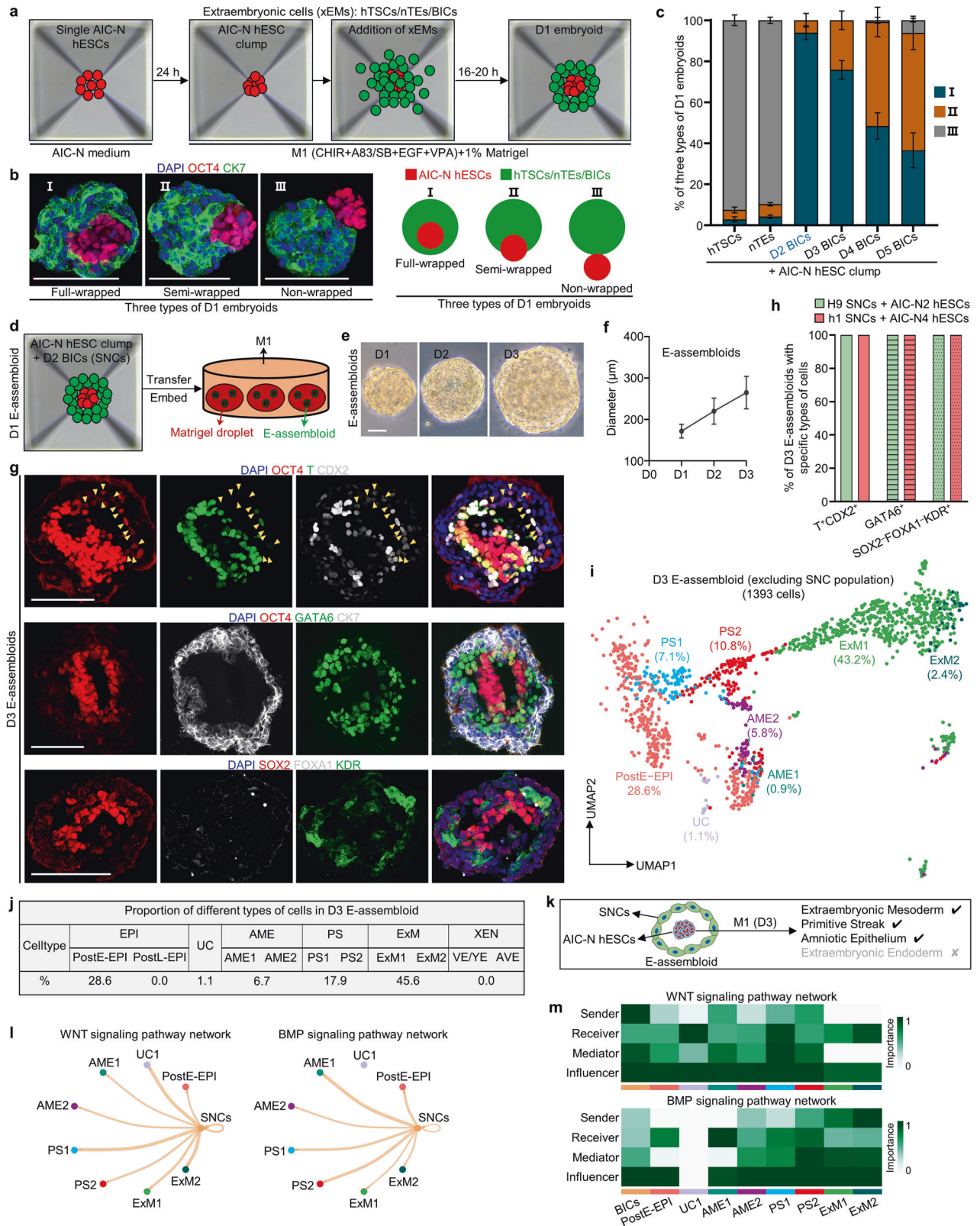
WNT and BMP signaling pathways play key roles in early human embryonic development.^{9,14,31,44–46} To investigate whether the extraembryonic tissues regulate the development of embryonic compartment by secreting WNT and BMP signals, we analyzed the WNT and BMP signaling interactions between extraembryonic and embryonic compartments in 3D-cultured human embryos before E10. As expected, the peripheral TrB and XEN send WNT and BMP signals to the internal EPI, respectively (Supplementary information, Fig. S4a, b). Similarly, D2 BICs, the intermediates that express amnion markers transiently during trophoblast induction, show upregulated expression of WNT and BMP ligands (Supplementary information, Fig. S4c), suggesting that they may have a regulatory function of extraembryonic tissue for embryonic compartment. Together, even though the counterpart of D2 BICs in the natural human embryo is unclear, they express key signaling ligand genes and provide an extraembryonic-like nest for AIC-N hESC clump, similar to the strategy reported in two recent studies using cells from hESCs induced by BMP4 to instruct EPI development into early gastrulation cell types¹⁰ or experimentally engineered morphogen signaling center that functions as an organizer to instruct development of embryo-like entities (embryoids).⁴⁷ We thus termed D2 BICs signaling nest cells (SNCs), and the embryo-like assemblage generated through integrating SNCs and naive hESCs were termed E-assembloid.

We next examined the developmental potential of E-assembloids. Under the M1 condition, E-assembloids embedded in Matrigel droplets gradually grew in size (Fig. 3d–f). On D3, a large number of T⁺/CDX2⁺ cells (putative primitive streak-like cells, PSLCs or amniotic epithelium-like cells, AMELCs), and GATA6⁺ ExM-like cells (ExMLCs) that expressed ExM markers KDR and COL6A1, but not XEN markers FOXA1, OTX2 and EOMES, emerged in E-assembloids (Fig. 3g, h; Supplementary information, Fig. S4d, e). The E-assembloids generated from different cell lines showed similar results (Fig. 3h). To confirm cell lineage origins, E-assembloids were generated using female SNCs and male AIC-N hESCs. Single-cell gene expression analysis of Y-chromosome specific gene *RPS4Y1* and lineage markers showed the presence of ExMLCs and PSLCs/AMELCs and their origins from AIC-N hESCs (Supplementary information, Fig. S4f), which was further confirmed by immunofluorescence staining of E-assembloids constructed from SNCs and mCherry-labeled AIC-N hESCs (Supplementary information, Fig. S4g). To characterize D3 E-

assembloids, the scRNA-seq data from E10–14 human embryos were used as a reference for comparative transcriptome analysis. UMAP analyses revealed a clear distinction between SNC and TrB populations, which was further confirmed by differential expressions of trophoblast and amnion marker genes in the two populations (Supplementary information, Fig. S4h, i). Thus, SNCs in E-assembloids fail to recapitulate the development of TrB-related cell lineages. To determine how AIC-N hESC derivatives in E-assembloids correspond to their embryonic counterparts, we excluded the TrB and SNC populations from the scRNA-seq data of embryos and E-assembloids, respectively. Consistent with the immunostaining results, further integrated UMAP analyses showed that AIC-N hESC derivatives consisted of cells with molecular profiles comparable to PostE-EPI (28.6%), UC (1.1%), AME (6.7%), PS (17.9%) and ExM (45.6%), but lacked PostL-EPI-like cells and XEN-like cells (XENLCs) (Fig. 3i, j; Supplementary information, Fig. S4j). In the E-assembloids, there are no XENLCs, but a large number of ExMLCs (Fig. 3j), suggesting that these ExMLCs are progenies of EPILCs, consistent with a recent study.⁴⁶ Together, E-assembloids cultured in the M1 condition recapitulate the development of human ExM, PS and AME but not PostL-EPI and XEN (Fig. 3k).

WNT and BMP signaling pathways orchestrate post-implantation lineage development

In contrast to E-assembloids, most of AIC-N hESC clumps grown alone in the M1 condition maintained pluripotency even on D9 (Supplementary information, Fig. S4k, l), suggesting that there are inductive signals from SNCs to drive E-assembloid development. Since SNCs express WNT and BMP ligands (Supplementary information, Fig. S4c), and the specifications of ExM, PS and amnion are closely related to WNT and BMP signaling pathways.^{8–10,14,46} To investigate whether the SNCs provide WNT and BMP signals to induce embryonic and extraembryonic lineage differentiation in E-assembloids, we analyzed the intercellular communication networks in D3 E-assembloids by single-cell transcriptomes, and confirmed that BICs do produce WNT and BMP signals that are received by EPI-like cells (EPILCs) and other types of cells (Fig. 3l, m). We therefore manipulated WNT and BMP signaling pathways to assess their effects on the development of E-assembloids, AIC-N hESC clumps and human embryos. Specifically, we first inhibited WNT and BMP pathways using chemical inhibitors during extended culture of E-assembloids (Fig. 4a). Contrary to the development of E-assembloids in the M1 condition, inhibition of either WNT or BMP signaling pathway in E-assembloids largely delayed EPILC development toward embryonic and extraembryonic lineages, specifications of which were even completely blocked by dual inhibitions of WNT and BMP signaling pathways (Fig. 4a, b; Supplementary information, Fig. S5a, b). Notably, both WNT and BMP inhibitions significantly reduced ExM specification, but with differential effects on AME and PS specifications (Fig. 4a–d; Supplementary information, Fig. S5a, c, d). WNT inhibition abolished PS specification but promoted AME development. In contrast, BMP inhibition



abolished AME development and reduced PS specification (Fig. 4a–d; Supplementary information, Fig. S5a, c, d). Consistent with chemical inhibition assays, development of ExM, AME and PS was significantly inhibited in E-assemblies generated from $BMP4^{-/-}/BMP7^{-/-}$ SNCs (Fig. 4e, f; Supplementary information,

Fig. S5e, f). To further validate the functions of WNT and BMP signals, different combinations of WNT and BMP agonists and inhibitors were used to treat AIC-N hESC clumps that were not wrapped in SNCs (Fig. 4g). The results showed that ExM specification required dual activation of WNT and BMP signaling

Fig. 3 Generation and extended culture of E-assembloids. **a** Diagram for generating human embryoids by assembly of AIC-N hESCs and different types of xEMs. **b** Representative staining images (left) and schematics (right) depicting three types of D1 embryoids according to the state of AIC-hESCs wrapped by extraembryonic cells (xEMs). CK7 for xEMs, and OCT4 for AIC-N hESCs. **c** Quantification of three types of D1 embryoids derived from AIC-N hESCs and different types of xEMs. Three independent experiments; more than 100 embryoids were quantified for each experimental group. Data are presented as mean \pm SD. A large number of dead cells surround the embryoid assembled from D1 BICs (Supplementary information, Fig. S3l), and therefore, these embryoids were not used for further study. **d** Diagram of extended culture of human E-assembloids assembled from AIC-N hESCs and D2 BICs (SNCs). **e** Representative phase contrast images of E-assembloids during extended culture in the M1 condition. **f** Dynamic diameters of E-assembloids during extended culture in the M1 condition. $n = 3$ independent experiments; at least 100 E-assembloids were quantified in each experiment; data are presented as mean \pm SD. **g** Immunostaining of D3 E-assembloids with specific lineage markers. Yellow arrowheads indicate AMELCs. **h** Quantification of different types of E-assembloids indicated in **g**. Three independent experiments; at least 100 E-assembloids were quantified in each experiment; data are presented as mean \pm SD. **i** UMAP visualization of integration analysis of the remaining cell types after excluding TrB and SNC populations from human embryos and D3 E-assembloids, respectively. **j** Proportion of different subtypes of cells indicated in **i** based on analysis of scRNA-seq data. **k** Schematic diagram showing the specifications of indicated cell lineages in D3 E-assembloids. **l** The inferred WNT and BMP signaling outputs from SNCs; line weight represents the communication probability. **m** Heatmap shows the relative importance of each cell group based on the computed four network centrality measures of WNT and BMP signaling network. Scale bars, 100 μ m. See also Supplementary information, Figs. S3 and S4.

pathways, and PS and AME specifications required activation of WNT and BMP signaling pathways, respectively (Fig. 4h; Supplementary information, Fig. S5g). Finally, we used 3D-cultured human embryos⁴ to further confirm that dual activation of WNT and BMP signals is indispensable for ExM specification (Supplementary information, Fig. S5h). Together, our results demonstrate that ExM specification requires both WNT and BMP signals, whereas WNT and BMP signaling pathways are required for PS and AME development, respectively (Fig. 4i).

Signaling regulatory mechanisms controlling XEN specification

Even with modulated WNT and BMP signaling pathways and extended culture, E-assembloids and AIC-N hESC clumps did not give rise to XENLCs (Fig. 4d, f, h). Since Nodal signaling is activated in XEN (Fig. 5a, b), we asked whether the presence of Activin/Nodal inhibitors (A83-01 + SB431542, AS) in the M1 medium might have blocked XENLC specification in E-assembloids and AIC-N hESC clumps. To this end, we first induced 2D differentiation of AIC-N hESCs with or without AS (Fig. 5c), and the results showed that CHIR + BMP4 induced differentiation of AIC-N hESCs into a heterogeneous population containing both ExMLCs and XENLCs (Fig. 5d, Conditions (2) and (3)). Importantly, inhibition of endogenous Nodal signaling with AS blocked XENLC specification (Fig. 5d, Conditions (1) and (4)). Interestingly, the effect of CHIR + BMP4 treatment was time-dependent: a transient (2 days) induction of CHIR + BMP4 was beneficial for XENLC specification (Fig. 5d, e, Condition (3)), whereas their continuous (4 days) treatment promoted ExMLC differentiation at the expense of XENLCs (Fig. 5d, Condition (2)). Notably, activation of WNT or BMP signaling pathway alone was ineffective in specifications of XENLCs and ExMLCs (Fig. 5f), suggesting that their specifications require a synergy between WNT and BMP signaling pathways. Together, effective differentiation of AIC-N hESCs into XENLCs requires short-time activation of both WNT and BMP signaling pathways in a Nodal-dependent manner (Fig. 5g, h).

Establishment of an improved E-assembloid by regulating signaling pathways

Based on the principle of specifications of embryonic and extraembryonic lineages involving WNT, BMP and Nodal signaling pathways (Figs. 4i and 5g, h), we established an improved protocol for assembly and extended culture of E-assembloids (Fig. 6a). The improved protocol was based on the M1 minus Activin/Nodal inhibitors and included three steps: (1) lineage induction by endogenous signals secreted from SNCs and exogenous WNT agonist (1 μ M CHIR) from D0 to D2; (2) inhibition of excessive ExM specification by blocking endogenous BMP signals with LDN and removing exogenous WNT agonist (CHIR withdrawal) from D2 to D4; and (3) spontaneous lineage specification and self-organization

without exogenous signal interference from D4 onwards (Fig. 6a). This improved protocol allowed E-assembloids to progressively grow and develop with similar sizes to 3D-cultured human embryos (Fig. 6b; Supplementary information, Fig. S6a). Remarkably, immunostaining for lineage markers revealed dynamic development of XENLCs and ExMLCs in E-assembloids (Fig. 6c, d; Supplementary information, Fig. S6b–d). On D1, no XENLCs or ExMLCs were observed (Fig. 6d; Supplementary information, Fig. S6b). On D2, abundant GATA6⁺OCT4^{weak}NANOG^{weak} cells and small numbers of SOX17⁺BLIMP1⁺ cells were detected in > 90% of E-assembloids; few FOXA1⁺, OTX2⁺, SOX2⁻FOXA1⁻KDR⁺ and OCT4^{weak} KDR⁺ cells were also specified in approximately 10%, 20%, 10% and 10% E-assembloids, respectively; and no PODXL1⁺-GATA6⁺ XENLCs and BST2⁺ ExMLCs were observed (Fig. 6d; Supplementary information, Fig. S6b–d), implying an initial XENLC and ExMLC specifications. On D3, > 90% of E-assembloids produced abundant SOX17⁺OTX2⁺ and GATA6⁺PODXL1⁺ XENLCs and a few OCT4^{weak}/KDR⁺ ExMLCs, and approximately 80% of E-assembloids gave rise to small numbers of BST2⁺ ExMLCs. On D4, almost all E-assembloids contained abundant GATA6⁺, FOXA1⁺, SOX17⁻BLIMP1⁺, PODXL1⁺GATA6⁺ XENLCs that surrounded OCT4⁺ EPILC compartment, accompanied by specification of SOX2⁻KDR⁺ ExMLCs (Fig. 6d; Supplementary information, Fig. S6b). Together, during XENLC and ExMLC specifications in E-assembloids, GATA6 is first activated; followed by SOX17 and BLIMP1; and OTX2, FOXA1, KDR, PODXL1 and BST2 are activated last.

On D8, > 90% of E-assembloids contained two expanded cavities, the yolk sac-like and amniotic cavity-like structures, surrounded by ExMLCs (Fig. 6b–d). In > 90% of E-assembloids, notably, we observed XENLCs expressing CER1 and LEFTY1 (Fig. 6e, f), suggestive of an AVE-like identity.^{24,31} Similar results were obtained from the E-assembloids constructed from two different cell lines (Supplementary information, Fig. S6e). These results demonstrated that the XENLC deficiency in E-assembloids was solved by the improved protocol (Fig. 6a). Furthermore, AMELCs were specified at the prospective dorsal pole of the amniotic cavity-like structure, and primordial germ cell-like cells (PGCLCs) were detected within AME-like tissue or near the junction of EPILCs and XENLCs (Fig. 6g–i; Supplementary information, Fig. S6f, g). In addition, in approximately 20% of E-assembloids, some T⁺ putative PSLCs are observed (Supplementary information, Fig. S6h, i). However, in addition to being at the putative posterior junction of the yolk sac-like structure and EPILC compartment, some of these putative PSLCs, which co-expressed CDX2, also appeared on opposite side of the putative posterior domain (Supplementary information, Fig. S6f), exhibiting a disorganized localization. To answer whether the effects of BICs can be simulated by adding exogenous agonists of WNT and BMP signaling pathways, we transiently treated AIC-N hESC clumps with CHIR + BMP4 for 12 h (Supplementary information, Fig. S6j), which resulted in the

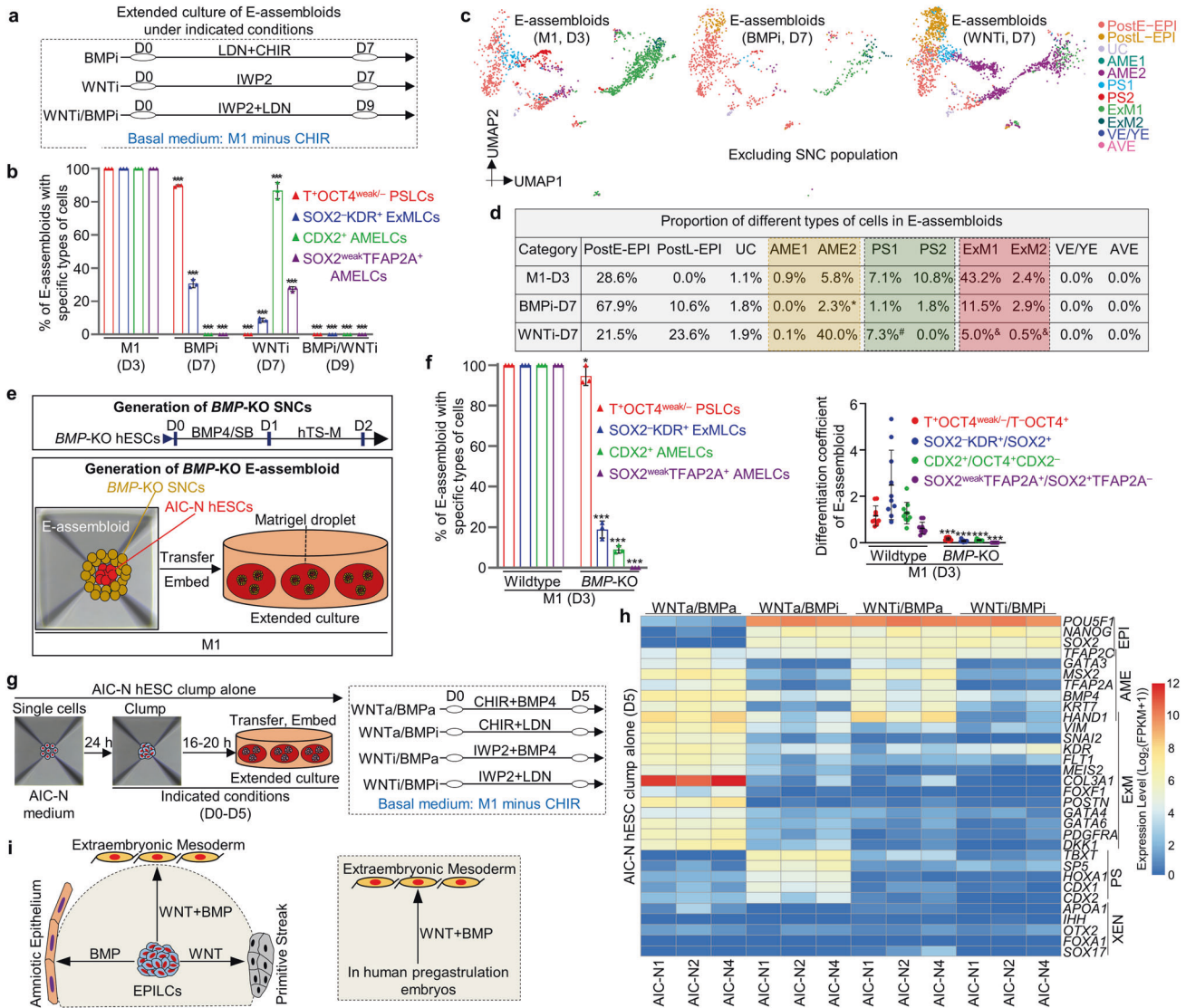


Fig. 4 WNT and BMP signaling pathways orchestrate lineage development. **a** Schematic diagram of different culture conditions for E-assembloids. BMPi and WNTi represent BMP inhibition and WNT inhibition, respectively. Compared to the M1 condition, BMP or/and WNT inhibition delayed/blocked EPILC development in E-assembloids toward embryonic and extraembryonic lineages, and we therefore prolonged culture of E-assembloids to observe the potential effects of WNT or/and BMP inhibition (7 days for individual inhibition and 9 days for dual inhibition). **b** Quantification of different types of E-assembloids cultured in the indicated conditions at the indicated time points. $n = 3$ independent experiments; at least 100 E-assembloids were quantified in each experiment; data are presented as mean \pm SD. $***P \leq 0.001$ with Student's *t*-test. **c** UMAP visualization of integration analysis of the remaining cell types after excluding TrB and SNC populations from human embryos and E-assembloids grown in the indicated conditions, respectively. **d** Proportion (left) and differentiation coefficient (right) of different types of E-assembloids (wild type and BMP-KO) grown in the M1 condition on D3. *These cells were clustered into AME2, but weakly co-expressed AME, PS and ExM marker genes, indicating an uncertain identity; #these cells were clustered into PS1 and co-expressed pluripotency and AME but not PS marker genes, implying a nascent amnion identity; ^athese cells were clustered into ExM1/2 and highly expressed AME but not ExM marker genes, implying an amnion identity. **e** Diagram of generation of BMP-KO H9 SNCs (top) and BMP-KO E-assembloids (bottom, see also Supplementary information, Fig. S5e). **f** Proportion (left) and differentiation coefficient (right) of different types of E-assembloids (wild type and BMP-KO) grown in the M1 condition on D3. The differentiation coefficient represents the ratio of differentiated cells to pluripotent cells. For the proportion, $n = 3$ independent experiments; at least 100 E-assembloids were quantified in each experiment; for the differentiation coefficient, 10 E-assembloids were randomly selected for statistics in each group. Data are presented as mean \pm SD. $*P \leq 0.05$ and $***P \leq 0.001$ with Student's *t*-test. **g** Schematic diagram of AIC-N hESC clumps cultured alone in different culture conditions for 5 days. WNTa and BMPa represent WNT activation and BMP activation, respectively. **h** Heatmap of representative marker genes of different lineages from three AIC-N hESC lines cultured in the indicated conditions. **i** Functional diagram of WNT and BMP signaling pathways on embryonic lineage development. CHIR CHIR99021, LDN LDN193189-2HCl, SB SB431542. See also Supplementary information, Fig. S5.

generation of a large number of embryonic and extraembryonic lineages on D6, including EPILCs ($OCT4^+NANOG^+SOX2^+$), PSLCs and AMELCs (T^+CDX2^+ and $TFAP2A^+$), XENLCs ($GATA6^+FOXA1^+$) and ExMLCs ($GATA6^+FOXA1^-$ and $SOX2^-KDR^+$) (Supplementary information, Fig. S6k, l), showing that exogenous activation of

WNT and BMP signaling pathways could partially simulate the induction effect of BICs. However, in these resulting AIC-N hESC spheres, the different types of cells are mostly arranged irregularly (Supplementary information, Fig. S6k). Therefore, BICs, in addition to their effect of inducing different lineage specifications, may also

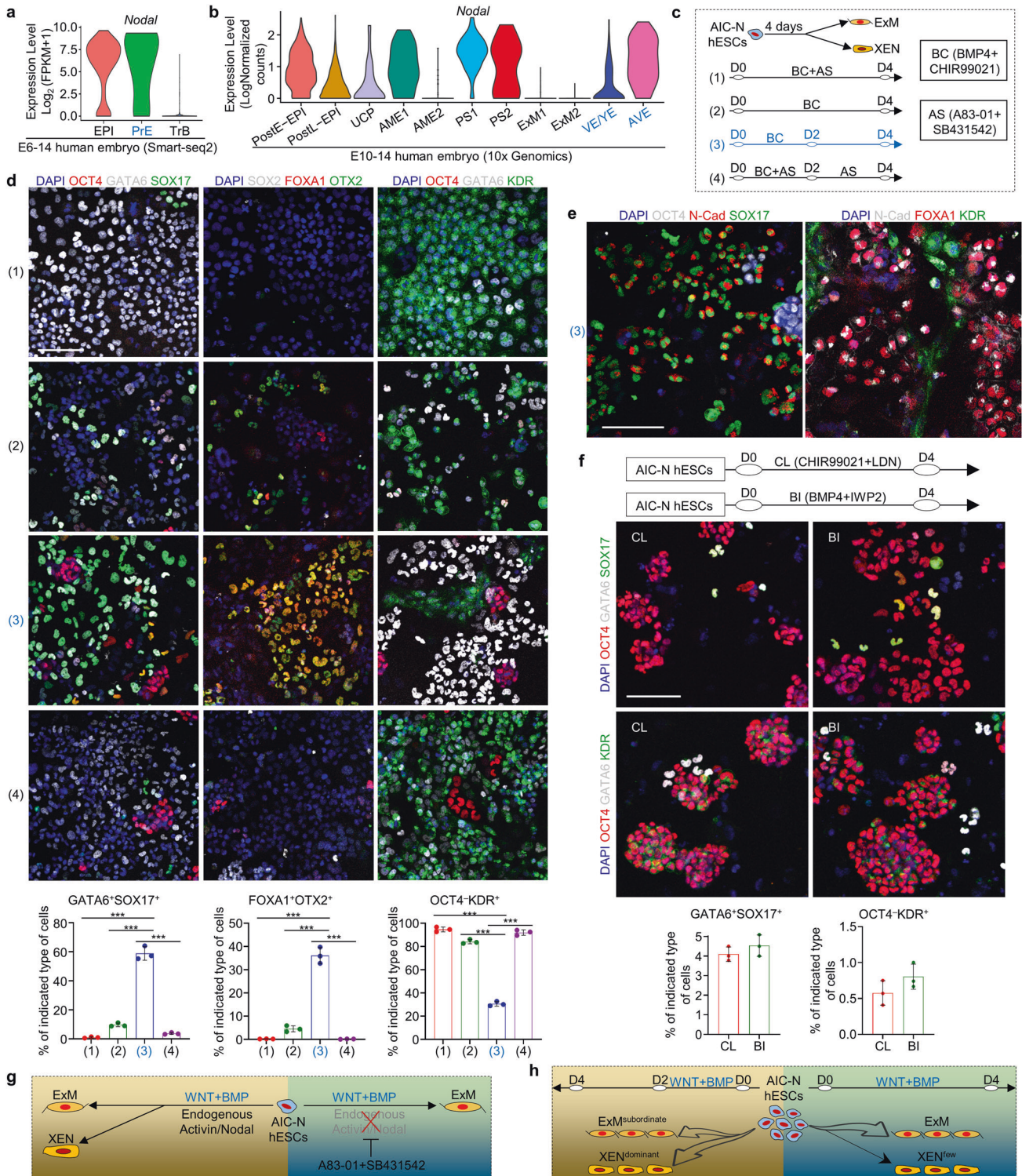


Fig. 5 Extraembryonic endoderm specification. **a, b** Violin plots of *Nodal* gene expressed in different types of cells by scRNA-seq data. **a** published smart-seq2 data.⁴ **b** 10x Genomics data. **c** Functional experiment of WNT, BMP and *Nodal* signals on XENLC and ExMLC specifications. **d** Representative staining images showing the specifications of XENLCs and ExMLCs and proportion of the indicated cell types under different culture conditions. $n = 3$ independent experiments; data are presented as mean \pm SD. *** $P \leq 0.001$ with Student's *t*-test. **e** Representative staining images showing the specifications of XENLCs and ExMLCs under Condition (3). **f** Schematic diagram of culture conditions (top), representative staining images (middle) and proportion of different cell types (bottom) under the indicated culture conditions. $n = 3$ independent experiments; data are presented as mean \pm SD. **g, h** Functional schematic diagram of WNT, BMP and *Nodal* signals on the specifications of XENLCs and ExMLCs. LCs, -like cells; scale bars, 100 μ m.

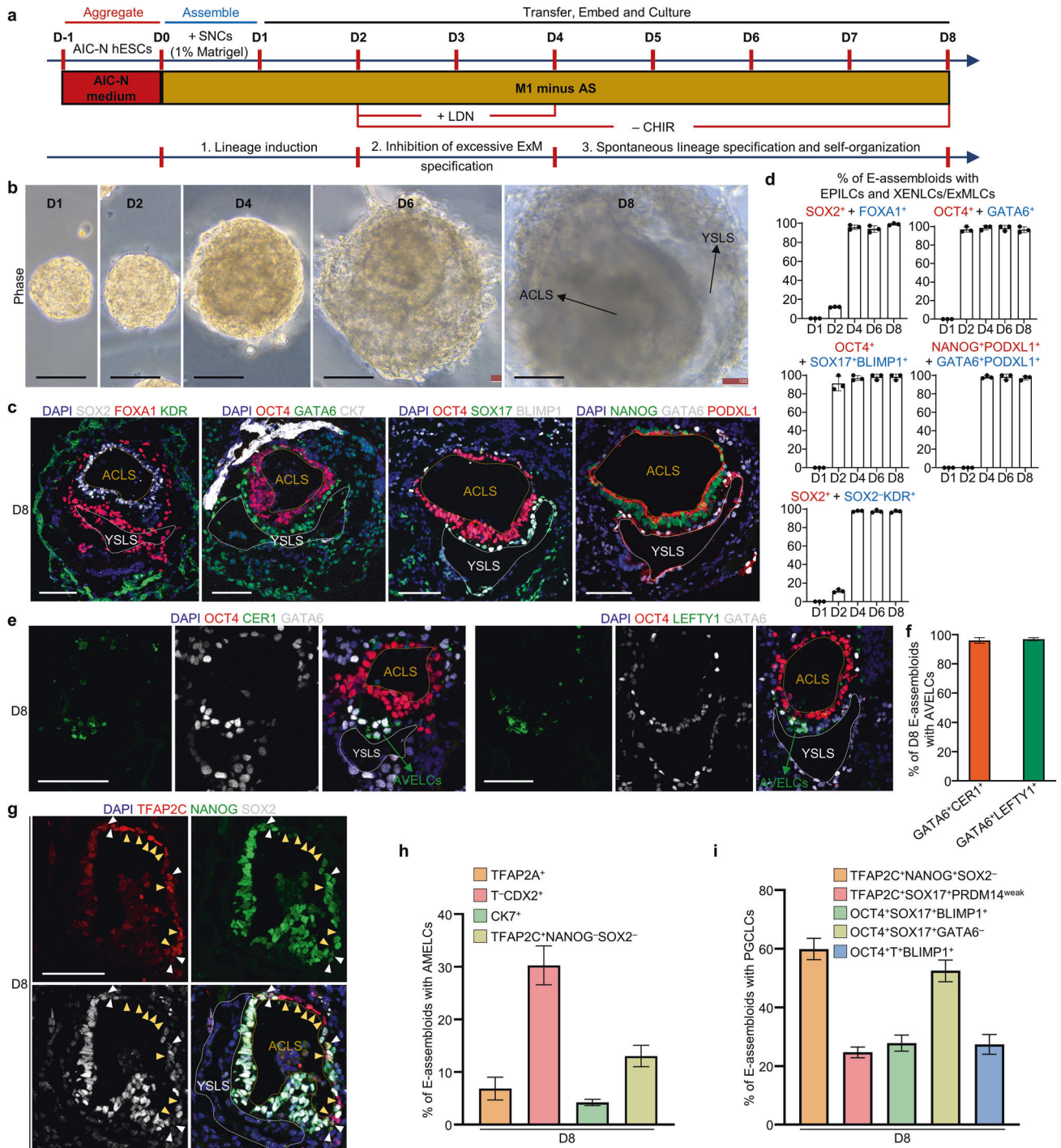


Fig. 6 E-assembloids recapitulate early post-implantation embryogenesis. **a** Schematic representation of improved protocol for assembly and extended culture of E-assembloids. AS, A83-01 + SB431542; CHIR, CHIR99021; LDN, LDN193189-2HCl; '+' and '-' represent 'add' and 'withdraw', respectively. **b** Representative phase contrast images of E-assembloids during extended culture. **c** Representative immunostaining images showing the formation of ACLS and YSLS in D8 E-assembloids. **d** Quantification of the E-assembloids with EPILCs (red) and XENLCs/ExMLCs (blue) at different time points. $n = 3$ independent experiments; more than 100 structures were quantified in each experiment; data are presented as mean \pm SD. **e** Representative immunostaining images showing the generation of AVELCs in D8 E-assembloids. **f** Quantification of the E-assembloids with AVELCs indicated in **e**. $n = 3$ independent experiments; more than 100 structures were quantified in each experiment; data are presented as mean \pm SD. **g** Representative immunostaining images showing the specifications of AMELCs (yellow arrowheads) and PGCLCs (white arrowheads) in D8 E-assembloids. **h, i** Quantification of D8 E-assembloids containing AMELCs and PGCLCs. $n = 3$ independent experiments; more than 100 structures were quantified in each experiment; data are presented as mean \pm SD. AC amniotic cavity, YS yolk sac, LS -like structure. Scale bars, 100 μ m. See also Supplementary information, Fig. S6.

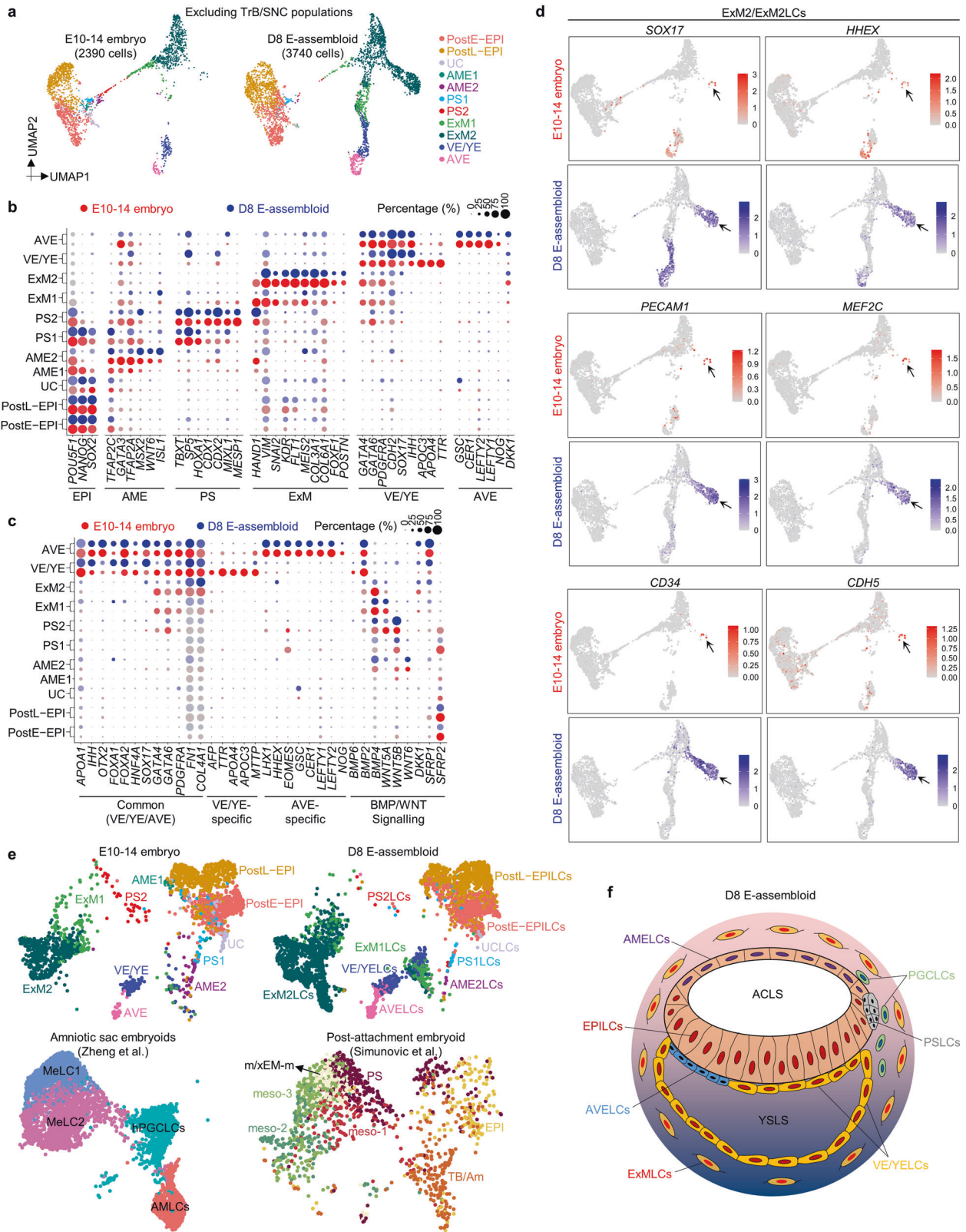


Fig. 7 Comparative single-cell transcriptomics of E-assembloids and 3D-cultured embryos. **a** UMAP visualization of integration analysis of the remaining cell types after excluding TrB and SNC populations in human embryos and D8 E-assembloids, respectively. **b, c** Dot plots of candidate genes specific for the indicated cell subtypes in human embryos and D8 E-assembloids. **d** UMAP plots of the indicated genes expressed in D8 E-assembloids and 3D-cultured embryos. Some cells in ExM2 population exhibit haemato-endothelial characteristics (arrow). **e** Comparative integration analysis of scRNA-seq datasets from 3D-cultured human embryos, E-assembloids and two other embryo models.^{10,14} **f** Schematic diagram showing 3D morphological feature of D8 E-assembloid. MeLC mesoderm-like cell, m/xEM-m mesoderm/extraembryonic mesoderm, meso mesoderm, TB/Am trophoblast/amnion. See also Supplementary information, Fig. S7.

provide mechanical effects that favor ordered organization of different cell lineages.

To further characterize D8 E-assembloids, comparative transcriptome analysis of integrated scRNA-seq data of 3D-cultured human embryos and D8 E-assembloids was conducted (Supplementary information, Fig. S7a). UMAP plots and differential expressions of trophoblast and amnion marker genes reveal an obvious difference between SNC and TrB populations (Supplementary information, Fig. S7a, b), but similar compositions of HB- and EPI-related lineages in human embryos and E-assembloids, including postE-EPI, postL-EPI, UC, AME2, PS1/2, ExM1/2, VE/YE and AVE (Fig. 7a; Supplementary information, Fig. S7c, d). AME1, however, was not detected by scRNA-seq in E-assembloids (Fig. 7a). We further split the integrated UMAP data of 3D-cultured human embryos according to different developmental time points and found that the D8 E-assembloids are similar to the E13/14 embryos in terms of cell lineage compositions (Supplementary information, Fig. S7e). Although *AFP*, *TTR*, *APOA4*, *APOC3* and *MTPP* were only expressed in a small number of VE/YE-like cells (VE/YELCs) in E-assembloids, very similar expression patterns of lineages and signaling regulatory genes were observed in other populations between human embryos and E-assembloids (Fig. 7b, c; Supplementary information, Fig. S7f). For BMP and WNT signal ligands, *BMP2* was mainly expressed in VE/YELCs and AVELCs; *BMP4* in ExMLCs, PSLCs and AMELCs; *WNT5A* in ExM1LCs and PS2LCs; *WNT5B* in PSLCs; and *WNT6* in AME2LCs (Fig. 7c). Intercellular communication analysis further showed that ExMLCs, AMELCs, XENLCs and PSLCs are the main senders of BMP signaling pathway, while AMELCs and PSLCs are also the main senders of WNT signaling pathway (Supplementary information, Fig. S7g). For antagonists of WNT, BMP and Nodal signaling pathways, *HHEX*, *CER1*, *NOG* and *LEFTY1/2* were mainly expressed in AVELCs; *SFRP1* in AVELCs; and *SFRP2* in EPILCs (Fig. 7c). These results suggest that extensive intercellular interactions between different cell lineages exist in E-assembloids, similar to the human embryos (Fig. 1i; Supplementary information, Figs. S1h and S7g). Interestingly, in the ExM2/ExM2LC populations, we observed a population of special cells, which were negative for *POSTN*, *BMP4* and *BMP2*, but highly expressed *SOX17*, *HHEX* and haemato-endothelial markers *KDR*, *S100A6*, *PECAM1*, *MEF2C*, *CD34* and *CDH5* (Fig. 7d; Supplementary information, Fig. S7h), possibly suggesting an initial hematopoiesis in both human embryos and E-assembloids.

DISCUSSION

In this study, we provided a complete cell atlas of peri-implantation embryonic lineages and identified XEN and ExM markers. Our data can serve as a resource useful for benchmarking human embryo models and embryonic/extraembryonic lineages in vitro. The bilaminar disc, amniotic cavity, yolk sac and outer ExM are vital tissue structures in human pre-gastrulation embryos, but how these tissues develop together remains unclear. Amniotic sac embryoids developed by Zheng et al. reconstruct the formation of amniotic cavity or bipolar embryonic sac, and the specifications of early gastrulating cells and PGCLCs, but cannot recapitulate the development of yolk sac and ExM (Fig. 7e).¹⁴ Although the recently reported post-attachment embryoid (also known as embryo assembloid) provide a novel strategy for studying the anteroposterior symmetry breaking and onset of gastrulation in the human EPI, this model cannot mimic the formation and specification of amniotic cavity, yolk sac and ExM (Fig. 7e).¹⁰ Our E-assembloid, a model of peri-implantation human development, can recapitulate developmental hallmarks and 3D structures of human pre-gastrulation EPI and HB derivatives, especially the bilaminar embryonic disc consisting of the amniotic cavity and the yolk sac, and the surrounding ExMs, which have never been presented simultaneously in previous embryo models (Fig. 7e, f). These advanced features place our E-assembloid as an important

model for decoding developmental mechanisms of human peri-implantation embryogenesis, such as the interactions between embryonic and extraembryonic tissues, pluripotent state transitions in the EPI,⁴⁸ XEN specification, BMP signal source that drives amnion specification, and functions of specific genes and signaling pathways. Human ExM is specified prior to gastrulation and its origin remains currently unclear.^{49,50} In 3D-cultured human embryos, RNA velocity map appeared to support EPI or early PS origin of ExM (Fig. 1g), but clustering showed that ExM is more closely related to XEN (Fig. 1h). Our results and a recently reported study have shown that EPILCs can differentiate into ExMLCs without XENLC specification (Figs. 3k, 4i and 5g),⁴⁶ however, whether ExM can originate from XEN is still unknown. With abundant XENLCs and ExMLCs (Fig. 7f), combined with live imaging and lineage tracing, our E-assembloid will provide an important platform for deciphering the developmental origin of ExM. Due to the low efficiency of trophoblast-related cells in wrapping naive hESCs, SNCs were used to construct E-assembloid. Thus, E-assembloid is not suitable for modeling peri-implantation human trophoblast development. Compared to complete human embryo models such as blastoids,^{17–23} a lack of trophoblast-related cell lineages in E-assembloids makes the model less ethically concerning. Therefore, whether the E-assembloid can continue to develop and mimic human gastrulation and organogenesis is well worth further exploring.

Using 3D-cultured human embryos and E-assembloids, as well as 2D and 3D differentiation of AIC-N hESCs, we systematically dissected the functions of WNT, BMP, and Nodal signaling on the specifications of PS, AME, ExM and XEN, and deciphered the mechanisms underlying XEN and ExM specifications (Figs. 4i and 5g, h). Our findings will be useful for improving differentiation protocols of naive hPSCs into XENLCs⁴⁴ or ExMLCs.⁴⁶ In contrast to ExMLCs, our data show that specification of XENLCs functionally requires endogenous Nodal signaling, consistent with high *Nodal* expression in human XEN (Fig. 5a, b).⁵¹ Although it has been reported that Nodal inhibition does not affect the number of GATA6⁺ cells in human and marmoset preimplantation embryos,^{45,52} another study has shown that SOX17⁺ HB cells are undetectable in Nodal inhibitor-treated human blastocysts.⁵³ These seemingly conflicting results may be caused by different choices of markers — SOX17 marks HB/XEN more specifically than GATA6.⁵⁴

Our study has several limitations. First, the mechanisms affecting the capacity of xEMs to wrap naive hESC clumps are not revealed. A previous study proposed that robust patterning during tissue morphogenesis results from cellular self-organization based on a combinatorial adhesion code instructed by morphogen-directed gradient.⁵⁵ The view is in line with the hypothesis about the cell sorting model of differential adhesion proposed by Steinberg in 1970⁵⁶ and has recently been further validated in mouse synthetic embryos.⁵⁷ Therefore, detecting differential adhesion molecules between naive hESCs and xEMs may be beneficial to answer why SNCs, rather than hTSCs and nTEs, efficiently wrap naive hESC clumps. Second, although SNCs can gradually acquire a hTSC-like identity in hTSC medium⁴³ under 2D culture condition, they cannot effectively give rise to TrB-related lineages in the extended cultured E-assembloids and therefore fail to recapitulate TrB development in human embryos. The results show that the fate of SNCs is influenced by the culture conditions and growing environment, but the underlying molecular mechanisms are unknown. Finally, in the extendedly cultured E-assembloid, it is not fully understood how the endogenous signals secreted by SNCs and exogenous signal interference truly represent signaling events during human peri-implantation development. These limitations are largely due to the lack of in vitro counterparts of human TE and HB. The establishment of faithful in vitro counterparts of TE and HB, combined with the strategy of mouse ETX embryo approach,⁵⁸

may help overcome the current challenges in developing complete human embryo model.

In conclusion, this study offers important information for advancing knowledge of human embryology and reproduction, and the E-assembloid provides a useful model for future exploration of lineage diversification, signaling interactions and tissue patterning during human peri-implantation development.

MATERIALS AND METHODS

Ethics statement

The research about human embryos and derivation of new stem cell lines from human blastocysts is a continuation of our previous work and had been approved by the Medicine Ethics Committee of The First People's Hospital of Yunnan Province (2017LS[K]NO.035 and KHLL2020-KY064).⁴ All donated embryos were supernumerary frozen embryos after IVF clinic treatment. The informed consent process for embryo donation complied with International Society for Stem Cell Research (ISSCR) Guidelines (2016/2021) and Ethical Guidelines for Human Embryonic Stem Cell Research (2003) jointly issued by the Ministry of Science and Technology and the Ministry of Health of People's Republic of China. All the donor couples had signed the informed consent for voluntary donations of supernumerary embryos for this study. No financial inducements were offered for the donations.

All experiments of in vitro culture for human E-assembloids complied with the 2016/2021 ISSCR Guidelines.^{59,60} All E-assembloids were terminated by no later than Day 8 and did not develop beyond the appearance of primitive streak or generate the initial nervous system, meeting the internationally recognized ethical limit for human embryo model.^{59,60} The research was approved by the Medical Ethics Committee of Yunnan Key Laboratory of Primate Biomedical Research (LPBR-YX003).

Embryo manipulation

To control the variability of the blastocysts, we evaluated the quality of the embryos before use. According to the Gardner's scoring system,⁶¹ thawed blastocysts were given numerical scores from 1 to 6 based on their expansion degree and hatching status. The blastocysts with expansion and hatching status above 3 and with visible inner cell mass above grade B were used in the study. Frozen-thawed human blastocysts (5–6 days post-fertilization) were gently treated by acidic Tyrode's solution (Sigma, T1788) to remove the zona pellucida.⁴ In vitro 3D culture, frozen section staining and single-cell isolation of human embryos were performed according to the previously published methods.⁴ To verify the function of WNT and BMP signaling, IWP2 (2 μ M, Selleck, S7085) and LDN193189-2HCI (LDN, 200 nM, Selleck, S7507) were used during 3D-culture of human embryos from E6–14 (Supplementary information, Fig. S5h).

In this study, 3D-cultured human embryos were selected for 10 \times sequencing according to previously established two criteria: (1) obvious expansion over culture; and (2) absence of obvious cell death or fragmented phenotype.⁴ The numbers and the size of human embryos used for 10 \times sequencing are shown in Supplementary information, Fig. S1a.

Derivation and culture of naive hESC lines in 21% O₂

Blastocysts were seeded on mouse embryonic fibroblast (Millipore, PMEF-CFL) feeder cells (hereafter referred to as feeders) inactivated by mitomycin C (Sigma, M0503). After 48 h, hESC-like outgrowths became visible in the attached embryos that were then mechanically transferred into 50% TrypLE (ThermoFisher Scientific, 12605-028) and incubated for 10–15 min at 37 °C. Alternatively, if the hESC-like outgrowth was invisible, the whole attached embryo was incubated with 50% TrypLE for 10–15 min at 37 °C. Next, the digested hESC-like outgrowth (or whole embryo) was transferred into a drop of AIC-N medium and dissociated into single cells by mildly pipetting up and down with a finely drawn Pasteur pipette. All dissociated single cells were scattered onto feeders. After 3–5 days, some obvious naive hESC colonies were observed and further dissociated into single cells by 50% TrypLE for passaging. Routinely, the AIC-N medium was changed every 2 days and the newly established AIC-N hESC lines were passaged every 3–4 days at a 1:5 to 1:10 split ratio. AIC-N medium was supplemented with 10 μ M Y27632 (Selleck, S1049) during the derivation and maintenance of AIC-N hESCs. In this study, AIC-N hESCs have been retained over 50 passages in a stable naive state. AIC-N medium was composed of modified N2B27 medium supplemented with 10 ng/mL Activin-A (Peprotech, 120-

14E), 2 μ M IWP2, 0.3 μ M CHIR99021 (Selleck, S2924), 1 μ M PD0325901 (Selleck, S1036); 2 μ M Gö6983 (TOCRIS, 2285) and 10 ng/mL recombinant human LIF (Peprotech, 300-05). 500 mL modified N2B27 medium was composed of 240 mL DMEM/F12 (ThermoFisher Scientific, 10565-018), 240 mL Neurobasal (ThermoFisher Scientific, 21103-049), 5 mL N2 supplement (ThermoFisher Scientific, 17502-048), 10 mL B27 supplement (ThermoFisher Scientific, 17504-044), 0.5% GlutaMAX (ThermoFisher Scientific, 35050-061), 1% nonessential amino acids (NEAA, ThermoFisher Scientific, 11140-050), 0.1 mM β -mercaptoethanol (Sigma, M7522), 0.38 mL 7.5% BSA (Sigma, A1933), 50 μ g/mL L-ascorbic acid 2-phosphate (Sigma, A8960), 0.5 mL Chemically Defined Lipid Concentrate (ThermoFisher Scientific, 11905-031), 12.5 μ g/mL Insulin (Roche, 11376497001), 1.25 mL Sodium Pyruvate (ThermoFisher Scientific, 11360-070), 0.02 μ g/mL Progesterone (Sigma, P8783).

Conversion of primed hPSCs from the conventional medium to the AIC-N medium

Conventional cultured-primed hPSCs (CC-hPSCs, KSR/bFGF medium) were digested into single cells with 50% TrypLE and inoculated at a density of 3×10^4 cells per 3.5-cm dish onto feeders. In the AIC-N medium, CC-hESC derivatives maintained a stable primed state after extended single-cell passaging (every 3–4 days at a 1:5 to 1:10 split ratio), with no obvious cell differentiation, death and pluripotent state transition. We recommend supplementing the AIC-N medium with 5 μ M Y27632 in the first 48 h of conversion.

To reset human CC-hPSCs, single CC-hPSCs were inoculated at a density of 8×10^4 cells per 3.5-cm dish onto feeders in AIC-N medium supplemented with 1 mM VPA (Sigma, P4543).³⁶ After 4–6 days, cell cultures were digested into single cells using 50% TrypLE and were replated onto feeders at a split ratio of 1:10 followed by withdrawal of VPA, and majority of colonies displayed domed morphology at passage 1. Compared with the flat colonies, the dome-shaped colonies were first detached from feeders by exposure to 1 mg/mL Collagenase type IV (ThermoFisher Scientific, 17104-019). Based on the different detaching speeds of colonies, we purified domed colonies in the context of Collagenase type IV during the initial (1–2) passages, the detached colonies were collected and dissociated into single cells with 50% TrypLE for sub-culturing. Thereafter, cell cultures can be split every 3–4 days at a 1:5 ratio via single-cell dissociation with 50% TrypLE, the medium was routinely changed every other day. AIC-N medium was supplemented with 10 μ M Y27632 during reset and maintenance of naive hESCs.

Establishment and culture of hTSC lines

Establishment and culture of blastocyst-derived hTSC lines was performed according to the published method.⁴³ Briefly, blastocyst was placed into Col IV (Corning, 354233)-coated 96-well plate and cultured in 150 μ L of hTSC medium. After 4–5 days of culture, the blastocyst outgrowth was dissociated with 50% TrypLE for 10–20 min at 37 °C and passaged to a new Col IV-coated 96-well plate. hTSCs were routinely passaged every 3–4 days at a 1:3 split ratio and pure hTSC lines were established within 5–10 passages. hTSC medium was composed of the following ingredients: DMEM/F12 supplemented with 0.1 mM β -mercaptoethanol, 0.2% FBS (BI, 04-002-1 A), 0.3% BSA, 1% ITS-X supplement (ThermoFisher Scientific, 51500-056), 50 μ g/mL L-ascorbic acid 2-phosphate, 50 ng/mL EGF (R&D Systems, 236-EG-01M), 2 μ M CHIR99021, 0.5 μ M A83-01, 1 μ M SB431542, 0.8 mM VPA and 5 μ M Y27632. Alternatively, hTSCs can also be cultured on feeders instead of Col IV in hTSC medium.

Differentiation of AIC-N hESCs into nTEs and hTSCs

The differentiation of nTEs was performed according to the previously published methods,^{41,42} with some minor modifications. Briefly, single AIC-N hPSCs were inoculated at a density of 1×10^5 cells per 3.5-cm dish onto Matrigel (Corning, 354277)-coated dishes in the nTE induction medium, which was changed every 2 days. After 72 h, nTEs were dissociated and collected to assemble embryoids. nTE induction medium was composed of modified N2B27 medium supplemented with 2 μ M PD0325901, 2 μ M A83-01 and 5 μ M Y27632.

The differentiation of hTSCs was performed as described,^{41,42} with some minor modifications. Briefly, single AIC-N hPSCs were inoculated at a density of 5×10^4 cells per 3.5-cm dish onto feeders, after 3 days treatment with 1 μ M PD0325901 and 1 μ M A83-01 in the modified N2B27 medium,³⁵ the initial medium was replaced with hTSC medium⁴³ and the cultures were passaged at day 5. Henceforth, cells were routinely passaged every

3–4 days at a 1:3 split ratio, and pure hTSC lines were established within 2–3 passages. All culture medium was supplemented with 5 μM Y27632 and replaced every 2 days. 500 mL modified N2B27 medium was composed of 240 mL DMEM/F12, 240 mL Neurobasal, 5 mL N2 supplement, 10 mL B27 supplement, 0.5% GlutaMAX, 1% NEAA, 0.1 mM β -mercaptoethanol, 0.38 mL 7.5% BSA, 50 $\mu\text{g}/\text{mL}$ L-ascorbic acid 2-phosphate, 0.5 mL Chemically Defined Lipid Concentrate, 12.5 $\mu\text{g}/\text{mL}$ Insulin, 1.25 mL Sodium Pyruvate and 0.02 $\mu\text{g}/\text{mL}$ Progesterone.

Differentiation of AIC-N hESC-derived hTSCs into syncytiotrophoblast cells and extravillous cytotrophoblast cells

The differentiation methods of AIC-N hESC-derived hTSCs into syncytiotrophoblast cells (2D and 3D) and extravillous cytotrophoblast cells were performed as previously reported.⁴³

AIC-hPSC culture, BIC induction and terminal differentiation

AIC-hPSCs grown on Matrigel-coated dishes/plates were cultured in the AIC medium as described previously.³⁵ Briefly, the AIC medium was changed every 2 d and AIC-hPSCs were passaged every 3–4 days at a 1:10 to 1:20 split ratio by single-cell dissociation. AIC medium was composed of modified N2B27 medium supplemented with 10 ng/mL Activin-A, 2 μM IWP2 and 0.6 μM CHIR99021.

For generation of BICs, AIC-hPSCs were digested into single cells with 50% TrypLE and inoculated at a density of 1×10^5 cells per 3.5-cm dish onto Matrigel-coated dishes in the BIC induction medium. After 24 h, BIC induction medium was replaced with hTSC medium⁴³ and BICs were passaged at Day 4 after induction. Routinely, hTSC medium was changed every 2 days, and BICs were passaged every 3–4 days at 1:5 to 1:10 and expanded over 10 passages. D2 BICs were termed SNCs in the following context. The medium was supplemented with 5 μM Y27632 during induction and maintenance of BICs. BIC induction medium is composed of DMEM/F12, 15% knockout serum replacement (KSR, ThermoFisher Scientific, A3181502), 1% NEAA, 0.1 mM β -mercaptoethanol and 12.5 $\mu\text{g}/\text{mL}$ Insulin supplemented with 10 ng/mL BMP4 and 8 μM SB431542.

The differentiation of BICs into mature trophoblast cells was performed according to the published method.⁶² Briefly, BICs were passaged using 50% TrypLE and replated on Matrigel-coated plates in feeder-conditioned medium (DMEM/F12 containing 15% KSR, 1% NEAA, 0.1 mM β -mercaptoethanol and 12.5 $\mu\text{g}/\text{mL}$ insulin was pre-conditioned on feeders for 24 h, termed FCM) supplemented with 10 ng/mL BMP4, 8 μM SB431542 and 5 μM Y27632. Routinely, the medium (termed FCM + BMP4/SB) was changed every 2 days. These cells were fixed and stained at Day 6.

Specifications of XEN and ExM

For the functional experiments of WNT, BMP and Nodal signals on specifications of XENLCs and ExMLCs, AIC-N hESCs were plated at 1500 cells/well-96 onto feeders in indicated differentiation conditions (Fig. 5c, f). The medium was changed every other day and the cells were fixed and stained at Day 4. The differentiation medium was composed of modified N2B27 medium supplemented with 50 ng/mL recombinant human FGF4 (Peprotech, 100-31-25), 1 $\mu\text{g}/\text{mL}$ heparin (Sigma, H3149) and 10 μM Y27632, 2 μM CHIR99021, 10 ng/mL BMP4 (R&D Systems, 314-BP-050), 0.5 μM A83-01 (Tocris, 2939), 1 μM SB431542 (Cellagen, C7243), 2 μM IWP2 and 200 nM LDN were added to the differentiation medium according to the indicated combinations in Fig. 5c, f.

Derivation of AAVS1-knockin hESC lines

2.5×10^5 hESCs (AIC-hESCs or AIC-N hESCs) were electroporated with 1 μg of donor plasmid AAVS1-CAG-hrGFP (Addgene, #52344) or AAVS1-Pur-CAG-mCherry (Addgene, #80946), and 1 μg of sgRNA-Cas9 expression vector by 4D-Nucleofector (Lonza) using P3 Primary Cell 4D-Nucleofector X kit (Lonza). Transfected cells were plated onto irradiated DR4 MEFs (the Cell Bank of the Chinese Academy of Sciences, Shanghai, China) in AIC or AIC-N media. 3 days later, cells were selected with puromycin (0.25 $\mu\text{g}/\text{mL}$) and expanded in the same media for 2 w, GFP/mCherry-positive and puromycin-resistant clones were picked and identified by PCR.

Generation of AIC-hESCs with simultaneous knockout of BMP4 and BMP7

Pre-validated gRNA sequences targeting the exon 3 of *BMP4* or *BMP7* gene were obtained from genome-wide databases provided by GenScript

(<https://www.genscript.com/gRNA-database.html>),⁶³ and cloned into the pGL3-U6-2sgRNA-ccbB-EF1a-Cas9ZF-IRES (Addgene, Plasmid#115519) vector. 2 μg of assembled pGL3 plasmid were electroporated into 2.5×10^5 AIC-H9-hESCs³⁵ by 4D-Nucleofector (Lonza) using P3 Primary Cell 4D-Nucleofector X kit (Lonza). Transfected cells were plated onto irradiated DR4 MEFs (the Cell Bank of the Chinese Academy of Sciences, Shanghai, China) in the AIC medium.³⁵ After 24 h, cells were selected with puromycin (0.25 $\mu\text{g}/\text{mL}$) for 3 days. About one week later, the single-cell derived clones were picked for genotype analysis around the sgRNA targeting site by genomic PCR and Sanger sequencing. The PCR Primer information and the sgRNA information are in Supplementary information, Table S1.

Teratoma formation

Teratoma assay was performed according to NIH guidelines and animal procedures and approved in advance by the Animal Care and Use Committee of Yunnan Key Laboratory of Primate Biomedical Research. Approximately 10^6 cells were resuspended in 75 μL of AIC-N medium including 20 μM Y27632 and co-injected subcutaneously with 75 μL of Matrigel (Corning, 354234) into the groin of 4-week-old NOD/SCID female SPF mice (*Prkdc^{scid}/NcrCrI*, Beijing Vital River Laboratory Animal Technology Co., Ltd). After 10–12 w, teratomas were excised, fixed with 4% paraformaldehyde (PFA), sectioned and stained with hematoxylin and eosin.

G-banding karyotype analysis

Cells were used to perform karyotype analysis 1 d before passaging. After incubated for 2–4 h with fresh medium, cells were treated by Colcemid Solution (BI, 12-004-1D) at a final concentration of 0.02 $\mu\text{g}/\text{mL}$ for 1 h. The cells were washed twice in PBS, trypsinized into single cells and centrifuged. Next, the pellet was resuspended in 5 mL of hypotonic solution (0.075 M KC1) and left at 37 °C for 15 min. 1 mL of ice-cold fixative (3:1 methanol: acetic acid) was added dropwise to the hypotonic suspension and left at room temperature for 5 min. After spinning and removing the supernatant, 5 mL of ice-cold fixative was added dropwise to the suspension, left at room temperature for 30 min and spun down. The fixing step was repeated for three times. Finally, the pellet was resuspended in a final volume of 3 mL ice-cold fixative and placed in –40 °C freezer. Subsequent G-banding karyotype analysis was performed at The First People's Hospital of Yunnan Province, Kunming, China. For each analysis, at least 20 metaphases were counted.

Blastoid formation

Blastoid formation was performed according to a previously described protocol¹⁷ with minor modifications. Briefly, AIC-N hESC colonies were detached from feeders by exposure to 1 mg/mL Collagenase type IV for 60–90 min. The detached colonies were collected, dissociated into single cells with 50% TrypLE, filtered through a 20- μm cell strainer (Miltenyi Biotec, 130-101-812), pelleted by centrifugation for 4 min at 1000 rpm, suspended in modified N2B27 medium containing 10 μM Y27632 and counted using a hemocytometer. AggreWell™ 400 24-well Plates (STEMCELL Technologies, 34415) were prepared following the manufacturer's instructions. 2 mL of cell suspension (including 8×10^4 AIC-N hESCs) per well was added into the Aggrewells. After 24 h, the modified N2B27 medium containing 10 μM Y27632 was replaced with modified PALLY medium.¹⁷ The modified PALLY medium was composed of modified N2B27 medium supplemented with 1 μM PD0325901, 1 μM A83-01, 4 μM 1-Oleoyl lysophosphatidic acid sodium (LPA; Sigma, L7260), 10 ng/mL recombinant human LIF and 10 μM Y27632. After 72 h, the modified PALLY medium was replaced with modified N2B27 medium containing 4 μM LPA and 10 μM Y27632 and maintained for another 2 d. At day 5, blastoids were collected for staining. Cultures were maintained in a humidified incubator under 37 °C, 5% CO₂ and 5% O₂ along the whole process, and the medium was refreshed every 24 h.

Generation of human embryoids

AIC-N hESC colonies were detached from feeders by exposure to 1 mg/mL Collagenase type IV for 60–90 min. The detached colonies were collected, dissociated into single cells with 50% TrypLE, filtered through a 20- μm cell strainer, pelleted by centrifugation for 4 min at 1000 rpm, suspended in AIC-N medium and counted using a hemocytometer. AggreWell™ 400 24-well plates were prepared following the manufacturer's instructions. 500 μL of cell suspension (including 3×10^4 AIC-N hESCs) per well was added into the Aggrewells. After 24 h, AIC-N hESCs formed tight and round

clumps. Different types of xEMs (including hTSCs, nTEs and BICs/SNCs) were dissociated into single cells, filtered, centrifuged, suspended in M1, and counted as described above. AIC-N medium were largely removed and 1 mL of cell suspension (including 3.5×10^5 xEMs) per well was slowly added to the Aggrewwells. The time of cell seeding is considered as the starting point (0 h). After 1 h, additional 1 mL M1 with 2% Matrigel (Corning, 354230) per well was slowly added into the Aggrewwells. Within 16–20 h, xEMs were aggregated with AIC-N hESC clumps to form embryoids. The embryoids assembled by AIC-N hESC clumps and SNCs (D2 BICs) were hereinafter collectively referred to as E-assembloids. The ingredients of M1 medium were as follows: 3:1 mixture of DMEM/F12 and Neurobasal supplemented with 0.1 mM β -mercaptoethanol, 0.5% NEAA, 0.1% FBS, 2.5% KSR, 0.25% N2 supplement, 0.5% B27 supplement, 0.15% BSA, 0.25% GlutaMAX, 1% ITS-X supplement, 50 μ g/mL L-ascorbic acid 2-phosphate, 50 nM Methylene blue (Selleck, S4535), 25 ng/mL EGF, 1 μ M CHIR99021, 0.25 μ M A83-01, 0.5 μ M SB431542, 0.4 mM VPA. Additional chemicals described in this study contain: IWP2 (2 μ M) and LDN (200 nM). All media were supplemented with 10 μ M Y27632 during generation and culture of human embryoids.

'3D embedded' extended culture of human E-assembloids

Human E-assembloids were transferred from the microwells into the precooled 1.5-mL micro-centrifuge tubes inserted in ice. After about 10 min, the supernatant was aspirated, and ice-cold 1:2 mixture of M1 and Matrigel (Corning, 354230) was added to resuspend the pellet at a final concentration of 600 E-assembloids per mL. After mixed thoroughly (To ensure the uniform distribution of E-assembloids in the suspension, each 1.5-mL micro-centrifuge tube contained no more than 300 μ L of suspension, which was thoroughly mixed while being dropped to 3.5 cm dish), 10–20 μ L/droplet E-assembloid suspension was plated into 3.5-cm dish (about 20 droplets per dish), the dish was quickly turned upside down to prevent E-assembloids from falling to the bottom of the droplet, allowed to solidify at 37 °C for 20 min and overlaid with 2 mL pre-warmed M1 per dish. Matrigel, media and tubes were kept on ice during the whole process of embedding, and all media used for embedding and culturing human E-assembloids were supplemented with 10 μ M Y27632. Routinely, the medium was changed every 2 d unless otherwise noted. For the extended culture of human E-assembloids, IWP2/LDN addition and CHIR/A83-01/SB431542 withdrawal was performed based on the M1 condition during indicated timeframe, and 2 protocols for assembly and extended culture of E-assembloids were established (Figs. 3a, d and 6a). For the functional experiments of WNT and BMP signaling pathways, IWP2/LDN addition and CHIR withdrawal were performed based on the M1 condition (Fig. 4a).

'3D embedded' culture of AIC-N hESC clumps

AIC-N hESC clumps was prepared as described in 'Generation of human embryoids'. After the formation of clumps, the AIC-N medium was largely removed and 2 mL of indicated medium per well was gently added to the Aggrewwells. Within 16–20 h, these AIC-N hESC clumps were transferred from the microwells and cultured according to the '3D embedded' culture method as described above. Briefly, these AIC-N hESC clumps were transferred into the precooled 1.5-mL micro-centrifuge tubes inserted in ice. After about 10 min, the supernatant was aspirated, and ice-cold 1:2 mixture of indicated medium and Matrigel was added to resuspend the pellet with a final concentration of 1000 clumps/mL. 10–20 μ L/droplet AIC-N hESC clump suspension was plated into 3.5-cm dish (about 20 droplets per dish), allowed to solidify at 37 °C for 20 min and overlaid with 2 mL pre-warmed medium per dish. Matrigel, media and tubes were kept on ice during the whole process of embedding, and all media used for embedding and culturing AIC-N hESC clumps was supplemented with 10 μ M Y27632. Routinely, the medium was changed every 2 d unless otherwise noted. During extended culture, AIC-N hESC clumps were cultured alone to different time points in M1 condition (Supplementary information, Fig. S4k). For the functional experiments of WNT and BMP signaling, CHIR99021 (3 μ M), BMP4 (20 ng/mL), LDN (200 nM) and IWP2 (2 μ M) were added to M1 medium minus CHIR99021 according to the indicated combinations (Fig. 4g).

Inducing '3D embedded' AIC-N hESC clumps using CHIR99021 and BMP4

AIC-N hESC clumps was prepared as described in 'Generation of human embryoids'. 24 h after seeding, these formed AIC-N hESC clumps were

transferred from the microwells and cultured for 6 d in the indicated condition (Supplementary information, Fig. S6j) according to the '3D embedded' culture method as described above. The basal medium was composed of modified N2B27 medium supplemented with 10 μ M Y27632. For the induction of WNT and BMP signaling, CHIR99021 (2 μ M) and BMP4 (1 ng/mL) were used to treat the AIC-N hESC clumps during the first 12 h of extended culture (Supplementary information, Fig. S6j).

In this study, we usually used AIC-N hESCs at passages 20–40 to construct embryoids and did not observe obvious differences caused by the number of passages. Unless otherwise specified, cell or embryoid culture experiments were performed in a humidified incubator under 21% O₂ and 5% CO₂ at 37 °C. Cell lines were routinely checked for mycoplasma contaminations using MycoAlert Mycoplasma Detection Kit (LONZA, LT07-318) every two weeks, and all cell samples used in this study have been ruled out of mycoplasma contamination.

Flow cytometry

To detect cell-surface markers using flow cytometry, AIC-N hESC colonies were detached from feeders by exposure to 1 mg/mL Collagenase type IV for 60–90 min, the detached colonies were collected and dissociated into single cells with 50% TrypLE, centrifuged, washed with ice-cold PBS containing 1% FBS (BI, 04-001-1 A). For flow cytometric analysis of Tra-1-85, the cells were incubated at 4 °C for 30 min with an anti-Tra-1-85 antibody (Abcam, ab108308, 1:100) diluted in PBS including 1% BSA and then stained with a AlexaFluor 555-conjugated donkey anti-rabbit IgG secondary antibody (ThermoFisher Scientific, A-31572, 1:1000) for 30 min. For Tra-1-60 and SUSD2, the cells were incubated at 4 °C for 30 min with a PE-conjugated anti-Tra-1-60 antibody (ThermoFisher Scientific, 12-8863-82, 1:5) or a PE-conjugated anti-SUSD2 antibody (Biolegend, 327406, 1:100) diluted in PBS including 1% BSA. Normal rabbit IgG (Abcam, ab172730, 1:100), PE-conjugated mouse IgM (ThermoFisher Scientific, MGM04, 1:10), and PE-conjugated mouse IgG1 (BD Pharmingen, 554680, 1:100) antibody was used as isotype control, respectively. Flow cytometry was performed using a FACSAria III, and the data were analyzed using FlowJo software.

Immunofluorescence staining

All adherently growing cells in the study were fixed with 4% PFA for 20 min at room temperature and washed thrice in PBS. For embryoids at D1 (aggregates in Aggrewwells), aggregates were transferred from the microwells into the precooled 1.5-mL micro-centrifuge tubes inserted in ice. After about 10 min, the supernatant was aspirated, aggregates were resuspended and fixed in 4% PFA at 4 °C for 3 h. For '3D embedded' culture, embryoids or hESC spheres were fixed in 4% PFA at 4 °C for 3 h, then transferred into the 1.5-mL micro-centrifuge tubes. All structures (embryoids or hESC spheres) were washed thrice in PBS, dehydrated overnight in PBS including 20% sucrose at 4 °C, embedded into O.C.T. (Sakura Finetek, 4583) and sectioned by a Leica frozen slicer at a thickness of 10 μ m. After permeabilization and blocking with PBS including 0.2% Triton X-100, 100 mM Glycin and 3% BSA at room temperature for 60 min, the cells or sections were incubated with primary antibodies at 4 °C overnight, washed thrice with PBS including 0.05% Tween-20, incubated with secondary antibodies for 2 h at room temperature and washed thrice with PBS including 0.05% Tween-20. DAPI (Roche Life Science, 10236276001) was used for staining the nuclei. Pictures were taken by Leica SP8 laser confocal microscope. The antibodies were listed in Supplementary information, Table S2.

Envelopment efficiency evaluation, quantifications of different types of cells or structures, cell number counts and diameter measurements

To evaluate the efficiency of AIC-N hESC clumps enveloped by xEMs (hTSCs or nTEs or BICs/SNCs), frozen sections of embryoids were stained with OCT4 and CK7 for AIC-N hESCs and xEMs, respectively. According to the expression pattern of lineage markers, these structures are divided into 3 types (Fig. 3b). The percentage of different types of structures were quantified manually using the confocal microscope (Leica SP8). Statistical analysis and plotting were performed with GraphPad Prism 9.

To quantify the proportion of different types of cells or embryoids, and number of specific cell types in embryoids, 2D cell cultures and frozen sections of different types of embryoids were stained with the indicated markers. The percentage of different types of cells or embryoids, and number of specific cell types in embryoids were then quantified manually using the confocal microscope (Leica SP8). Statistical analysis and plotting were performed with GraphPad Prism 9.

To measure the diameters of developing E-assembloids, E-assembloids at different time points were fixed in 4% PFA, and randomly photographed by phase contrast microscope after Matrigel was depolymerized. Images were processed and diameters of E-assembloids were measured with Image J software. Statistical analysis and plotting were performed with GraphPad Prism 9.

RNA-seq and data analysis

Adherently growing cells (BICs grown on Matrigel and hTSCs grown on Col IV) were collected by dissociating into single cells with 50% TrypLE. For AIC-N hESC and hTSCs grown on feeders, the colonies were detached from feeders by exposure to Collagenase type IV for 60–90 min, and the detached colonies were collected. For differentiated AIC-N hESC clumps embedded in Matrigel droplets, the cultures were subjected to Cell Recovery Solution (Corning, 354253) at 4 °C for 1 h. After Matrigel was depolymerized, the differentiated AIC-N hESC clumps were collected. All cultures were washed twice in PBS. Total RNA of AIC-N hESCs, hTSCs, BICs and differentiated AIC-N hESC clumps was isolated with the TRIzol™ Reagent (ThermoFisher Scientific, 15596018). The 2 × 150 bp paired-end libraries were sequenced with Illumina HiSeq X Ten or NovaSeq 6000 instrument. Library construction and sequencing were performed by Annonoad Gene Technology (<http://www.annonoad.com/>). The published data used in this study are from reset hPSCs,³³ 3iL hPSCs,⁶⁴ human EPS cells,^{65,66} Si/LAF hPSCs,⁶⁷ AIC-hPSCs,³⁵ NHSM4i hPSCs,^{68,69} HENSM/HENSM-ACT hPSCs,³⁴ HNES cells,⁷⁰ human rESCs⁷¹ and conventional (primed) hPSCs. Reads were mapped to human genome hg38 using HISAT2 (v2.2.1) with GENECODE v30. The counts and FPKM values for each gene were calculated with StringTie (v2.1.1). We identified the most variable genes through fitting a non-linear regression curve between average \log_2 (FPKM) and the square of coefficient of variation according to the methods described.^{41,54} PCA was performed using princomp function from the R stats package based on the covariance matrix. Heatmaps were generated using pheatmap package from the R software.

Correlation analysis among human trophoblast cells,⁴ monkey amnion cells,³⁰ AIC-hPSCs,³⁵ BICs different days after induction and hTSCs were performed by Pearson correlation. We calculated the average gene expression level of different cell types using the AverageExpression function in the Seurat. Then, specific thresholds were applied along the x-axis (average \log_2 FPKM > 5) and y-axis ($\log CV^2 > 0.2$) to identify the most variable genes. Finally, the remaining genes are used to calculate the Pearson correlation among different cell types.

BS-seq and data analysis

Library construction and sequencing were completed by Annonoad Gene Technology (<http://www.annonoad.com/>). Paired-end sequencing was performed on HiSeq X Ten platform (Illumina).

The sequencing quality of paired end 150 bp raw reads were checked by FastQC (v0.11.8) (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Reads were then aligned to human reference genome (GRCh38) using BSMAP (v2.7.4) allowing 2 mismatches (-v 2), up to one multiple hits (-w 1), map to all four possible strands (-n 1).⁷² Methylation levels over each cytosine were calculated using BSMAP (v2.7.4) methratio.py script. Methylation levels over different genomic regions were extracted using a customized Python script (<https://github.com/MoPie/20210607LTQProject/extractMeth.py>). Promoters were defined as upstream 1 kb and downstream 1 kb of transcription start sites. CGI coordinates were downloaded from UCSC genome browser (<http://hgdownload.cse.ucsc.edu/goldenpath/hg38/database/cpgIslandExt.txt.gz>). Promoter CGIs on chromosome X were calculated using Bedtools intersect function with default parameter.⁷³ Promoter non-CpG islands were defined as promoter with no CGIs. Random 2 kb bins were permuted with Bedtools shuffle function by excluding CGIs and promoter regions.⁷³ Data were then visualized in R with boxplot() function and pheatmap() function from pheatmap package (<https://cran.r-project.org/web/packages/pheatmap/>). Global comparison of methylation between samples was calculated by averaging CpG methylation levels over indicated windows using the bisulfite feature methylation pipeline in SeqMonk. Pseudocolour scatter plots of methylation levels over indicated windows were generated using R.

Single cell dissociation, RNA-seq and data processing

Single cell dissociation of 3D-cultured human embryos was optimized based on our previous work and other report.^{4,74} Briefly, embryos were washed in PBS three times, washed with TrypLE two times and incubated

in TrypLE at 37 °C for 25–30 min. The embryos were dissociated into single cells by mouth pipetting in 1% DFBS/PBS. E-assembloids grown in indicated conditions at different time points were subjected to Cell Recovery Solution at 4 °C for 1 h. After Matrigel was depolymerized, E-assembloids were transferred into TrypLE containing 1 mg/mL Dispase (Corning, 354235) and incubated for 30 min at 37 °C. E-assembloids were gently dissociated into single cells by pipetting up and down, filtered through a 20 µm cell strainer, centrifuged, suspended in PBS containing 0.04% BSA and counted using a hemocytometer. Single-cell suspension of E-assembloids and embryos were loaded into the 10 × Genomics Chromium system within 30 min after dissociation. 10 × Genomics v3 or v3.1 libraries were prepared according to the manufacturer's instructions. Libraries were sequenced with a minimum coverage of 30,000 raw reads per cell on Illumina HiSeq X Ten or NovaSeq 6000 with 150-bp paired-end sequencing, which was performed by Annonoad Gene Technology (<http://www.annonoad.com/>). Sequencing data were aligned and quantified using the Cell Ranger Pipeline v3.0.1 (10 × Genomics) against the GRCh38 reference genome. Data from 10 × Genomics for E-assembloids or 3D-cultured embryos were filtered based on number of expressed genes and expression level of mitochondrial genes (below 20%). Cell doublets were removed by DoubletFinder (v2.0.3) with assuming multiplet rate according to the loaded cells number (refer to Multiplet Rate Table provided in the 10 × Genomics User Guide).

Further analyses were performed using Seurat package (v 4.0.3).⁷⁵ The raw counts were normalized and scaled with default parameters. Top 2000 most variable genes were identified and used for dimensionality reduction with PCA followed by non-linear dimensionality reduction using UMAP. Cell types were defined based on the lineage markers and clusters identified through FindClusters function. Data was visualized with the UMAP dimensionality reduction. DEGs were identified with the FindAllMarkers() function in Seurat and filtered with P_{adj} of Wilcoxon's rank-sum test < 0.05, \log_2 (FC) > 0.25 and expressed in > 25% of cells of the given cluster.

For 10 × Genomics data from E-assembloids, *RPS4Y1* expression was used to help determine the source of the cells as the two cell lines used to construct E-assembloids are of different genders. The male AIC-N hESC derivatives were further filtered with *RPS4Y1* expression (normalized and natural-log (log1p) transformed value > 1), the remaining cells were integrated with epiblast or HB derivatives from 3D-cultured embryos. Data integration was performed with IntegrateData() function. Cells from E-assembloids were classified based on 10 × Genomics data of 3D-cultured human embryos with the TransferData() function in Seurat, and predicted cell types were displayed on UMAP of integrated Seurat object.

The intercellular communication networks of 3D-cultured human embryos and E-assembloids were analyzed following the published method⁷⁶ implemented in CellChat (<https://github.com/sqjin/CellChat>), and the netVisual_signalingRole function was used to visualize the communication pattern among different cells. Expression of different genes were displayed with VlnPlot () or DotPlot () function in Seurat.

Statistical analysis

Statistical tests were performed on GraphPad Prism 9 software and Microsoft office Excel 2019. Data were checked for normal distribution and equal variances before each parametric statistical test was performed. Where appropriate, *t*-tests were performed with Welch's correction if variance between groups was not equal. ANOVA tests were performed with a Dunnett's multiple comparisons test if variance between groups was not equal. Error bars represent standard deviation in all cases, unless otherwise noted. Figure legends indicate the number of independent experiments and statistical subjects performed in each analysis.

DATA AVAILABILITY

The raw sequence data from our study have been deposited in the Genome Sequence Archive in National Genomics Data Center (<https://www.cnbc.ac.cn/>) under the BioProject accession code PRCA017779.

REFERENCES

- Shahbazi, M. N. et al. Self-organization of the human embryo in the absence of maternal tissues. *Nat. Cell Biol.* **18**, 700–708 (2016).
- Deglinerti, A. et al. Self-organization of the in vitro attached human embryo. *Nature* **533**, 251–254 (2016).

3. Zhou, F. et al. Reconstituting the transcriptome and DNA methylome landscapes of human implantation. *Nature* **572**, 660–664 (2019).
4. Xiang, L. et al. A developmental landscape of 3D-cultured human pre-gastrulation embryos. *Nature* **577**, 537–542 (2020).
5. Shahbazi, M. N. & Zernicka-Goetz, M. Deconstructing and reconstructing the mouse and human early embryo. *Nat. Cell Biol.* **20**, 878–887 (2018).
6. Itskovitz-Eldor, J. et al. Differentiation of human embryonic stem cells into embryoid bodies compromising the three embryonic germ layers. *Mol. Med.* **6**, 88–95 (2000).
7. Warmflash, A., Sorre, B., Etoc, F., Siggia, E. D. & Brivanlou, A. H. A method to recapitulate early embryonic spatial patterning in human embryonic stem cells. *Nat. Methods* **11**, 847–854 (2014).
8. Martyn, I., Kanno, T. Y., Ruzo, A., Siggia, E. D. & Brivanlou, A. H. Self-organization of a human organizer by combined Wnt and Nodal signalling. *Nature* **558**, 132–135 (2018).
9. Simunovic, M. et al. A 3D model of a human epiblast reveals BMP4-driven symmetry breaking. *Nat. Cell Biol.* **21**, 900–910 (2019).
10. Simunovic, M., Siggia, E. D. & Brivanlou, A. H. In vitro attachment and symmetry breaking of a human embryo model assembled from primed embryonic stem cells. *Cell Stem Cell* **29**, 962–972.e4 (2022).
11. Moris, N. et al. An in vitro model of early anteroposterior organization during human development. *Nature* **582**, 410–415 (2020).
12. Shao, Y. et al. Self-organized amniogenesis by human pluripotent stem cells in a biomimetic implantation-like niche. *Nat. Mater.* **16**, 419–425 (2017).
13. Shao, Y. et al. A pluripotent stem cell-based model for post-implantation human amniotic sac development. *Nat. Commun.* **8**, 208 (2017).
14. Zheng, Y. et al. Controlled modelling of human epiblast and amnion development using stem cells. *Nature* **573**, 421–425 (2019).
15. Yamanaka, Y. et al. Reconstituting human somitogenesis in vitro. *Nature* **614**, 509–520 (2023).
16. Miao, Y. et al. Reconstruction and deconstruction of human somitogenesis in vitro. *Nature* **614**, 500–508 (2023).
17. Kagawa, H. et al. Human blastoids model blastocyst development and implantation. *Nature* **601**, 600–605 (2022).
18. Yanagida, A. et al. Naive stem cell blastocyst model captures human embryo lineage segregation. *Cell Stem Cell* **28**, 1016–1022.e4 (2021).
19. Yu, L. et al. Blastocyst-like structures generated from human pluripotent stem cells. *Nature* **591**, 620–626 (2021).
20. Sozen, B. et al. Reconstructing aspects of human embryogenesis with pluripotent stem cells. *Nat. Commun.* **12**, 5550 (2021).
21. Fan, Y. et al. Generation of human blastocyst-like structures from pluripotent stem cells. *Cell Discov.* **7**, 81 (2021).
22. Liu, X. et al. Modelling human blastocysts by reprogramming fibroblasts into iBlastoids. *Nature* **591**, 627–632 (2021).
23. Tu, Z. et al. Modeling human pregastrulation development by 3D culture of blastoids generated from primed-to-naive transitioning intermediates. *Protein Cell* **14**, 337–349 (2023).
24. Mole, M. A. et al. A single cell characterisation of human embryogenesis identifies pluripotency transitions and putative anterior hypoblast centre. *Nat. Commun.* **12**, 3679 (2021).
25. Kinoshita, M. et al. Capture of mouse and human stem cells with features of formative pluripotency. *Cell Stem Cell* **28**, 453–471.e8 (2021).
26. Yu, L. et al. Derivation of intermediate pluripotent stem cells amenable to primordial germ cell specification. *Cell Stem Cell* **28**, 550–567.e12 (2021).
27. Nakamura, T. et al. A developmental coordinate of pluripotency among mice, monkeys and humans. *Nature* **537**, 57–62 (2016).
28. Bergmann, S. et al. Spatial profiling of early primate gastrulation in utero. *Nature* **609**, 136–143 (2022).
29. Boroviak, T. et al. Single cell transcriptome analysis of human, marmoset and mouse embryos reveals common and divergent features of preimplantation development. *Development* **145**, dev167833 (2018).
30. Ma, H. et al. In vitro culture of cynomolgus monkey embryos beyond early gastrulation. *Science* **366**, eaax7890 (2019).
31. Sasaki, K. et al. The germ cell fate of cynomolgus monkeys is specified in the nascent amnion. *Dev. Cell* **39**, 169–185 (2016).
32. Theunissen, T. W. et al. Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. *Cell Stem Cell* **15**, 471–487 (2014).
33. Takashima, Y. et al. Resetting transcription factor control circuitry toward ground-state pluripotency in human. *Cell* **158**, 1254–1269 (2014).
34. Bayerl, J. et al. Principles of signaling pathway modulation for enhancing human naive pluripotency induction. *Cell Stem Cell* **28**, 1549–1565.e12 (2021).
35. Ai, Z. et al. Modulation of Wnt and Activin/Nodal supports efficient derivation, cloning and suspension expansion of human pluripotent stem cells. *Biomaterials* **249**, 120015 (2020).
36. Guo, G. et al. Epigenetic resetting of human pluripotency. *Development* **144**, 2748–2763 (2017).
37. Liu, X. et al. Comprehensive characterization of distinct states of human naive pluripotency generated by reprogramming. *Nat. Methods* **14**, 1055–1062 (2017).
38. Di Stefano, B. et al. Reduced MEK inhibition preserves genomic stability in naive human embryonic stem cells. *Nat. Methods* **15**, 732–740 (2018).
39. Bredenkamp, N., Stirparo, G. G., Nichols, J., Smith, A. & Guo, G. The cell-surface marker sushi containing domain 2 facilitates establishment of human naive pluripotent stem cells. *Stem Cell Rep.* **12**, 1212–1222 (2019).
40. Bi, Y. et al. Identification of ALPPL2 as a naive pluripotent state-specific surface protein essential for human naive pluripotency regulation. *Cell Rep.* **30**, 3917–3931.e5 (2020).
41. Guo, G. et al. Human naive epiblast cells possess unrestricted lineage potential. *Cell Stem Cell* **28**, 1040–1056.e6 (2021).
42. Io, S. et al. Capturing human trophoblast development with naive pluripotent stem cells in vitro. *Cell Stem Cell* **28**, 1023–1039.e13 (2021).
43. Okae, H. et al. Derivation of human trophoblast stem cells. *Cell Stem Cell* **22**, 50–63.e6 (2018).
44. Linneberg-Agerholm, M. et al. Naive human pluripotent stem cells respond to Wnt, Nodal and LIF signalling to produce expandable naive extra-embryonic endoderm. *Development* **146**, dev180620 (2019).
45. Boroviak, T. et al. Lineage-specific profiling delineates the emergence and progression of naive pluripotency in mammalian embryogenesis. *Dev. Cell* **35**, 366–382 (2015).
46. Pham, T. X. A. et al. Modeling human extraembryonic mesoderm cells using naive pluripotent stem cells. *Cell Stem Cell* **29**, 1346–1365.e10 (2022).
47. Xu, P. F. et al. Construction of a mammalian embryo model from stem cells organized by a morphogen signalling centre. *Nat. Commun.* **12**, 3277 (2021).
48. Smith, A. Formative pluripotency: the executive phase in a developmental continuum. *Development* **144**, 365–373 (2017).
49. Ross, C. & Boroviak, T. E. Origin and function of the yolk sac in primate embryogenesis. *Nat. Commun.* **11**, 3760 (2020).
50. Ai, Z., Yin, Y., Niu, B. & Li, T. Deconstructing human peri-implantation embryogenesis based on embryos and embryoids. *Biol. Reprod.* **107**, 212–225 (2022).
51. Yan, L. et al. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat. Struct. Mol. Biol.* **20**, 1131–1139 (2013).
52. Van der Jeught, M. et al. Treatment of human embryos with the TGFbeta inhibitor SB431542 increases epiblast proliferation and permits successful human embryonic stem cell derivation. *Hum. Reprod.* **29**, 41–48 (2014).
53. Blakeley, P. et al. Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. *Development* **142**, 3613 (2015).
54. Stirparo, G. G. et al. Integrated analysis of single-cell embryo data yields a unified transcriptome signature for the human pre-implantation epiblast. *Development* **145**, dev158501 (2018).
55. Tsai, T. Y. et al. An adhesion code ensures robust pattern formation during tissue morphogenesis. *Science* **370**, 113–116 (2020).
56. Steinberg, M. S. Does differential adhesion govern self-assembly processes in histogenesis? Equilibrium configurations and the emergence of a hierarchy among populations of embryonic cells. *J. Exp. Zool.* **173**, 395–433 (1970).
57. Bao, M. et al. Stem cell-derived synthetic embryos self-assemble by exploiting cadherin codes and cortical tension. *Nat. Cell Biol.* **24**, 1341–1349 (2022).
58. Sozen, B. et al. Self-assembly of embryonic and two extra-embryonic stem cell types into gastrulating embryo-like structures. *Nat. Cell Biol.* **20**, 979–989 (2018).
59. Lovell-Badge, R. et al. ISSCR guidelines for stem cell research and clinical translation: the 2021 update. *Stem Cell Rep.* **16**, 1398–1408 (2021).
60. Daley, G. Q. et al. Setting global standards for stem cell research and clinical translation: the 2016 ISSCR guidelines. *Stem Cell Rep.* **6**, 787–797 (2016).
61. Gardner, D. K., Lane, M., Stevens, J., Schlenker, T. & Schoolcraft, W. B. Blastocyst score affects implantation and pregnancy outcome: towards a single blastocyst transfer. *Fertil. Steril.* **73**, 1155–1158 (2000).
62. Horii, M. et al. Human pluripotent stem cells as a model of trophoblast differentiation in both normal development and disease. *Proc. Natl. Acad. Sci. USA* **113**, E3882–E3891 (2016).
63. Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784 (2014).
64. Chan, Y. S. et al. Induction of a human pluripotent state with distinct regulatory circuitry that resembles preimplantation epiblast. *Cell Stem Cell* **13**, 663–675 (2013).
65. Yang, Y. et al. Derivation of pluripotent stem cells with in vivo embryonic and extraembryonic potency. *Cell* **169**, 243–257.e25 (2017).
66. Gao, X. et al. Establishment of porcine and human expanded potential stem cells. *Nat. Cell Biol.* **21**, 687–699 (2019).
67. An, C. et al. Overcoming autocrine FGF signaling-induced heterogeneity in naive human ESCs enables modeling of random X chromosome inactivation. *Cell Stem Cell* **27**, 482–497.e4 (2020).

68. Irie, N. et al. SOX17 is a critical specifier of human primordial germ cell fate. *Cell* **160**, 253–268 (2015).
69. Sperber, H. et al. The metabolome regulates the epigenetic landscape during naive-to-primed human embryonic stem cell transition. *Nat. Cell Biol.* **17**, 1523–1535 (2015).
70. Guo, G. et al. Naive pluripotent stem cells derived directly from isolated cells of the human inner cell mass. *Stem Cell Rep.* **6**, 437–446 (2016).
71. Wu, J. et al. An alternative pluripotent state confers interspecies chimaeric competency. *Nature* **521**, 316–321 (2015).
72. Xi, Y. & Li, W. BSMAP: whole genome bisulfite sequence MAPping program. *BMC Bioinformatics* **10**, 232 (2009).
73. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
74. Yang, R. et al. Amnion signals are essential for mesoderm formation in primates. *Nat. Commun.* **12**, 5126 (2021).
75. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).
76. Jin, S. et al. Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* **12**, 1088 (2021).

ACKNOWLEDGEMENTS

This work was supported by the National Key R&D Program of China (2022YFA1103100 and 2021YFA0805700), the National Natural Science Foundation of China (32130034, 82192874, 32100655 and 82160294), the National Key R&D Program of China (2020YFA0112700), the Major Basic Research Project of Science and Technology of Yunnan (202102AA100007, 202001BC070001 and 2019FY002), the Joint Special Funds for the Department of Science and Technology of Yunnan Province-Kunming Medical University (202101AY070001-260 and 202301AY070001-298). We thank Yan Bi for sharing ALPPL2 antibody. We acknowledge Ran Yang for providing technical details of the single-cell dissociation and preparation of human embryos.

AUTHOR CONTRIBUTIONS

T.L. and W.J. initiated the project. T.L. supervised the entire project. T.L. and Z.A. designed the experiments and wrote the manuscript. T.L. supervised Z.A., Y.Y., K.D. and B.N. to perform data analysis. Z.A., B.N., Y.Y., S.W., Y.H., C.Z., C.T.Z., R.K., T.C., N.L., S.Z. and X.Z. performed stem cell isolation and culture, embryoid assembly and

identification, culture system development and mechanistic studies. L.X., G.S., L.R., X.M., Y.L. and Z.W. performed human embryo culture. Y.Y., Z.A., K.D., Y.G. and W.L. analyzed sequencing data. Y.Z. and J.F. contributed to sequencing data analysis and manuscript preparation.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41422-023-00846-8>.

Correspondence and requests for materials should be addressed to Zongyong Ai, Weizhi Ji or Tianqing Li.

Reprints and permission information is available at <http://www.nature.com/reprints>



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023