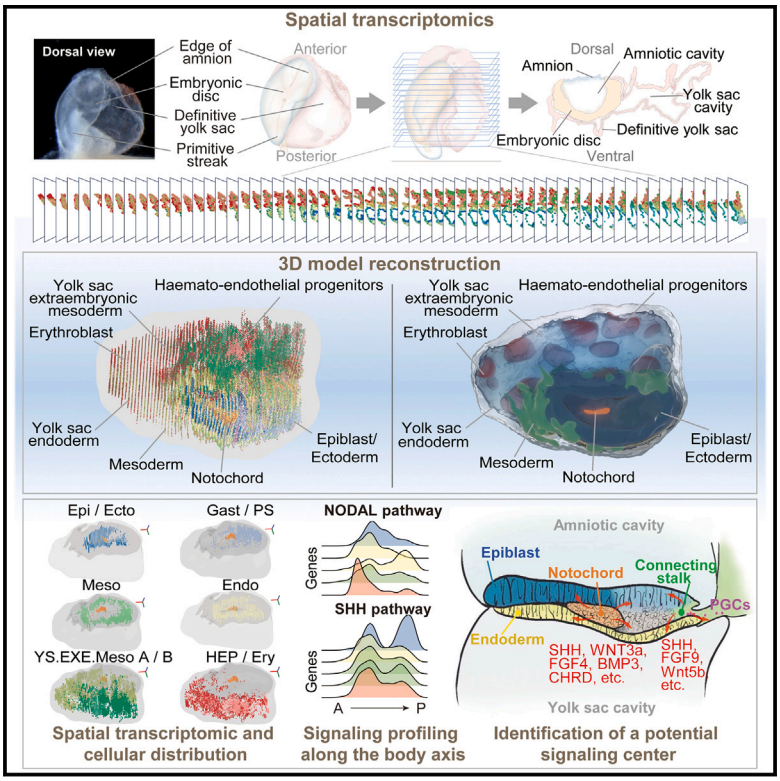
Cell

Resource

# 3D reconstruction of a gastrulating human embryo

## Graphical abstract




## Highlights

- Reconstructing a 3D model of a gastrulating human embryo using spatial transcriptomics
- Annotating cell lineages by integrating gene expression patterns and spatial information
- Identifying a potential signaling center at the posterior end of the human embryo
- Investigating the dynamic activity of signaling pathways along the embryonic body axis


## Authors

Zhenyu Xiao, Lina Cui, Yang Yuan, ..., Yulei Wei, Jingtao Guo, Leqian Yu

## Correspondence

xyan.wang@ioz.ac.cn (X.W.), yulei.wei@cau.edu.cn (Y.W.), jingtao.guo@ioz.ac.cn (J.G.), leqianyu@ioz.ac.cn (L.Y.)



## In brief

A 3D model of an intact Carnegie stage 8 human embryo is constructed from spatial transcriptomics profiling of serial transverse sections.


Xiao et al., 2024, Cell *187*, 2855–2874
May 23, 2024 
https://doi.org/10.1016/j.cell.2024.03.041

CellPress

Cell

CellPress

Resource

# 3D reconstruction of a gastrulating human embryo

Zhenyu Xiao,[1,2,6,7,8] Lina Cui,[1,5,6,7,8] Yang Yuan,[1,3,5,6,7,8] Nannan He,[4,8] Xinwei Xie,[2] Sirui Lin,[3] Xiaolong Yang,[2] Xin Zhang,[3] Peifu Shi,[2] Zhifeng Wei,[2] Yang Li,[2] Hongmei Wang,[1,2,5,6,7,9] Xiaoyan Wang,[1,6,7,*] Yulei Wei,[3,*] Jingtao Guo,[1,5,6,7,*] and Leqian Yu[1,5,6,7,10,*]

[1]Key Laboratory of Organ Regeneration and Reconstruction, State Key Laboratory of Stem Cell and Reproductive Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China
[2]School of Life Science, Beijing Institute of Technology, Beijing 100081, China
[3]State Key Laboratory of Animal Biotech Breeding, College of Biological Sciences, China Agricultural University, Beijing 100193, China
[4]Department of Gynecology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan 450052, China
[5]University of Chinese Academy of Sciences, Beijing 100049, China
[6]Institute for Stem Cell and Regeneration, Chinese Academy of Sciences, Beijing 100101, China
[7]Beijing Institute for Stem Cell and Regenerative Medicine, Beijing 100101, China
[8]These authors contributed equally
[9]Senior author
[10]Lead contact
*Correspondence: xyan.wang@ioz.ac.cn (X.W.), yulei.wei@cau.edu.cn (Y.W.), jingtao.guo@ioz.ac.cn (J.G.), leqianyu@ioz.ac.cn (L.Y.)
https://doi.org/10.1016/j.cell.2024.03.041

## SUMMARY

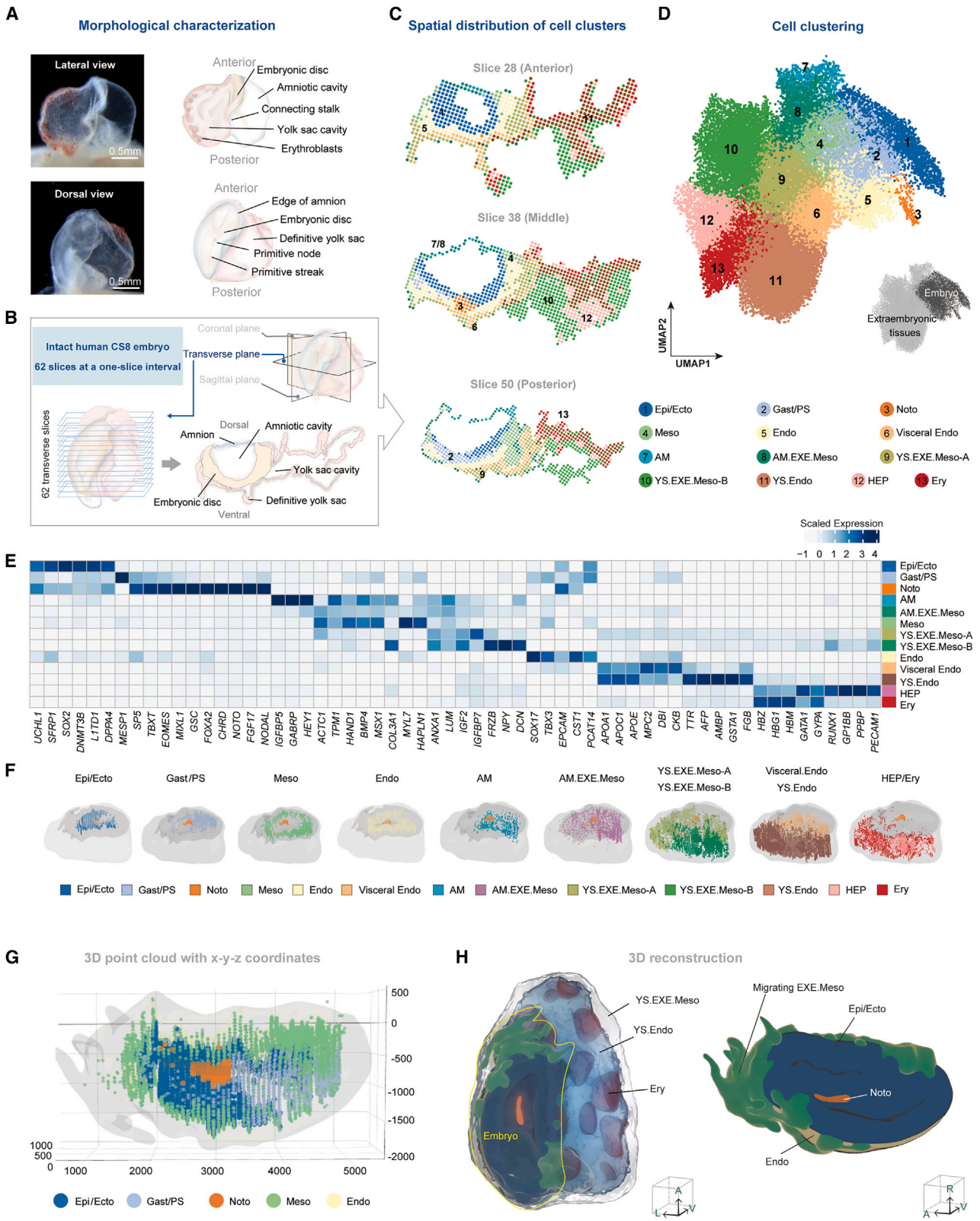
Progress in understanding early human development has been impeded by the scarcity of reference datasets from natural embryos, particularly those with spatial information during crucial stages like gastrulation. We conducted high-resolution spatial transcriptomics profiling on 38,562 spots from 62 transverse sections of an intact Carnegie stage (CS) 8 human embryo. From this spatial transcriptomic dataset, we constructed a 3D model of the CS8 embryo, in which a range of cell subtypes are identified, based on gene expression patterns and positional register, along the anterior-posterior, medial-lateral, and dorsal-ventral axis in the embryo. We further characterized the lineage trajectories of embryonic and extra-embryonic tissues and associated regulons and the regionalization of signaling centers and signaling activities that underpin lineage progression and tissue patterning during gastrulation. Collectively, the findings of this study provide insights into gastrulation and post-gastrulation development of the human embryo.

## INTRODUCTION

Gastrulation orchestrates the transformation of the bilaminar embryonic disc within the blastula into a complex multi-dimensional structure known as the gastrula.[1] This intricate event lays the foundation for distinct cell lineages and spatial patterning. Human gastrulation occurs between embryonic days 14 and 21 (E14–21) after fertilization, corresponding to Carnegie stages (CS) 7 to 9. Within this period, the epiblast (Epi) are allocated to precursors and derivatives of the primary germ layers and extra-embryonic tissues. In addition to cell lineage determination, establishing the body axis becomes paramount for orchestrating migration, positioning, and patterning of differentiating cells. Dynamic signaling centers such as notochord, regulate the establishment of the body axis and guide the cells to develop into their designated destinations.[2] Disruptions in this coordination can lead to severe congenital anomalies. Nonetheless, the study of human embryogenesis presents unique challenges as the scarcity of *in utero* early-stage human embryo materials limits direct investigations. Consequently, our understanding of human gastrulation heavily relies on data derived from alternative sources.

Animals, such as mice, are the major sources for acquiring knowledge about gastrulation. Particularly, the application of single-cell RNA sequencing (scRNA-seq) has provided an increasingly detailed view of the developmental gene expression and developmental path of each lineage.[3–9] However, the uniqueness of primate embryogenesis in terms of anatomical architecture, developmental timing, and lineage fate determination leads to the difficulty in understanding primate embryos solely by replying on studying the model animals. Recently developed primate blastocyst *in vitro* culture approach[10–17] and pluripotent stem-cell-based embryo models[18–27] allowed us to study, outside the uterus, the embryonic lineage ontogeny and have revealed several developmental landmarks of early primate embryos. However, the lack of relevant references makes it difficult to judge whether these stem-cell-derived embryo models can faithfully replicate embryonic development. For the in-uterus primate embryo, recent transcriptomic characterization of an early gastrulating human embryo (CS7)[28] and cynomolgus monkey gastrulating embryos,[29,30] provided crucial insights into primate gastrulation. Despite these advances, these datasets lack spatial positioning information, making it challenging to




*(legend on next page)*

accurately annotate cell subtypes. During gastrulation, how cell fate is determined at different locations and how various subtypes of cells were orchestrated and developed along the body axis remain fragmented. Lately, spatial profiling has been conducted on marmoset embryos at CS5 to CS7 following implantation.[31] This research utilized laser capture microdissection in conjunction with Smart-seq2, leading to the discovery of significant developmental events.

In this study, we employed serial transverse sectioning across 62 slices that cover a fully intact gastrulating human CS8 embryo, which allowed us to combine the spatial transcriptomes of all slices from the anterior side to the posterior side, resulting in a complete 3D reconstruction of the embryo. This dataset presents an opportunity to investigate the essential cellular and molecular characteristics of gastrulating events in humans and provides invaluable information for the development of advanced stem-cell-derived human embryo models.

## RESULTS

### Spatial transcriptomics profiling and 3D reconstruction of an intact CS8 human gastrula

The CS8 embryo showed an intact morphology, including an oval-shaped embryonic disc with a transparent amniotic cavity, a primitive node, a primitive streak extending to about one-third of the anterior-posterior axis, a yolk sac with noticeable erythroblasts, and an opaque connecting stalk connected with chorionic villi (Figure 1A). We conducted cryosection transverse slices spanning along the anterior-posterior (A-P) axis of the embryo (Figures 1B and S1A). In total, 62 transverse slices with a one-slice interval were subjected to Stereo-seq,[32] resulting in the generation of a spatial transcriptome atlas of the entire embryo. Specifically, to perform unsupervised spatially constrained clustering (SCC),[33] bin50 was utilized to represent approximately one cell. After quality control, 38,562 spots were retained for downstream analysis (Figure S1B). The sequencing data were processed and integrated into the bright field image of each slice (Figure S1A), resulting in the generation of 2D spatial transcriptome maps (Figure 1C). The transcriptomic datasets were clustered by Scanpy and visualized using uniform manifold approximation and projection (UMAP) dimension reduction.

By integration of the top differentially expressed genes (DEGs) (Table S1), known markers, and spatial information, thirteen major clusters were identified (Figures 1C–1E), including epiblast/ectoderm (Epi/Ecto, cluster 1, expressing *SOX2*, *DNMT3B*, on the dorsal part), gastrulating cell/primitive streak (Gast/PS, cluster 2, expressing *TBXT*, *MESP1*, at the middle line of the posterior region), notochord (Noto, cluster 3, expressing *NOTO*, *CHRD*, at the middle line of the anterior and middle regions), mesoderm (Meso, cluster 4, expressing *HAND1*, *MYL7*), endoderm (Endo, cluster 5, expressing *SOX17*, *TBX3*, on the ventral part), visceral endoderm (Visceral.Endo, cluster 6, expressing *MPC2*, *CKB*), amnion (AM, cluster 7, expressing *GABRP*, *IGFBP5*, covering the amniotic cavity), amniotic extra-embryonic mesoderm (AM.EXE.Meso, cluster 8, co-expressing several AM and Meso markers, at the outer layer of the amniotic membrane), two types of yolk sac extra-embryonic mesoderm (YS.EXE.Meso-A and YS.EXE.Meso-B, cluster 9 and cluster 10, expressing *IGFPB7*, *RUNX1*$^{low}$ and *IGFPB7*, *RUNX1*$^{high}$, respectively, at the outer layer of the yolk sac cavity), yolk sac endoderm (YS.Endo, cluster 11, expressing *AFP*, *TTR*, at the inner layer of the yolk sac cavity), haemato-endothelial progenitors (HEP, cluster 12, expressing *RUNX1*, *GP1BB*), and erythroblasts (Ery, cluster 13, expressing *HBZ*, *HBG1*) (Figures 1E, S1C, and S1D; Video S1; Table S2). The spots which co-expressing the trophoblast markers (e.g., *KRT7*, *LEP*, and *PSG8*) were excluded from further analysis (Figure S1C). Cell cycle analysis showed that embryonic cells are more proliferative than extra-embryonic cells (Figure S1E).

With the efficient and faithful capture of spatial transcriptome of the embryo, we performed a 3D reconstruction of the spatial transcriptome of the entire sample. The sections were aligned by PASTE,[34] which leverages both transcriptional similarity and physical distances between spots in adjacent slices and allows us to assign x-y-z 3D coordinates to each spot (Figures 1F and 1G). We then generated a 3D point cloud that represents different cell types, which was further refined, enhanced, and locally rendered, resulting in surfaces rather than separated points. This approach allowed us to generate a 3D model depicting the spatial domains of cell types in the embryo, with or without the presence of extra-embryonic tissues (Figures 1H and S1F).

### 3D spatial distribution of cell types and associated molecular regulatory programs

Our 3D model allowed us to extract the 2D (x-y axis) transcriptomes and cell distributions of all slides spanning and to visualize along the z axis (A-P axis) (Figure 2A; Table S3). Extra-embryonic cell types can be found in all slices, while embryonic

**Figure 1. Spatial transcriptomic profiling and 3D reconstruction of a CS8 human embryo**

(A) Lateral (top) and dorsal (bottom) views of the intact CS8 human embryo with the complete embryonic disc, amnion, yolk sac, and primitive streak. Scale bar, 0.5 mm.

(B) Schematic diagram of the sectioning strategy for spatial transcriptomic profiling. The yolk sac cavity was pushed to one side of the embryo after optimal cutting temperature (OCT) embedding.

(C) Unsupervised spatially constrained clustering of representative CS8 human embryo section from anterior to posterior end. Bins/dots are colored by their annotations from (D).

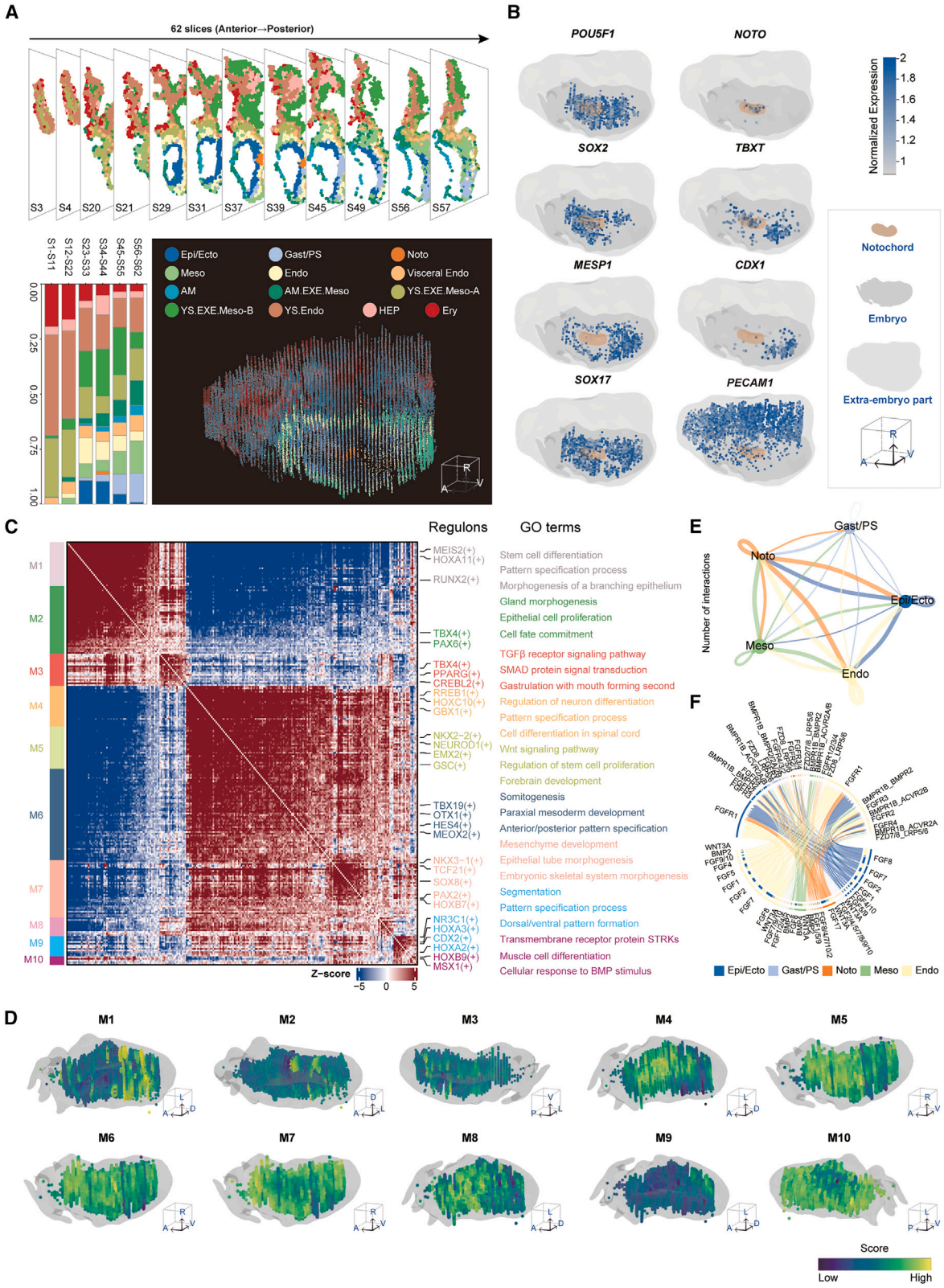
(D) UMAP of all the spots computed from the expression of highly variable genes. Spots from the whole sample can be grouped into 13 clusters, including 5 clusters of embryonic tissues and 8 clusters of extra-embryonic tissues. Bins/dots are colored by their cell type annotations.

(E) Heatmap showing the normalized expression of specific marker genes for each cluster.

(F) Spatial distribution of major cell types in the 3D reconstructed model.

(G) The 3D scatterplot shows the representation of the point cloud with x-y-z coordinates for each spot fitting inside the 3D reconstructed embryo model. Bins/dots are colored by their annotations.

(H) The 3D reconstructed models of the whole embryo with (left) or without (right) the extra-embryonic tissues.




*(legend on next page)*

clusters were present in Slice (S) 21-62 (Figure 2A and S2A). We examined the 3D expression patterns of known markers, and specifically, *POU5F1* displayed expression throughout the entire dorsal region of the embryonic disc, *SOX2* in the anterior part, and *SOX17* in the posterior region of the ventral part. *NOTO* was primarily expressed within the Noto, while *TBXT* and *CDX1* were primarily expressed posterior to the Noto (Figure 2B). We created a webpage (cs8.3dembryo.com) to provide easy access to the spatial patterns of additional genes.

Next, we applied Hotspot analysis[35] to explore the Gene Ontology (GO) and pathway enrichment in the 3D structure. First, we applied SCENIC analysis[36] to generate a heatmap illustrating the top 5 transcription factors (TFs) whose regulons are enriched in different clusters (Figures 2C and S2B). Subsequent gene regulatory network analysis showed the relationship of gene regulation, such as *SOX2*, which expressed highly in Epi/Ecto, its targets *FGF2* and *UCHL1* in the regulation network, and also *NOTO*, *FGF17* with *CHRD*, *NODAL*, *NANOG*, *MIXL1*, and *TBXT* (Figure S2C). To further identify spatial regulon modules, we used the regulon specificity scores (RSS) to identify 10 regulon modules in tissue clusters based on 281 total regulons (Figures 2C, S2D, and S2E). The resulting regulon programs exhibited well-defined spatial patterns that largely corresponded to anatomical regions. Projection of the regulon programs into 3D space showed that TFs regulated gastrulation in a spatially specific manner (Figures 2C, 2D, and S2E). For instance, module M5 (e.g., *NEUROD1*, *EMX2*, *GSC*), enriched in "WNT pathways", "stem-cell proliferation", and "forebrain development", was highly expressed in the anterior-ventral part (Figures 2C, 2D, and S2E). Module M6 (e.g., *MEOX2*, *HES4*, *TBX19*), showed enrichment in "anterior/posterior pattern specification", "somitogenesis", and "paraxial mesoderm development", consistent with prior findings about *MEOX2*[37,38] and *HES4*.[29,39] Module M9 was at the posterior part (e.g., *HOXA3*, *HOXD11*, *CDX2*), with enriched terms in "pattern specification process", "dorsal/ventral pattern formation", and "segmentation".

We also investigated cell communication regulation using our dataset using CellChat analysis[40] (Figures 2E, 2F, S2F, and S2G). The network diagram of cell interactions provided insights into the specific interactions among tissue clusters within the CS8 embryo (Figures 2E, 2F, S2F, and S2G). Noto was revealed as the source of FGF, IGF, BMP, and WNT pathways and ligand-receptor pairs such as *FGF17-FGFR1/2/3*, *IGF2-IGF1R*, *BMP4-BMPR2*, and *WNT3A-FZD2/7/8* (Figures 2F, S2G, and S2H), which were known to regulate cell migration and differentiation.[41]

## Spatial transcriptomic resolution of mesoderm formation

We next characterized the spatial allocation and differentiation of distinct Meso subtypes. Based on refined clustering analysis of related clusters (clusters 2, 3, and 4 in Figure 1D), and in the context of the spatial localization, 5 Meso subtypes can be further identified (Figures 3A and 3B; Video S2) —including axial mesoderm (Axial.Meso), proximal to the posterior end of Noto and expressing *SFRP2*, *DCLK1*, and *ESRG*, and weakly expressing Noto related genes such as *CHRD* and *NOTO*; lateral plate mesoderm (LP.Meso), extending along the lateral periphery of the embryonic disc, spanning from the anterior to the posterior end of the embryo, and expressing LP.Meso-specific markers such as *KDR*, *MYL7*, and *HAPLN1*; paraxial mesoderm (Para.Meso), positioned at the rear of the Axial.Meso and flanking the Axial.Meso from both sides, and expressing *MESP1/2*, *MSGN1,* and *HES7*; intermediate mesoderm (Inter.Meso), positioned between the LP.Meso and Para.Meso, and expressing Inter.Meso-specific markers *OSR1*, *EYA1*, and *CDH5*; extra-embryonic mesoderm progenitor (EXE.Meso.Prog), located at the outermost region of the embryonic disc—but have not yet migrated beyond the embryo's confines, and expressing early EXE.Meso markers *LUM*, *POSTN*, and *IGF2*. Moreover, we captured a cluster of cells located at the posterior end of the embryo and expressing *CDX1* and *CDX2*, which are known as the connecting stalk markers,[29,42] and thus annotated it as the connecting stalk (Figures 3B, 3D, S3A, and S3B). Interestingly, these connecting stalk cells also highly expressed several essential morphogens such as *SHH*, *WNT5B*, and *FGF9*, which will be discussed in the subsequent part. In addition, we used SCENIC to analyze the segmented clusters and identified enrichment for TFs with activities in different Meso subtypes (Figure 3E).

The specification of Meso in the vertebrate embryo is initiated and driven by dynamic morphogen gradients along the anterior-posterior axis of the embryo and the PS. After gastrulation, Meso patterning into subtypes strictly correlates with their spatial and temporal allocation.[43,44] However, how Meso is patterned in the human embryo is largely unknown. The human pluripotent stem cells-based Meso differentiation experiments suggested that, during gastrulation, cells at different locations in the PS responded to different inductions and generated progenitors to form distinct Meso subtypes along the PS.[45] In agreement with this model, in the CS8 human embryo, we observed the distinct subtypes of Meso were already specified within the PS structure (ahead of migration) and arranged along the midline of the PS in order (A to P, Axial.Meso, Para.Meso, Inter.Meso, LP.Meso, EXE.Meso.Prog) (Figures 3A, 3C, and S3B). Notably, the 3D

**Figure 2. 3D spatial distribution of cell types and associated molecular regulatory programs**

(A) Top: diagram of aligning the slices along the anterior-posterior axis. Bottom left: bar plot showing the distribution of cell types in slices along the A-P axis. Bottom right: spatial distribution of major cell types organized in all slices along the A-P axis.

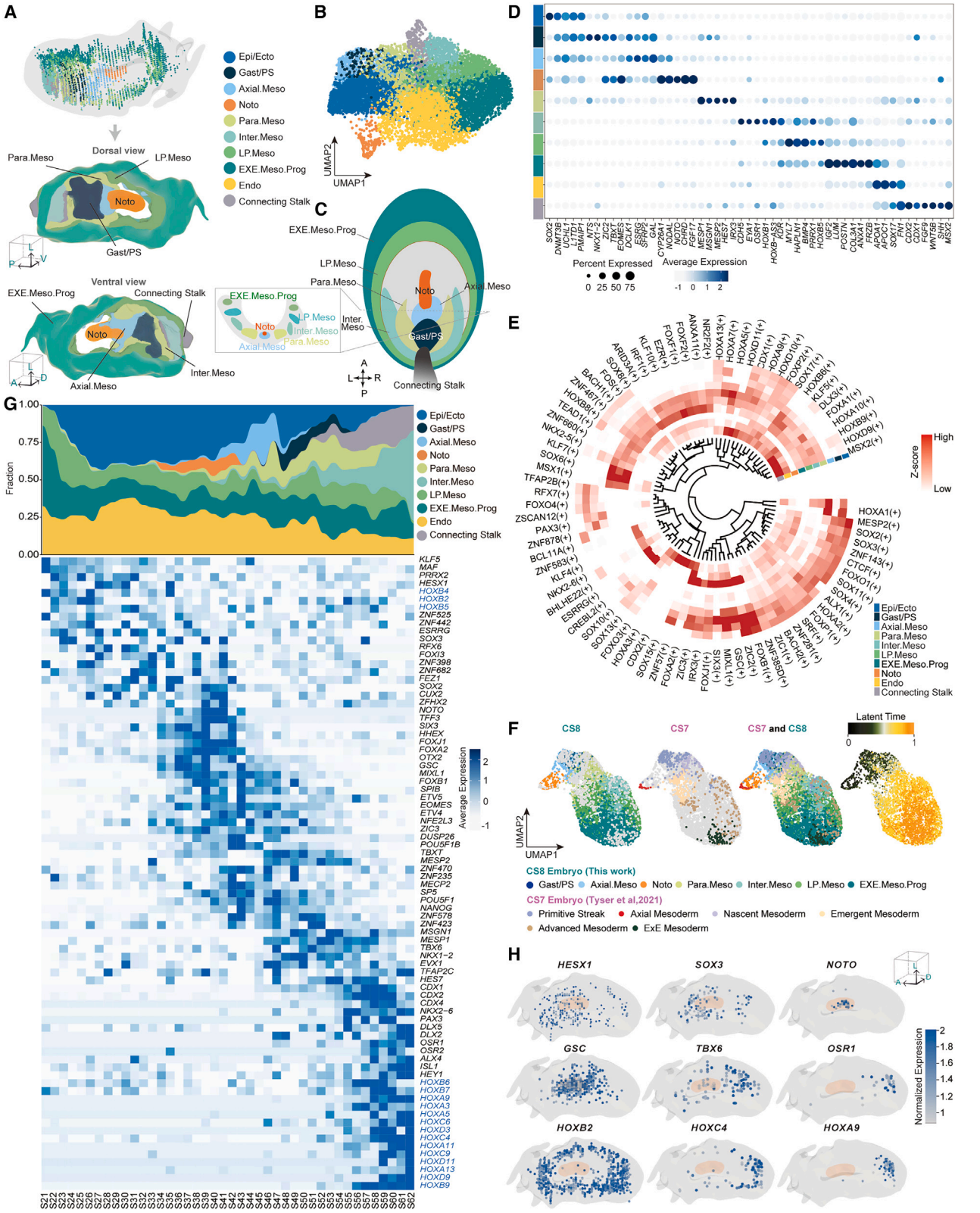
(B) Spatial expression patterns of key representative genes in the 3D reconstructed CS8 human embryo. Colors indicate normalized gene expression levels.

(C) Heatmap showing the regulons with significant spatial correlation grouped into 10 modules based on pairwise spatial correlations, with the corresponding GO terms related to each regulon module highlighted on the right side.

(D) Spatial patterns of functional regulon modules identified by hotspot analysis in the 3D reconstructed model.

(E) The chord diagram shows the ligand-receptor interaction pairs between Noto and other tissue clusters including Endo, Meso, and Epi/Ecto. The thickness of the line represents the count of ligand-receptor pairs.

(F) The chord diagrams showing the significant ($p < 0.05$) interaction pairs associated with key signaling pathways (including WNT, IGF, FGF, and BMP) among Endo, Meso, Epi/Ecto, and Noto.




*(legend on next page)*

spatial distribution pattern shows that the LP.Meso and EXE.Meso.Prog has already been distributed throughout the lateral side of the embryonic disc. Meanwhile, other types of Meso (e.g., Para.Meso, Inter.Meso) remain along the midline of the PS and have not migrated. This suggests that the migration of LP.Meso and EXE.Meso.Prog may occur earlier than that of others. In the previous CS7 human embryo dataset, Meso were categorized based on their developmental "age" with classifications ranging from advanced mesoderm to emergent mesoderm and nascent mesoderm, representing early to later emerging cells.[28] Subsequently, we integrated the CS7 datasets and conducted a pseudotime analysis. The results revealed a close clustering of primitive streak (CS7) cells with Gast/PS (CS8) cells (Figure 3F). Using Gast/PS as the reference point and considering how far these mesoderm subtypes had emerged from the primitive streak, the forthcoming results indicate that EXE.Meso.Prog and LP.Meso (corresponding to ExE mesoderm and advanced mesoderm in CS7) emerge relatively earlier, while Inter.Meso, Para.Meso, and Axial.Meso (corresponding to the emergent and nascent mesoderm in CS7) emerge later during gastrulation (Figures 3F and S3C). This finding aligns with the previously reported observations.[28]

## Patterning of mesoderm along the A-P axis

We then investigated the cell distribution and gene expression patterns along the embryo A-P axis. We observed that the LP.Meso and EXE.Meso.Prog are evenly spread along the A-P axis and the other subtypes are distributed at the posterior end in the order Axial.Meso, Para.Meso, and Inter.Meso (Figures 3C, 3G, and S3B). In these Meso, key genes, such as the HOX family, also displayed dynamic expression patterns along the A-P axis. The HOX proteins are known to interact with each other to regulate development programs and specify the regions of the body plan.[46,47] We have identified the HOX gene localization from the most anterior end (e.g., *HOXB4*, *HOXB2*, and *HOXB5*) to the most posterior end (*HOXA5*, *HOXC9*, and *HOXB9*) (Figures 3E and 3G), promoting further investigation to understand their potential roles.

We next performed clustering analysis for all TFs expressed in the Meso and obtained 10 distinct modules based on their A-P axis expression patterns (Figure S3D; Table S4). For instance, the *KLF5*, *MAF*, *PRRX2*, etc. (module 1) were specifically expressed in the most anterior end of embryos; *DLX 2/5*, *CDX 1/2/4*, *ALX4*, etc. (module 8) were expressed in the most posterior end; and *SIX3*, *OTX2*, *ZIC3*, etc. (modules 3 and 4) were expressed in the middle (Figures 3E, 3G, 3H, S3D, and S3E; Table S4). Multiple zinc finger (ZNF) genes exhibited specific distribution patterns along the A-P axis (Figures 3E and 3G) and in different Meso subtypes (Figure 3E). Given the critical roles of ZNFs to interact with DNA, RNA, and proteins to regulate multiple cellular processes,[48] indicating the potential functions of ZNFs in establishing the body plan and lineage development.

## Development of ectoderm and amnion

To explore Epi differentiation in humans, we performed sub-clustering of related cells (cluster 1 and 7 in Figure 1D), which resulted in four distinct cellular clusters: the Epi, characterized by the expression of *SOX2*, *POU5F1*, and *DNMT3B*; the Ecto, which co-expresses Epi markers alongside neuron-related markers such as *SOX3* and *PAX6*; the amniotic ectoderm (AM.Ecto), marked by *TFAP2A*, *ISL1*, and *GATA3* expression; and the advanced amnion (AM), expressing *GABRP*, *IGFBP3/5*, and *NPNT* (Figures 4A–4D and S4A). Spatially, the Ecto occupies the anterior region, while the Epi is predominantly positioned in the posterior part (Figures 4B and 4C). The AM.Ecto is positioned at the border between the embryonic disc and the AM, while the AM is distributed across the entire amniotic membrane (Figure 4B). Utilizing diffusion maps and RNA velocity analysis, we unveiled intricate developmental trajectories stemming from the Epi. These trajectories diverge into two streams, corresponding to Ecto or AM development, respectively (Figure 4A). We next performed immunostaining of the AM.Ecto marker *TFAP2C* and advanced AM markers *GABBR2*, *IGFBP3*, and *NPNT* (Figure 4D) and found the AM.Ecto cells (*TFAP2C*+) located at the interface between Epi and the advanced AM cells (*GABBR2*+, *IGFBP3*+, *NPNT*+) (Figure 4E), supporting the trajectory finding that AM.Ecto is a transitory cell type from Epi to AM.

A previous study suggested there are two independent waves of amniogenesis in primates, in which the AM formed during the early (1st) wave via a trophectoderm-like route and the late (2nd) wave derived from the nonneural ectoderm.[49] The PCA analysis and heatmap results indicated that the AM, at the CS8 stage, showed high similarity with the AM cells generated during the late wave of amniogenesis (late wave, AME-L) (Figure S4B and S4C), and the AM.Ecto serves as an intermediate cell type, during the late wave amniogenesis, between the Epi and AM.

We next identified the top variant genes during human AM development such as *PRTG*, *ID2*, and *SLC2A3* were up-regulated at the early stage of AM development and *POSTN*, *COL5A2*, *EPAS1*, and others that were highly expressed in the advanced AM. In contrast, *SNRPN*, *PKM*, and *JARID2* were highly expressed in the Ecto but gradually down-regulated during AM formation (Figures 4F, S4D, and S4E).

**Figure 3. Spatial distribution of Meso and the A-P pattern**

(A) 3D reconstructed model of Meso displaying distribution pattern of Meso subtypes. Ventral and dorsal views of the models were shown.

(B) UMAP plot showing the 10 major cell types identified in the human CS8 embryo.

(C) Schematic diagram showing the spatial location and cross-section view of Meso subtypes.

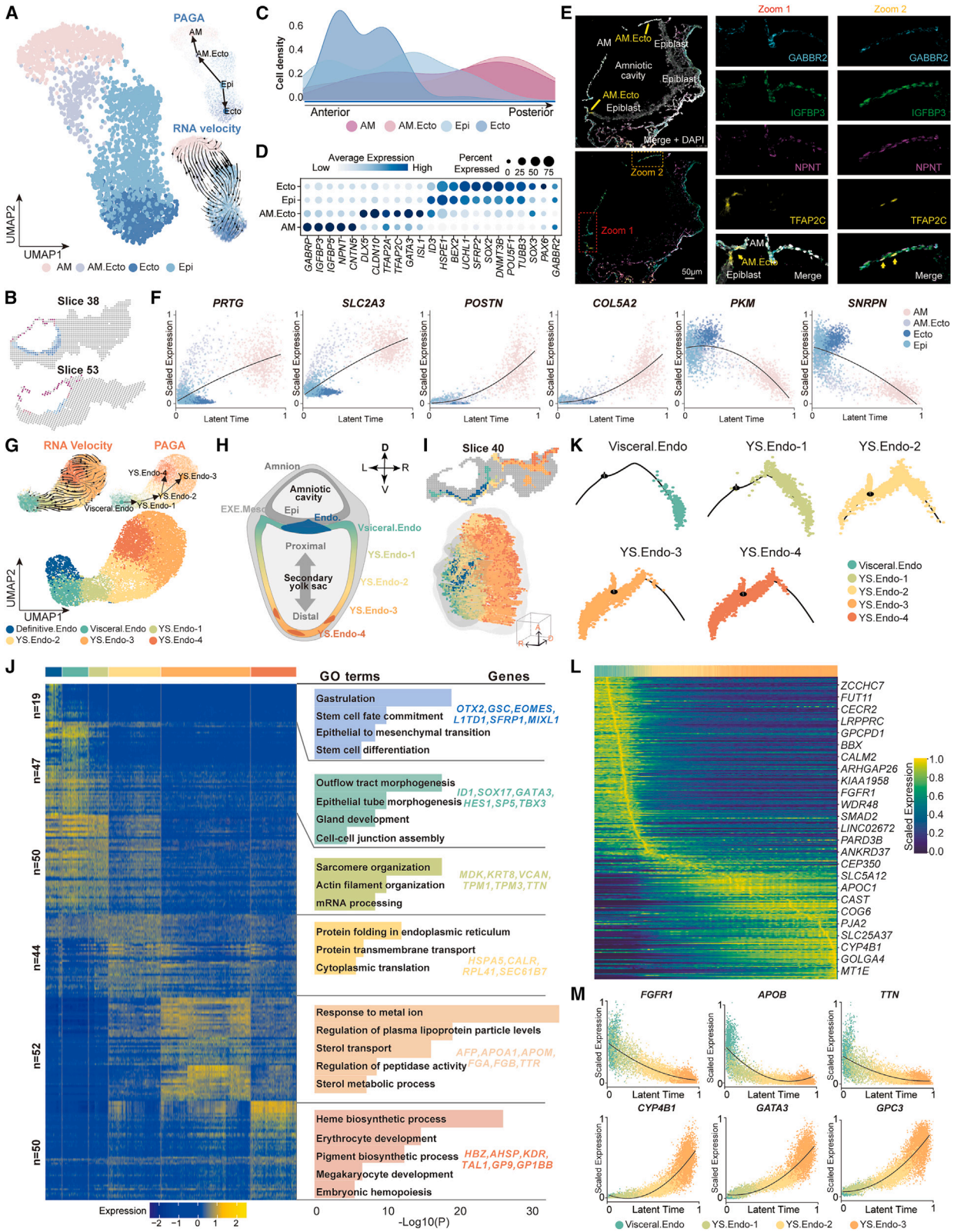
(D) Dotplot depicting the normalized expression levels of representative marker genes utilized to annotate Meso subtypes. The color-code of Meso subtypes is consistent with (A).

(E) Circular heatmap showing the representative TFs in different cell types.

(F) UMAP visualization of the Meso from human CS7 and CS8 gastrula. Colors indicate the cell types or latent time of Meso.

(G) Top: the stacked area plot showing the fraction of major human embryonic cell populations in the CS8 human embryo along the A-P axis. Bottom: the heatmap showing the gene expression patterns along the A-P axis, encompassing the representative genes from the HOX family (blue).

(H) Spatial expression patterns of representative genes. Colors indicate normalized gene expression levels.




*(legend on next page)*

### Development of endodermal cells

Recent studies reported that the Visceral.Endo also contribute to embryonic gut tube formation.[29,50] To understand Endo development, we performed an in-depth clustering of Endo-related lineages (cluster 5, 6, and 11 in Figure 1D) and identified six distinct clusters, including the definitive endoderm (Definitive.Endo), Visceral.Endo, and four YS.Endo subtypes (YS.Endo-1 to 4) (Figures 4G and S4F). We can then distinguish the locations of these cell types in the 3D model (Figures 4H and 4I). The Definitive.Endo occupies the underneath of the embryonic disc, and the Visceral.Endo is exclusively positioned at the area from the lateral side of the Definitive.Endo to the edge of the embryonic disc. The YS.Endo-1, YS.Endo-2, and YS.Endo-3 clusters are arranged sequentially along the membrane of the yolk sac in consecutive domains between the hypoblast/Visceral.Endo interface and the distant region of the yolk sac (Figures 4H and 4I). Notably, the spatial distribution of YS.Endo-4 highly overlaps with that of YS.Endo-3, although it manifests a distinctive gene expression pattern involving several hematopoietic-related genes (Figure 4J).

After implantation, the hypoblast differentiate into Visceral.Endo and YS.Endo. Both the VE and YS.Endo contributes to the formation of the primary yolk sac.[29,51] The organized sequence in which the YS.Endo subtypes are positioned along the proximal-distal axis of the yolk sac, prompting us to inquire whether these distinct cell types recapitulate the different stages. We conducted RNA velocity and Monocle analysis, revealing a trajectory originating from the Visceral.Endo and advancing through YS.Endo-1 to YS.Endo-2. Subsequently, it diverges into two parallel streams leading to YS.Endo-3 and YS.Endo-4 (Figures 4G and 4K). Moreover, the distribution of cells across the phases of the cell cycle reveals the YS.Endo-3 exhibits a reduced ratio in the G2/M phase compared to others (Figure S4G).

We next proceeded to investigate the gene expression dynamics and functions of distinct Endo subclusters. Through pseudotime analysis, we identified the top 25 genes that underwent dynamic changes during human YS.Endo development. For instance, *TTN*, *FGFR1*, and *GATA3* exhibited a gradual down-regulation, while *APOB*, *GPC3*, and *CYP4B1* were up-regulated as YS.Endo developed (Figures 4L, 4M, S4H, and S4I). Further GO enrichment analysis unveiled functional distinctions among YS.Endo subtypes (Figure 4J). Specifically, Definitive.Endo was enriched in terms like gastrulation and Epi to mesenchymal transition, while Visceral.Endo showed enrichment in terms related to Epi tube morphogenesis and gland development. YS.Endo-1 and YS.Endo-2 exhibited enrichment in terms such as actin filament organization and PROTEIN transmembrane transport, suggesting remodeling of cellular functions. YS.Endo-3 displayed enrichments in terms like response to meta lion and sterol metabolic process. YS.Endo-4, as previously mentioned, exhibited high expression of several hematopoietic-related markers and was enriched in terms related to erythrocyte development and megakaryocyte development (Figure 4J).

### Development of extra-embryonic mesoderm

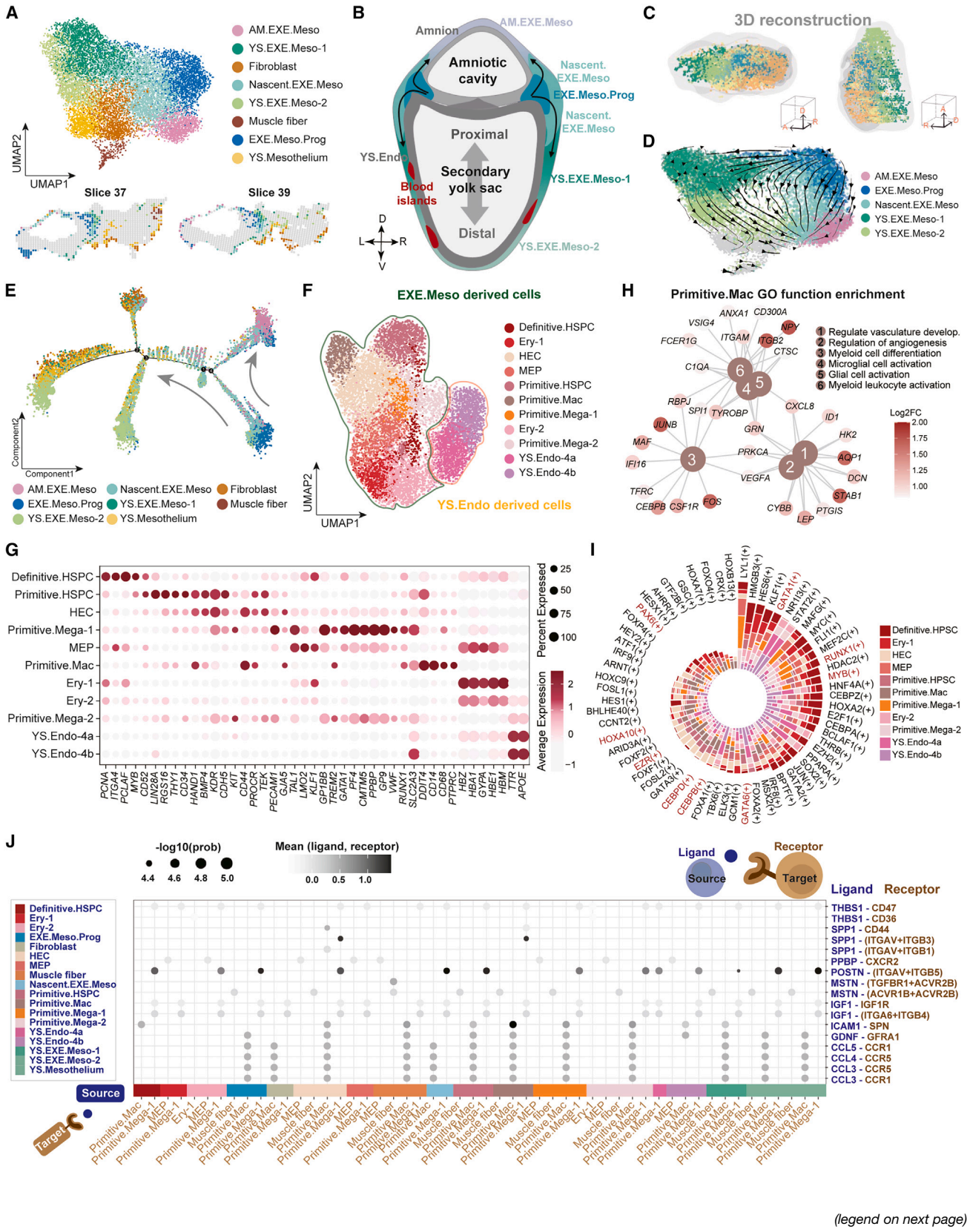
The EXE.Meso plays a role in early erythropoiesis and provides mechanical support to the developing embryo.[29,52] To understand EXE.Meso development, we sub-clustered the EXE.Meso-related cells, including AM.EXE.Meso (cluster 8 from Figure 1D), YS.EXE.Meso-A and YS.EXE.Meso-B (clusters 9 and 10 from Figure 1D), and EXE.Meso.Prog (from Figure 3B), yielding 8 subclusters (Figures 5A, 5B, S5A, and S5B). We further identified: EXE.Meso.Prog (*MSX1* and *COL6A3*), Nascent.EXE.Meso (*IGFBP5* and *IGFBP7*), AM.EXE.Meso (*MSX2* and *TEAD1*), YS.EXE.Meso-1 (*GNRH2* and *SPINK1*), YS.EXE.Meso-2 (*MT2A* and *A2M*), muscle fiber (*MYL9* and *ACTA2*), YS.Mesothelium (*PODXL* and *CDH1*) and fibroblast (*COL3A1* and *EGR1*) (Figure S5B). The EXE.Meso.Prog was located around the embryonic disc, while the Nascent.EXE.Meso was predominantly found in the proximal region (adjacent to the embryonic disc) of both amniotic and yolk sac cavities. The AM.EXE.Meso was specifically located above the amniotic cells. YS.EXE.Meso-1, YS.EXE.Meso-2, and YS.EXE.Meso-3 were sequentially distributed along the proximal to the distal end of the yolk sac (Figures 5A–5C). RNA velocity and monocle analysis revealed that the developmental trajectories start at EXE.Meso.Prog and then diverge into two streams, corresponding to AM.EXE.Meso or YS.EXE.Meso, respectively (Figures 5D and 5E).

### The specification of hemogenic cells in the yolk sac

Previous studies showed that primitive hematopoiesis in mice generates primitive nucleated erythrocytes and macrophages at around E7.5[4]. The transient definitive hematopoiesis commences when erythro-myeloid progenitors emerge from the yolk sac

**Figure 4. Spatial distribution of Ecto and Endo**

(A) UMAP plot showing 4 cell types from Epi and AM (left) with PAGA graph (top right) and RNA trajectories (bottom right) inferred from RNA velocity analysis.
(B) Spatial distribution of AM, AM.Ecto, Ecto, and Epi in the representative slices (slice 38 and slice 53).
(C) Density plot showing the indicated cell distributions along the A-P axis through cell ratio analysis.
(D) Dotplot depicting the normalized expression levels of representative genes in each cell type.
(E) IF staining results show the location of GABBR2, IGFBP3, NPNT, and TFAP2C in the human CS8 embryo. Scale bar, 50 μm.
(F) Scatterplots showing the expression dynamic of indicated developmental genes along the latent time inferred by RNA velocity analysis of AM and Ecto.
(G) UMAP plot showing the 6 subtypes of Endo (bottom) and RNA velocity analysis (top left) and PAGA path (top right) overlaid on the UMAP (top left).
(H) Schematic diagram showing the zonation of Endo types in the yolk sac.
(I) Spatial distribution of the 6 subtypes of Endo in the representative slice 40 (top) or 3D reconstructed model (bottom).
(J) Heatmap of the expression pattern of the top 50 genes of each cell type (left) with enriched GO terms (middle) and representative genes related to each GO term (right).
(K) Deconvolution of the pseudotime trajectory of (I) according to cell types.
(L) Heatmap showing the scaled expression pattern of top 25 genes identified along the latent time of YS.Endo.
(M) Scatterplots showing the expression dynamic of indicated developmental genes (top 50 significant genes) along the latent time from (G).




(legend on next page)

haemogenic endothelium at E8.25[4]. However, studying stage-matched human blood and immune cells is challenging. We re-clustered the EXE.Meso-A and EXE.Meso-B, Ery, HEP (clusters 9, 10, 12, and 13, respectively annotated in Figure 1D), and YS.Endo that displayed high expression of *TTR*, *APOE*, and hemoglobin genes such as *HBZ*, *HBQ1*, and *HBA1*. Eleven subclusters were identified based on the expression of known markers, including primitive hematopoietic stem/progenitor cells (Primitive.HSPC; *LIN28A*, *THY1*, and *HAND1*), definitive hematopoietic stem/progenitor cells (Definitive.HSPCs; *ITGA4*, *PCLAF*, and *CD52*), hemogenic endothelial cells (HECs; *BMP4* and *TEK*), megakaryocyte-erythroid progenitors (MEPs; *TAL1*, *LMO2*, and *KLF1*), primitive macrophage (Primitive.Mac; *CD14*, *CD68*, and *HAND1*), primitive megakaryocyte-1 (Primitive.Mega-1; *PF4*, *PPBP*, *CMTM5*, and *VWF*), primitive megakaryocyte-2 (Primitive.Mega-2; *PF4* $^{low}$, *PPBP* $^{low}$, *CMTM5* $^{low}$, *GP9* $^{low}$, and *VWF* $^{low}$), erythroid-1 (Ery-1; *HBE1*, *GYPA*, and *HBZ* ), erythroid-2 (Ery-2; *HBE1* $^{low}$, *HBM* $^{low}$, and *HBZ* $^{low}$), YS.Endo-4a and 4b (*HBE1* $^{low}$, *HBM* $^{low}$, *TTR*, and *APOE*) (Figures 5F and 5G). Distinct developmental trajectories from YS.Endo-4a into Ery-2 and YS.Endo-4b into Primitive.Mega-2 were revealed by RNA velocity analysis (Figure S5F).

Utilizing early and late HSPCs markers, we were able to distinguish between primitive and definitive.HSPCs fractions.[53] Definitive.HSPCs demonstrate a higher expression of cell cycle genes (Figure 5G). Primitive.HSPCs specifically expressed *DDIT4*, *LIN28A*, *RGS16*, and *TEK* with enriched GO terms associated with epithelial cell proliferation and regulation of RNA splicing (Figures 5G and S5E). Compared to Primitive.HSPCs, Definitive.HSPCs express highly critical genes for erythrocyte development (*GATA1* and *KLF1*), as well as *HBZ*, *CYPA*, and *HBM*, suggesting their potential for erythroid differentiation. *CD44* has been reported as a marker to distinguish arterial endothelial cells from other endothelial cells.[54] The expression of *CD44*, *GJA4*, and *GJA5* in the HEC cluster, along with the co-expression of *PROCR*, *MYB*, *CD34*, and *KIT*, signified the characteristic arterial endothelial and hematopoietic features of HECs in the yolk sac. GO analysis showed that HECs were enriched with terms associated with extracellular matrix organization, regulation of vascular development, and aorta development (Figures 5G and S5D). MEP was annotated with megakaryocytes and erythrocytes-related genes, which were enriched in GO terms such as platelet activation and erythrocyte differentiation (Figure S5D).

The GO terms associated with primitive megakaryocytes-1 and megakaryocytes-2 were associated with platelet activation and aggregation, while Ery-1 and Ery-2 were related to oxygen and gas transport and homeostasis (Figure S5D). Previous studies suggest that primitive macrophages, independent of monocytes, may serve as the main source of microglia in both mice and humans.[55,56] Consistently, primitive macrophages were identified based on expression markers of myeloid (*CD68* and *CD14*), endothelium (*CDH5* and *KDR*), pan-hematopoietic cells (*PTPRC*), and microglia (*TREM*). GO analysis revealed that the primitive macrophages were enriched in terms related to vascular development, glial and microglial cell activation (Figure 5H), further supporting their potential transition into microglia. Pearson correlation analysis revealed that Primitive.HSPCs exhibited the highest similarity with the Nascent.EXE.Meso and YS.EXE.Meso-2 clusters (Figure S5C). Furthermore, Definitive.HSPC clusters exhibited a higher correlation with HEC (Figure S5C), highlighting the potential emergence of definitive.HSPC from HEC during the transient definitive hematopoiesis.

## Dissection of key regulators for hemogenic cells and their communication with other yolk sac cells

To identify key regulators during hematogenic cell diversification, we performed the SCENIC analysis. *HOXA10* was significantly enriched in both primitive HSPCs and definitive HSPCs. *LYL1*, known to play a significant role in platelet production and function,[57] was significantly enriched in primitive megakaryocyte clusters (Figure 5I).

Next, we sought to identify the ligand-receptor interactions involved in primitive macrophage and their surrounding cells. We found that primitive macrophages interact with multiple cells in the yolk sac through ligand-receptor pairs such as *ICAM1-SPN*, *GDNF-GFRA1*, and *CCL3,4,5-CCR1/5* (Figure 5J). *CCL5*, through binding *CCR1/5*, is known to contribute to the transendothelial migration of monocytes in atherogenic lesions.[58] The interaction of *CCL5* with *CCR1* between primitive macrophages and HECs may play a role in promoting primitive macrophages migration into the embryo head. Furthermore, we observed that *SPP1*, the transcriptional target of *MAF*, is expressed in HEC, and its receptor genes *ITGAV* and *ITGB1/2* are upregulated in primitive megakaryocytes and primitive macrophages. Additionally, *POSTN* is enriched in primitive megakaryocytes, consistent with the role of *POSTN* in promoting the secretion of extracellular matrix proteins involved in tissue regeneration and

**Figure 5. Distinct developmental trajectories of EXE.Meso and hematopoietic cells**

(A) Schematic diagram indicating the development of EXE.Meso in CS8 human embryo.

(B) Top: UMAP plot showing 8 major EXE.Meso subtypes in the human CS8 embryo. Bottom: the spatial distribution patterns of the EXE.Meso-related clusters in the representative slices (37 and 39).

(C) The spatial distribution patterns of the EXE.Meso-related clusters in the 3D reconstructed model.

(D) RNA velocity analysis overlaid on the UMAP plot of major EXE.Meso cell types.

(E) Pseudotime analysis of EXE.Meso.Prog and its derivatives along two independent trajectories. The points are colored by cell types.

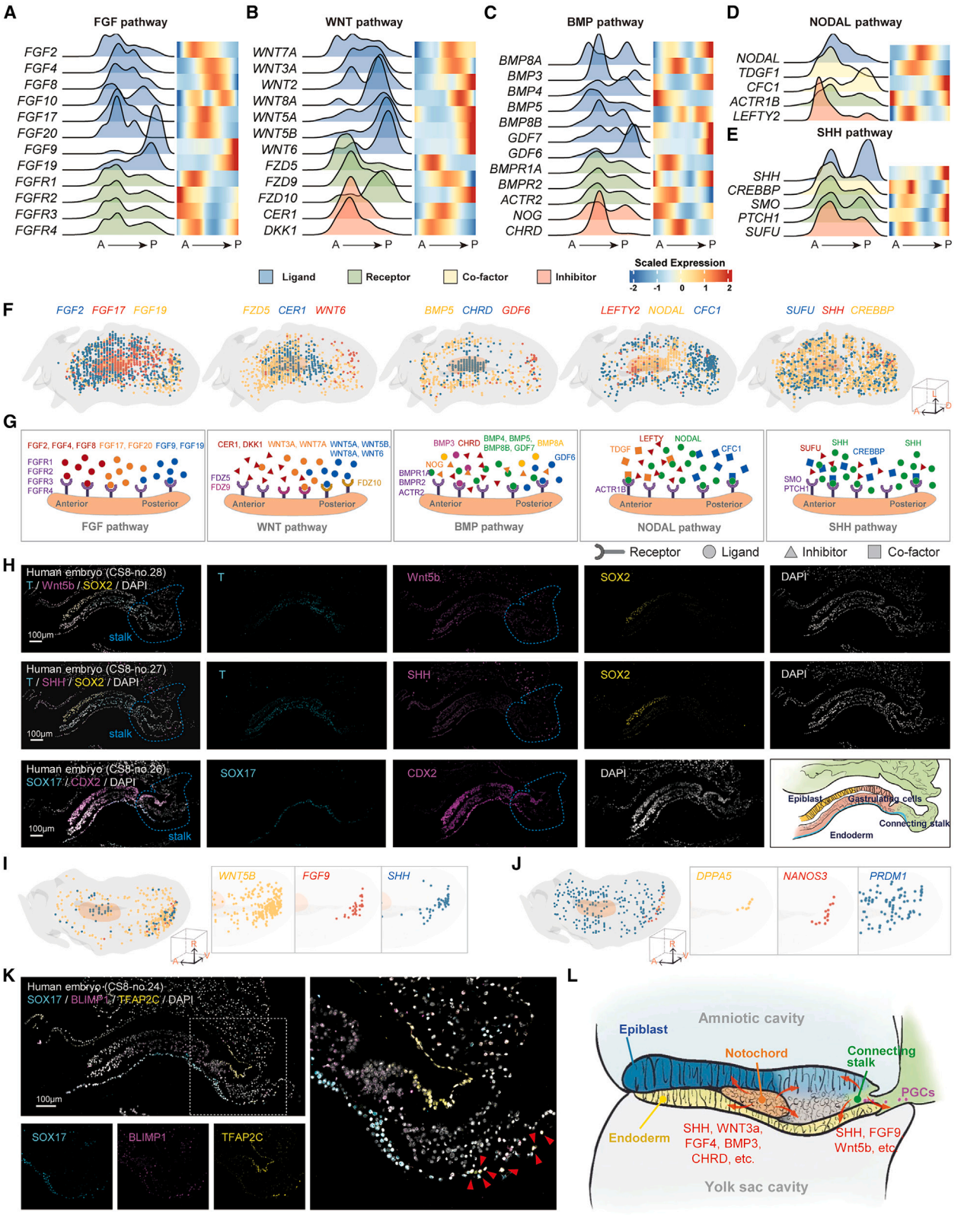
(F) UMAP plot showing the 11 major hematogenesis-related cell types in the human CS8 embryo.

(G) Dotplot depicting the normalized expression levels of representative marker genes within the distinct hematopoietic clusters.

(H) Functional networks of the GO enrichment results. Nodes represent GO terms and DEGs. Node color: fold change of DEGs; edge width: adjusted *p* value. The color bars from grey to red indicate the log2FC from low to high.

(I) Circular stacked barplot showing the enriched core regulatory TF regulons of each hematopoietic cell type in the human CS8 embryo.

(J) Bubble plot showing the selected ligand-receptor interactions. The means of the average expression level of interacting receptors in target cells and ligands in source cells are indicated by gray value.




*(legend on next page)*

wound healing mediated by megakaryocyte-derived platelets (Figure S5G).

### Spatial expression patterns of key signaling pathways critical for early human embryo development

How the morphogens pattern the gastrulating human embryos remains poorly understood. Given that Noto is a well-recognized signaling center, we conducted cell-cell communication analysis (Figures S6A, S6B, and S6E) to identify the cell types that display strong communications ($p < 0.05$) between Noto and other embryonic cell types. Combing the cell-cell communication results and existing knowledge,[59–70] we identified a list of key pathways that show dynamic expression in different cell types, such as FGF, WNT, NODAL, BMP, and SHH signaling pathways (Figures 6A–6E).

The FGF signaling is a large family consisting of 22 identified FGF members and 4 FGF receptors (FGFRs).[71,72] Different FGF members demonstrated different distribution and expression patterns along the A-P axis and were mainly distributed at the front and middle part of the embryo (e.g., *FGF2*, *FGF4*, and *FGF17*), and all 4 FGFRs were highly expressed in the anterior part (Figure 6A). Two exceptions, *FGF9* and *FGF19*, were exclusively expressed at the posterior end (Figures 6A, 6F, and 6G). Moreover, different FGF members demonstrated distinct expression patterns among different cell types, for instance, *FGF1/2* in Epi/Ecto, *FGF5/13* in Definitive.Endo, *FGF10/19* in LP.Meso, *FGF9/18* in connecting stalk, and a large number of FGFs (e.g., *FGF11*, *16*, *17*, *20*) in Noto (Figure S6A).

The majority of WNT ligands[73] (e.g., *WNT5B*, *WNT6*, and *WNT8A*) were specifically expressed at the posterior end, except for *WNT3A* and *WNT2*, which were distributed at the middle and anterior parts (Figure 6B). Consistently, WNT antagonists *DKK1* and *CER1* were predominantly expressed at the anterior end (Figures 6B, 6F, and 6G), which agrees with the prior mouse work that these antagonists may guide the formation of the A-P axis.[74–77] We also found the receptors FZDs demonstrated spatial-specific expression patterns (e.g., *FZD5*, *FZD9*, and *FZD10* in the anterior, middle, and posterior parts, respectively) (Figures 6B, 6F, and 6G). By calculating the pathway activity score, we also found the WNT pathway demonstrated a lower score in the anterior region than in the posterior region (Figure S6D).

Although TGFβ signaling regulates diverse biological processes during development,[78] we did not observe significant differences in their distribution and expression patterns along the A-P axis (Figure S6C). However, we found distinct spatial patterns of BMP in the A-P axis distribution. For instance, we found *BMP3* in the anterior region, *BMP6* in both the anterior and middle regions, *BMP4* in the middle and posterior regions, *GDF6* exclusively in the posterior region, and *BMP8A* in both the anterior and posterior regions but absent from the middle region. In contrast, BMP receptors and inhibitors (e.g., *NOG*) exhibited nearly uniform expression along the A-P axis (Figures 6C, 6F, and 6G). The pathway activity scoring analysis indicates the BMP pathway showed higher activity in the anterior region (Figure S6D).

The *NODAL* is highly expressed in the anterior and middle regions, while its co-factors such as *TDGF1*[79] and *CFC1*[80] were distinctively expressed in the anterior and posterior parts of the embryo (Figures 6D, 6F, and 6G). Besides, the receptors (e.g., *ACTR1B*) and inhibitors (e.g., *LEFTY2*) of NODAL were primarily expressed in the anterior region.

Retinoic acid (RA) is also an essential signaling for embryonic development. We found the RA synthetases ALDH[81] demonstrated a sequential distribution, from anterior to posterior end: *ALDH1A1*, *ALDH1A3*, and *ALDH1A2*, respectively (Figure S6C). The RA receptor *RARB*, displayed even expression patterns, while *CYP26A1*, which is responsible for metabolizing RA,[82] showed expression in the anterior and middle regions (Figure S6C). The pathway activity score suggested RA pathway was evenly activated along the A-P axis (Figure S6D).

Another pivotal morphogenic signaling pathway is SHH.[83] We unveiled two distinct expression peaks of *SHH*, one in the anterior and middle regions and another in the posterior region. Notably, the expression level in the posterior region surpassed that in the anterior region (Figures 6E–6G). To our best knowledge, *SHH* expression was not previously reported to be expressed in the posterior end of early primate embryos, while only a previous study conducted in mice had reported the observation of SHH distribution in the posterior end of ∼E7.0 to E9.75 mouse gastrulating embryos.[84] Likewise, the expression pattern of SHH receptor *PTCH1* and SHH co-factors *CREBBP* also exhibited a two-peak behavior (Figures 6E–6G). The pathway activity score implies SHH pathway has two high activity sites at the middle and posterior regions (Figure S6D).

Additionally, extending beyond the A-P axis, we also performed a similar analysis for the dorsal-ventral (D-V) axis and left-right (L-R) axis. The outcomes indicated elevated FGF

**Figure 6. Spatial expression patterns of key signaling pathways critical for early human embryo development**

(A–E) Ridgeline plots of gene distributions (left) and heatmap of gene expressions (right) showing dynamic expression patterns of ligands, receptors, co-factors, inhibitors, synthesis enzymes, and metabolic enzymes from FGF (A), WNT (B), BMP (C), NODAL (D), and SHH (E) pathways along the A-P axis.

(F) Spatial expression patterns of key representative genes from (A–E) in the 3D reconstructed CS8 human embryo. Spots with gene expression levels greater than 0 are highlighted in corresponding colors for display.

(G) Schematic summary of the interactions among key components of FGF, WNT, BMP, NODAL, or SHH signaling pathways along the A-P axis of the CS8 human embryo.

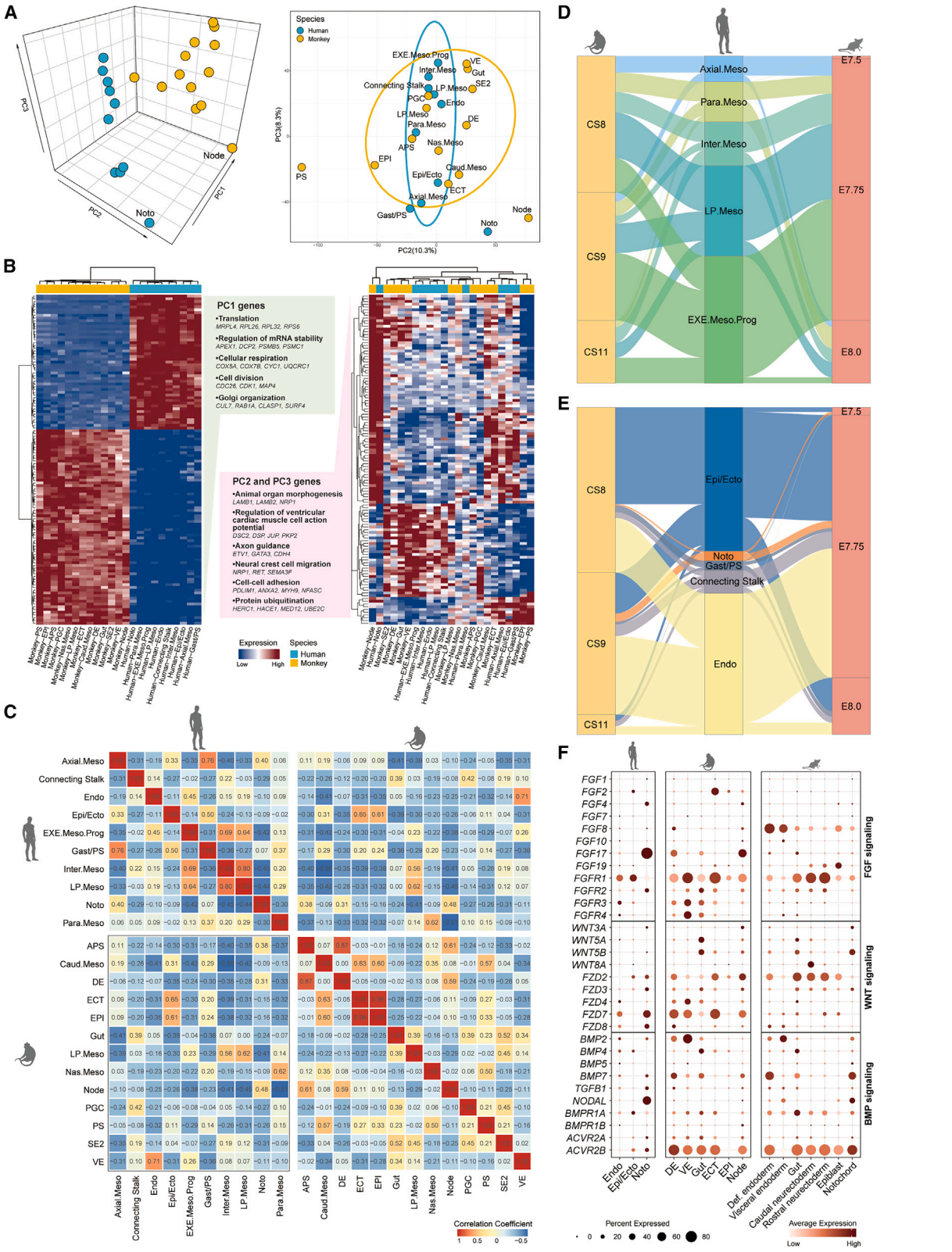
(H) Representative IF staining results showing localization of T+, WNT5B+, SOX2+, SHH+, SOX17+, or CDX2+ cells in a truncated CS8 embryo. The schematic diagram (bottom right) summarizes the distribution of indicated cell types within the region of interest, based on the IF images above.

(I) 3D spatial expression patterns of *WNT5B*, *FGF9*, and *SHH*. Spots with gene expression levels greater than 0 are highlighted in corresponding colors for display.

(J) 3D spatial expression patterns of *DPPA5*, *NANOS3*, and *PRDM1*. Spots with gene expression levels greater than 0 are highlighted in corresponding colors for display.

(K) Representative IF staining showing the localization of SOX17+, BLIMP1+, and TFAP2C+ cells (PGCs) in a human CS8 embryo.

(L) Schematic diagram summarizing key findings.




*(legend on next page)*

pathway activity on the dorsal side (Epi/Ecto), increased BMP and WNT pathway activity on the ventral side (Endo) and heightened RA, NODAL, SHH, and TGF pathway activities in the middle region (Meso) (Figure S6D). Regarding the L-R axis, except for a slightly higher activity score for the NODAL pathway on the left side (divided by the Noto), we did not observe significant differences in the activity scores of other pathways along the L-R axis (Figure S6D).

## Identification of a potential signaling center at the posterior region of the CS8 human embryo

We then investigated the specific cell types responsible for the expression of these morphogens. Regarding *SHH*, the main source is the Noto in the anterior and middle regions of the embryo (Figures 6E, S6A, S6B, and S6D). In the posterior region, the primary source of SHH is the connecting stalk, defined by the expression of *CDX1* and *CDX2*[29,42] (Figure 3C). In addition to *SHH*, we also identified that Noto express signaling factors such as *CHRD*, *BMP2*, and *WNT3A*, while connecting stalk cells express signaling factors such as *WNT5B* and *MDK* (Figures S6A, S6E, and S6F). For validation, we performed immunostaining using a separate truncated CS8 human embryo, confirming the presence of SHH and WNT5B signals at the posterior site (Figures 6H and 6I). A previous study has reported the robust expression of *SHH* and *SMO* in the posterior region of the mice embryos.[84] Collectively, it implied that the connecting stalk may serve as a signal-secreting center at the posterior region of the human early embryo (Figure 6L).

At the connecting stalk site, we identified several spots that express the primordial germ cell (PGC) markers[28] (e.g., *NANOS3*, *BLIMP1*, and *PRDM1*) (Figures 6J and 6K), which aligns with a previous study that observed the presence of monkey PGCs in the posterior embryonic disc of *in vitro* cultured (day 22) monkey embryos.[17] In primates, the origination of PGCs, from Epi or AM, is still controversial.[85–90] To understand the potential origin of PGCs, we projected the PGCs (*NANOS3*+) alongside Epi, AM, and AM.Ecto and observed the PGCs were closely clustered together (Figures S6G and S6H). Sub-clustering further categorized the epiblast into early (Epi-1) and late (Epi-2 and Epi-3) epiblast. Partition-based graph abstraction (PAGA) trajectory analysis suggested a common progenitor shared by PGCs and AM.Ecto (Figures S6G and S6I), consistent with a previous study.[91]

## Cross-species comparison of gastrulation

To understand the coordination and differentiation of lineages at key milestones during gastrulation across species, we compared our human CS8 embryo with published datasets of monkey and mouse gastrula. We started by comparing with cynomolgus monkey at the corresponding CS8 stage.[29] PCA revealed that major embryonic cell types from human and monkey exhibited distinct separation along PC1 (Figure 7A), which represents the major difference between species. We ranked the contribution of genes to PC1 and representative GO terms indicating the species-specific characteristics including 'Translation', 'Regulation of mRNA stability', and 'Cell division' were enriched (Figure 7B). The interlaced plotting of cell types from both species along PC2 and PC3 coordinates reflected their developmental hierarchy and cell-type diversification (Figure 7A). We determined genes of the major contribution to PC2 and PC3, which were associated with organogenesis, such as 'Animal organ morphogenesis', 'Regulation of ventricular cardiac muscle cell action potential', and 'Neural crest cell migration' (Figure 7B). A correlation analysis revealed that Endo, Epi/Ecto, Axial.Meso, LP.Meso, and Noto from humans exhibited the highest correlation with monkey Visceral.Endo, Epi/Ecto, Caud.Meso, LP.Meso and Node, respectively (Figure 7C). Thus, we identified conserved and divergent cellular composition, as well as molecular architecture between human and monkey CS8 gastrula.

Based on gene expression and spatial distribution, the Meso clusters exhibit a more diverse and sequential developmental trajectory across different species. Through further cross-species comparison using Meso of human and monkey CS8 gastrula, we noticed the comparative projection of cell types displayed in UMAP (Figure S7A). In addition, due to the spatially resolved resolution, our data reflects an accurate definition of Meso and reasonably monitors the heterogeneity of those from monkeys. Thus, a comparative analysis of the transcriptomic data from human and monkey CS8 embryos describes a conserved distribution of major cell types with approximate developmental progression between the two species.

Next, further cross-species analysis confirmed the comparative developmental stages among human CS8 and other species.[4] Specifically, reference mapping analysis showed that embryonic cell types of human CS8 are mainly allocated by monkey CS8 and mouse E7.75, and then by monkey CS9 and mouse E8.0 (Figures 7D and 7E). Next, we investigated the conservation and evolution of each germ layer. As expected, humans shared larger numbers of marker genes for Endo, Meso, and Ecto with monkeys than with mice (Figures S7B–S7D). Moreover, we observed consistent functional features among different germ layers across species, including ''digestive system development'' for Endo, ''heart development'' for Meso, and ''regulation of neuron differentiation'' for Ecto (Figure S7E). Upon

**Figure 7. Cross-species comparison of the gastrula from humans, monkeys, and mice**

(A) PCA plots of human and monkey CS8 embryo cells by one-to-one orthologs between the two species (12,152 genes). The color coding of species is as indicated.

(B) Heatmap of the expression of top 200 genes that contributed highly to the PC1 and PC2/PC3, and the associated GO terms.

(C) Heatmap of correlation coefficients among cell types from human and monkey CS8 embryo. The correlation coefficients were calculated using the average expression level of differentially expressed genes (DEGs) of each cluster (log2fc_threshold = 1.25). The black box highlighted the correlation coefficients between cell types from human and monkey embryos.

(D) Sankey diagram showing the corresponding relationships of Meso between human CS8 embryo and the developmental stages of monkeys and mice.

(E) Sankey diagram showing the corresponding relationships of Endo and Epi/Ecto between human CS8 embryo and the developmental stages of monkeys and mice.

(F) Dot plot showing the scaled expression of key components of the FGF, WNT, and BMP signaling pathways enriched in Epi and Ecto cells across species.

conducting a comprehensive mapping of Meso in the CS8, we observed notable correspondence between these cells and various stages and types annotated in monkeys and mice. This finding confirmed that LP.Meso, Inter.Meso, and EXE.Meso.Prog populations, which give rise to pharyngeal Meso represent more advanced states of mesoderm development (Figure S7F). Furthermore, the monkey Meso exhibit a more similar timing tempo with humans in comparison to that of mice.

Moreover, genes involved in FGF signaling, WNT signaling, and BMP signaling exhibited broad similarities in the Endo and Ecto of all three species. Certain specific differences were embedded in the high expression of *FGF2, FGF17, FGFR3*, and *FZD7* in primates (Figure 7F). Therefore, comparative analysis achieves a high consensus transcriptomic organization that is preserved among primates. Combined with mouse gastrula, the transcriptional similarity of cell types across species varies based on the evolutionary distance.

## DISCUSSION

Gaining insights into the post-implantation development of humans has been challenging. Recently, several stem-cell-derived human embryo models tried to provide an alternative resource to study human early development.[18–27] However, how to build and evaluate these embryo models is highly difficult due to the lack of natural embryo references. To fill this gap, we obtained an intact gastrulating human CS8 embryo, performed serial transverse sectioning, and generated an in-depth and high-resolution spatial transcriptomic dataset of a human CS8 embryo. By combining the spatial transcriptomes from all slices, we created a comprehensive 3D reconstruction of this CS8 embryo, along with a website (cs8.3dembryo.com) for an interactive survey of the results. This complete 3D transcriptome dataset offers opportunities, not only to determine the critical cellular and molecular features of key gastrulating events in humans but also to guide the generation of stem-cell-derived human embryo models.

In the conventional scRNA-seq data analysis processes, cell types were annotated based only on a limited set of known lineage-specific markers. This poses a particular challenge in early human embryos due to the lesser number of known genes, and variations in gene expression compared to other species. Integrating the spatial information of individual cells with their gene expression profiles allows for more accurate annotation of diverse cell subtypes. Here, we have demonstrated several advantages, including (1) the various subtypes of Meso, such as Axial.Meso, Para.Meso, Inter.Meso, LP.Meso, and EXE.Meso, could be identified and arranged sequentially along the A-P axis of the PS during gastrulation; (2) a type of cell was found located at the boundary between the AM and the embryonic disc, and both gene expression profiling and trajectory analysis indicate that they are in the process of differentiation from the Epi into amniotic cells, which indicates they exist as an intermediated cell type between Epi and amniotic cells; (3) the developmental process of YS.Endo can be divided into distinct stages, and they are arranged in a sequence from proximal to distal along the amniotic cavity, aligning with their migration direction; and (4) the EXE.Meso.Prog (originating from the Epi) migrate out from the embryo and follow two distinct upper and lower directions away from the embryonic disc, and subsequently give rise to the AM.EXE.Meso and the YS.EXE.Meso.

In the developing embryo, cells maintain continuous communication with neighboring cells, through inter-cellular signaling activity. The cells that produce morphogens serve as signaling organizers orchestrating the axial patterns and lineage specifications. Therefore, to gain insight into early human development and create stem-cell-derived embryo models, a fundamental focus lies in understanding the distribution of these morphogens and organizers. In this study, we investigated 3D spatial expression patterns of related ligands, receptors, inhibitors, and co-factors of the key signaling pathways involved in early development. The results suggest the different signaling pathways employ distinct strategies to establish differential activation along the embryo A-P axis. Furthermore, by analyzing the 3D spatial distributions of *SHH*, *FGF9*, and *WNT5B*, in line with a previous study conducted in mice,[84] we identified the connecting stalk as a potential signal-secreting center involved in the regulation of developmental processes in the posterior end of human embryos.

### Limitations of the study

Although image-based Stereo-seq allows for high-resolution exploration of spatial expression in samples, the commonly used Bin 50 (50 × 50 DNBs),[32] typically represents an expression matrix of approximately 1–3 cells. In instances where multiple cell types are nearby, the accuracy of cell segmentation and expression patterns may be compromised. Moreover, the challenge of obtaining precious donated embryos at such early stages restricts the availability of samples for extensive experimental validation and biological replicates. Although the stem-cell-based embryo models represent an optimal approach for validation, existing human embryo models only partially replicate early-stage human embryos, such as the pre-implantation blastocyst stage[18,19,24,25] or the very early peri-implantation stage.[20–22,27] A *bona fide in vitro* human embryo model representing CS8 or earlier stage is currently lacking. For cross-species comparison, the existing analysis approaches are challenging in considering the cases where different gene family members are deployed within conserved pathways.

## STAR★METHODS

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

- Cell cycle analysis
- Cluster annotation
- 3D modeling
- 3D gene expression patterns
- Cell type mapping and datasets integrated
- SCENIC analysis
- Hotspot analysis
- Functional enrichment analysis
- Cell-cell communication analysis
- Pseudotime analysis
- RNA velocity analysis
- Mfuzz analysis
- Cross-species comparative analysis
- Immunofluorescent staining

- QUANTIFICATION AND STATISTICAL ANALYSIS

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2024.03.041.

## ACKNOWLEDGMENTS

We would like to thank Dr. Qi Zhou for invaluable support and guidance. This work was supported by the National Key Research and Development Program of China (2022YFA1103100 and 2022YFA1104300 to L.Y.; 2022YFA1104100 and 2022YFC2702600 to J.G.); the start-up funding support from the Institute of Zoology and Chinese Academy of Sciences to L.Y. and J.G.; the Innovative Project of State Key Laboratory of Animal Biotech Breeding (2023SKLAB1-2), Pinduoduo-China Agricultural University Research Fund (PC2023B02016), STI 2030-Major Project (2023ZD0407504) to Y.W.; the National Natural Science Foundation Of China (32370871 to L.Y.; 82101764 to N.H.; 2022CX11015 and 82371685 to Z.X.). We acknowledge all the members of the Gynecologic Hysteroscopy Operating Room and the assistance with access to analytic instruments from the Translational Medicine Center at The First Affiliated Hospital of Zhengzhou University. We thank all the members from Yu Lab, Guo Lab, and Wei Lab for their discussions.

## AUTHOR CONTRIBUTIONS

L.Y. and J.G. conceptualized the idea. H.W. oversaw the overall project. L.Y., J.G., Z.X., and Y.W. designed, interpreted, and supervised the experiments. N.H. performed the human embryo sample collection. Z.X., Y.W., N.H., X.Z., Z.W., P.S., and Y.L. performed Stereo-seq library construction and sequencing. Y.Y., L.C., X.W., S.L., and X.Y. performed bioinformatics analyses. L.C. and X.W. constructed the online website. X.X. and N.H. performed immunostaining and confocal photography. L.Y., Z.X., J.G., and Y.W. performed manuscript writing, review, and editing with input and feedback from all authors.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 18, 2023
Revised: January 17, 2024
Accepted: March 26, 2024
Published: April 23, 2024

## REFERENCES

1. Muhr, J., Arbor, T.C., and Ackerman, K.M. (2023). Embryology, Gastrulation (StatPearls Publishing).
2. Anderson, C., and Stern, C.D. (2016). Chapter twenty-six organizers in development. Curr. Top. Dev. Biol. *117*, 435–454. https://doi.org/10.1016/bs.ctdb.2015.11.023.
3. Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F.J., et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. Nature *566*, 496–502. https://doi.org/10.1038/s41586-019-0969-x.
4. Pijuan-Sala, B., Griffiths, J.A., Guibentif, C., Hiscock, T.W., Jawaid, W., Calero-Nieto, F.J., Mulas, C., Ibarra-Soria, X., Tyser, R.C.V., Ho, D.L.L., et al. (2019). A single-cell molecular map of mouse gastrulation and early organogenesis. Nature *566*, 490–495. https://doi.org/10.1038/s41586-019-0933-9.
5. Argelaguet, R., Clark, S.J., Mohammed, H., Stapel, L.C., Krueger, C., Kapourani, C.-A., Imaz-Rosshandler, I., Lohoff, T., Xiang, Y., Hanna, C.W., et al. (2019). Multi-omics profiling of mouse gastrulation at single-cell resolution. Nature *576*, 487–491. https://doi.org/10.1038/s41586-019-1825-8.
6. Mohammed, H., Hernando-Herraez, I., Savino, A., Scialdone, A., Macaulay, I., Mulas, C., Chandra, T., Voet, T., Dean, W., Nichols, J., et al. (2017). Single-cell landscape of transcriptional heterogeneity and cell fate decisions during mouse early gastrulation. Cell Rep. *20*, 1215–1228. https://doi.org/10.1016/j.celrep.2017.07.009.
7. He, P., Williams, B.A., Trout, D., Marinov, G.K., Amrhein, H., Berghella, L., Goh, S.-T., Plajzer-Frick, I., Afzal, V., Pennacchio, L.A., et al. (2020). The changing mouse embryo transcriptome at whole tissue and single-cell resolution. Nature *583*, 760–767. https://doi.org/10.1038/s41586-020-2536-x.
8. Bouchereau, W., Jouneau, L., Archilla, C., Aksoy, I., Moulin, A., Daniel, N., Peynot, N., Calderari, S., Joly, T., Godet, M., et al. (2022). Major transcriptomic, epigenetic and metabolic changes underly the pluripotency continuum in rabbit preimplantation embryos. Development *149*, dev200538. https://doi.org/10.1242/dev.200538.
9. Mayshar, Y., Raz, O., Cheng, S., Ben-Yair, R., Hadas, R., Reines, N., Mittnenzweig, M., Ben-Kiki, O., Lifshitz, A., Tanay, A., and Stelzer, Y. (2023). Time-aligned hourglass gastrulation models in rabbit and mouse. Cell *186*, 2610–2627.e18. https://doi.org/10.1016/j.cell.2023.04.037.
10. Gong, Y., Bai, B., Sun, N., Ci, B., Shao, H., Zhang, T., Yao, H., Zhang, Y., Niu, Y., Liu, L., et al. (2023). Ex utero monkey embryogenesis from blastocyst to early organogenesis. Cell *186*, 2092–2110.e23. https://doi.org/10.1016/j.cell.2023.04.020.
11. Niu, Y., Sun, N., Li, C., Lei, Y., Huang, Z., Wu, J., Si, C., Dai, X., Liu, C., Wei, J., et al. (2019). Dissecting primate early post-implantation development using long-term in vitro embryo culture. Science *366*, eaaw5754. https://doi.org/10.1126/science.aaw5754.
12. Ma, H., Zhai, J., Wan, H., Jiang, X., Wang, X., Wang, L., Xiang, Y., He, X., Zhao, Z.-A., Zhao, B., et al. (2019). In vitro culture of cynomolgus monkey embryos beyond early gastrulation. Science *366*, eaax7890. https://doi.org/10.1126/science.aax7890.
13. Shahbazi, M.N., Jedrusik, A., Vuoristo, S., Recher, G., Hupalowska, A., Bolton, V., Fogarty, N.N.M., Campbell, A., Devito, L., Ilic, D., et al. (2016). Self-organization of the human embryo in the absence of maternal tissues. Nat. Cell Biol. *18*, 700–708. https://doi.org/10.1038/ncb3347.
14. Zhou, F., Wang, R., Yuan, P., Ren, Y., Mao, Y., Li, R., Lian, Y., Li, J., Wen, L., Yan, L., et al. (2019). Reconstituting the transcriptome and DNA methylome landscapes of human implantation. Nature *572*, 660–664. https://doi.org/10.1038/s41586-019-1500-0.
15. Xiang, L., Yin, Y., Zheng, Y., Ma, Y., Li, Y., Zhao, Z., Guo, J., Ai, Z., Niu, Y., Duan, K., et al. (2020). A developmental landscape of 3D-cultured human pre-gastrulation embryos. Nature *577*, 537–542. https://doi.org/10.1038/s41586-019-1875-y.
16. Deglincerti, A., Croft, G.F., Pietila, L.N., Zernicka-Goetz, M., Siggia, E.D., and Brivanlou, A.H. (2016). Self-organization of the in vitro attached human embryo. Nature *533*, 251–254. https://doi.org/10.1038/nature17948.
17. Zhai, J., Xu, Y., Wan, H., Yan, R., Guo, J., Skory, R., Yan, L., Wu, X., Sun, F., Chen, G., et al. (2023). Neurulation of the cynomolgus monkey embryo achieved from 3D blastocyst culture. Cell *186*, 2078–2091.e18. https://doi.org/10.1016/j.cell.2023.04.019.

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

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

20. Weatherbee, B.A.T., Gantner, C.W., Iwamoto-Stohl, L.K., Daza, R.M., Hamazaki, N., Shendure, J., and Zernicka-Goetz, M. (2023). Pluripotent stem cell-derived model of the post-implantation human embryo. Nature *622*, 584–593. https://doi.org/10.1038/s41586-023-06368-y.

21. Liu, L., Oura, S., Markham, Z., Hamilton, J.N., Skory, R.M., Li, L., Sakurai, M., Wang, L., Pinzon-Arteaga, C.A., Plachta, N., et al. (2023). Modeling post-implantation stages of human development into early organogenesis with stem-cell-derived peri-gastruloids. Cell *186*, 3776–3792.e16. https://doi.org/10.1016/j.cell.2023.07.018.

22. Oldak, B., Wildschutz, E., Bondarenko, V., Comar, M.-Y., Zhao, C., Aguilera-Castrejon, A., Tarazi, S., Viukov, S., Pham, T.X.A., Ashouokhi, S., et al. (2023). Complete human day 14 post-implantation embryo models from naive ES cells. Nature *622*, 562–573. https://doi.org/10.1038/s41586-023-06604-5.

23. Pedroza, M., Gassaloglu, S.I., Dias, N., Zhong, L., Hou, T.-C.J., Kretzmer, H., Smith, Z.D., and Sozen, B. (2023). Self-patterning of human stem cells into post-implantation lineages. Nature *622*, 574–583. https://doi.org/10.1038/s41586-023-06354-4.

24. Yu, L., Logsdon, D., Pinzon-Arteaga, C.A., Duan, J., Ezashi, T., Wei, Y., Ribeiro Orsi, A.E., Oura, S., Liu, L., Wang, L., et al. (2023). Large-scale production of human blastoids amenable to modeling blastocyst development and maternal-fetal cross talk. Cell Stem Cell *30*, 1246–1261.e9. https://doi.org/10.1016/j.stem.2023.08.002.

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

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

27. Okubo, T., Rivron, N., Kabata, M., Masaki, H., Kishimoto, K., Semi, K., Nakajima-Koyama, M., Kunitomi, H., Kaswandy, B., Sato, H., et al. (2024). Hypoblast from human pluripotent stem cells regulates epiblast development. Nature *626*, 357–366. https://doi.org/10.1038/s41586-023-06871-2.

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

29. Zhai, J., Guo, J., Wan, H., Qi, L., Liu, L., Xiao, Z., Yan, L., Schmitz, D.A., Xu, Y., Yu, D., et al. (2022). Primate gastrulation and early organogenesis at single-cell resolution. Nature *612*, 732–738. https://doi.org/10.1038/s41586-022-05526-y.

30. Nakamura, T., Okamoto, I., Sasaki, K., Yabuta, Y., Iwatani, C., Tsuchiya, H., Seita, Y., Nakamura, S., Yamamoto, T., and Saitou, M. (2016). A developmental coordinate of pluripotency among mice, monkeys and humans. Nature *537*, 57–62. https://doi.org/10.1038/nature19096.

31. Bergmann, S., Penfold, C.A., Slatery, E., Siriwardena, D., Drummer, C., Clark, S., Strawbridge, S.E., Kishimoto, K., Vickers, A., Tewary, M., et al. (2022). Spatial profiling of early primate gastrulation in utero. Nature *609*, 136–143. https://doi.org/10.1038/s41586-022-04953-1.

32. Chen, A., Liao, S., Cheng, M., Ma, K., Wu, L., Lai, Y., Qiu, X., Yang, J., Xu, J., Hao, S., et al. (2022). Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. Cell *185*, 1777–1792.e21. https://doi.org/10.1016/j.cell.2022.04.003.

33. Wang, M., Hu, Q., Lv, T., Wang, Y., Lan, Q., Xiang, R., Tu, Z., Wei, Y., Han, K., Shi, C., et al. (2022). High-resolution 3D spatiotemporal transcriptomic maps of developing Drosophila embryos and larvae. Dev. Cell *57*, 1271–1283.e4. https://doi.org/10.1016/j.devcel.2022.04.006.

34. Zeira, R., Land, M., Strzalkowski, A., and Raphael, B.J. (2022). Alignment and integration of spatial transcriptomics data. Nat. Methods *19*, 567–575. https://doi.org/10.1038/s41592-022-01459-6.

35. DeTomaso, D., and Yosef, N. (2021). Hotspot identifies informative gene modules across modalities of single-cell genomics. Cell Syst. *12*, 446–456.e9. https://doi.org/10.1016/j.cels.2021.04.005.

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

37. Tani, S., Chung, U.I., Ohba, S., and Hojo, H. (2020). Understanding paraxial mesoderm development and sclerotome specification for skeletal repair. Exp. Mol. Med. *52*, 1166–1177. https://doi.org/10.1038/s12276-020-0482-1.

38. Reijntjes, S., Stricker, S., and Mankoo, B.S. (2007). A comparative analysis of Meox1 and Meox2 in the developing somites and limbs of the chick embryo. Int. J. Dev. Biol. *51*, 753–759. https://doi.org/10.1387/ijdb.072332sr.

39. El Yakoubi, W., Borday, C., Hamdache, J., Parain, K., Tran, H.T., Vleminckx, K., Perron, M., and Locker, M. (2012). Hes4 controls proliferative properties of neural stem cells during retinal ontogenesis. Stem cells *30*, 2784–2795. https://doi.org/10.1002/stem.1231.

40. Jin, S., Guerrero-Juarez, C.F., Zhang, L., Chang, I., Ramos, R., Kuan, C.-H., Myung, P., Plikus, M.V., and Nie, Q. (2021). Inference and analysis of cell-cell communication using CellChat. Nat. Commun. *12*, 1088. https://doi.org/10.1038/s41467-021-21246-9.

41. Cui, G., Feng, S., Yan, Y., Wang, L., He, X., Li, X., Duan, Y., Chen, J., Tang, K., Zheng, P., et al. (2022). Spatial molecular anatomy of germ layers in the gastrulating cynomolgus monkey embryo. Cell Rep. *40*, 111285. https://doi.org/10.1016/j.celrep.2022.111285.

42. van Nes, J., de Graaff, W., Lebrin, F., Gerhard, M., Beck, F., and Deschamps, J. (2006). The Cdx4 mutation affects axial development and reveals an essential role of Cdx genes in the ontogenesis of the placental labyrinth in mice. Development *133*, 419–428. https://doi.org/10.1242/dev.02216.

43. Lawson, K.A., Meneses, J.J., and Pedersen, R.A. (1991). Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. Development *113*, 891–911. https://doi.org/10.1242/dev.113.3.891.

44. Tam, P.P., Parameswaran, M., Kinder, S.J., and Weinberger, R.P. (1997). The allocation of epiblast cells to the embryonic heart and other mesodermal lineages: the role of ingression and tissue movement during gastrulation. Development *124*, 1631–1642. https://doi.org/10.1242/dev.124.9.1631.

45. Mendjan, S., Mascetti, V.L., Ortmann, D., Ortiz, M., Karjosukarso, D.W., Ng, Y., Moreau, T., and Pedersen, R.A. (2014). NANOG and CDX2 pattern distinct subtypes of human mesoderm during exit from pluripotency. Cell Stem Cell *15*, 310–325. https://doi.org/10.1016/j.stem.2014.06.006.

46. Beccari, L., Moris, N., Girgin, M., Turner, D.A., Baillie-Johnson, P., Cossy, A.-C., Lutolf, M.P., Duboule, D., and Arias, A.M. (2018). Multi-axial self-organization properties of mouse embryonic stem cells into gastruloids. Nature *562*, 272–276. https://doi.org/10.1038/s41586-018-0578-0.

47. Garcia-Fernández, J., and Holland, P.W. (1994). Archetypal organization of the amphioxus Hox gene cluster. Nature *370*, 563–566. https://doi.org/10.1038/370563a0.

48. Cassandri, M., Smirnov, A., Novelli, F., Pitolli, C., Agostini, M., Malewicz, M., Melino, G., and Raschellà, G. (2017). Zinc-finger proteins in health

and disease. Cell Death Discov. *3*, 17071. https://doi.org/10.1038/cddiscovery.2017.71.

49. Rostovskaya, M., Andrews, S., Reik, W., and Rugg-Gunn, P.J. (2022). Amniogenesis occurs in two independent waves in primates. Cell Stem Cell *29*, 744–759.e6. https://doi.org/10.1016/j.stem.2022.03.014.
50. Nowotschin, S., Setty, M., Kuo, Y.-Y., Liu, V., Garg, V., Sharma, R., Simon, C.S., Saiz, N., Gardner, R., Boutet, S.C., et al. (2019). The emergent landscape of the mouse gut endoderm at single-cell resolution. Nature *569*, 361–367. https://doi.org/10.1038/s41586-019-1127-1.
51. Ross, C., and Boroviak, T.E. (2020). Origin and function of the yolk sac in primate embryogenesis. Nat. Commun. *11*, 3760. https://doi.org/10.1038/s41467-020-17575-w.
52. Jiang, X., Zhai, J., Xiao, Z., Wu, X., Zhang, D., Wan, H., Xu, Y., Qi, L., Wang, M., Yu, D., et al. (2023). Identifying a dynamic transcriptomic landscape of the cynomolgus macaque placenta during pregnancy at single-cell resolution. Dev. Cell *58*, 806–821.e7. https://doi.org/10.1016/j.devcel.2023.03.012.
53. Goh, I., Botting, R.A., Rose, A., Webb, S., Engelbert, J., Gitton, Y., Stephenson, E., Quiroga Londoño, M., Mather, M., Mende, N., et al. (2023). Yolk sac cell atlas reveals multiorgan functions during human early development. Science *381*, eadd7564. https://doi.org/10.1126/science.add7564.
54. Zeng, Y., He, J., Bai, Z., Li, Z., Gong, Y., Liu, C., Ni, Y., Du, J., Ma, C., Bian, L., et al. (2019). Tracing the first hematopoietic stem cell generation in human embryo by single-cell RNA sequencing. Cell Res. *29*, 881–894. https://doi.org/10.1038/s41422-019-0228-6.
55. Li, Y., Li, Z., Yang, M., Wang, F., Zhang, Y., Li, R., Li, Q., Gong, Y., Wang, B., Fan, B., et al. (2022). Decoding the temporal and regional specification of microglia in the developing human brain. Cell Stem Cell *29*, 620–634.e6. https://doi.org/10.1016/j.stem.2022.02.004.
56. Bian, Z., Gong, Y., Huang, T., Lee, C.Z.W., Bian, L., Bai, Z., Shi, H., Zeng, Y., Liu, C., He, J., et al. (2020). Deciphering human macrophage development at single-cell resolution. Nature *582*, 571–576. https://doi.org/10.1038/s41586-020-2316-7.
57. Chiu, S.K., Orive, S.L., Moon, M.J., Saw, J., Ellis, S., Kile, B.T., Huang, Y., Chacon, D., Pimanda, J.E., Beck, D., et al. (2019). Shared roles for Scl and Lyl1 in murine platelet production and function. Blood *134*, 826–835. https://doi.org/10.1182/blood.2019896175.
58. Chen, Z., Xie, X., Jiang, N., Li, J., Shen, L., and Zhang, Y. (2021). CCR5 signaling promotes lipopolysaccharide-induced macrophage recruitment and alveolar developmental arrest. Cell Death Dis. *12*, 184. https://doi.org/10.1038/s41419-021-03464-7.
59. Dorey, K., and Amaya, E. (2010). FGF signalling: diverse roles during early vertebrate embryogenesis. Development *137*, 3731–3742. https://doi.org/10.1242/dev.037689.
60. Ikeya, M., and Takada, S. (2001). Wnt-3a is required for somite specification along the anteroposterior axis of the mouse embryo and for regulation of cdx-1 expression. Mech. Dev. *103*, 27–33. https://doi.org/10.1016/s0925-4773(01)00338-0.
61. Isaacs, H.V., Pownall, M.E., and Slack, J.M. (1994). eFGF regulates Xbra expression during Xenopus gastrulation. EMBO J. *13*, 4469–4481. https://doi.org/10.1002/j.1460-2075.1994.tb06769.x.
62. Kiecker, C., and Niehrs, C. (2001). A morphogen gradient of Wnt/β-catenin signalling regulates anteroposterior neural patterning in Xenopus. Development *128*, 4189–4201. https://doi.org/10.1242/dev.128.21.4189.
63. Doniach, T. (1995). Basic FGF as an inducer of anteroposterior neural pattern. Cell *83*, 1067–1070. https://doi.org/10.1016/0092-8674(95)90133-7.
64. Martin, B.L., and Kimelman, D. (2009). Wnt signaling and the evolution of embryonic posterior development. Curr. Biol. *19*, R215–R219. https://doi.org/10.1016/j.cub.2009.01.052.
65. Muñoz-Espín, D., Cañamero, M., Maraver, A., Gómez-López, G., Contreras, J., Murillo-Cuesta, S., Rodríguez-Baeza, A., Varela-Nieto, I., Ruberte, J., Collado, M., and Serrano, M. (2013). Programmed cell senescence during mammalian embryonic development. Cell *155*, 1104–1118. https://doi.org/10.1016/j.cell.2013.10.019.
66. Farrell, J.A., Wang, Y., Riesenfeld, S.J., Shekhar, K., Regev, A., and Schier, A.F. (2018). Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. Science *360*, eaar3131. https://doi.org/10.1126/science.aar3131.
67. Li, B., Yan, Y.-P., He, Y.-Y., Liang, C., Li, M.-Y., Wang, Y., and Yang, Z.-M. (2023). IHH, SHH, and primary cilia mediate epithelial-stromal crosstalk during decidualization in mice. Sci. Signal. *16*, eadd0645. https://doi.org/10.1126/scisignal.add0645.
68. Anand, G.M., Megale, H.C., Murphy, S.H., Weis, T., Lin, Z., He, Y., Wang, X., Liu, J., and Ramanathan, S. (2023). Controlling organoid symmetry breaking uncovers an excitable system underlying human axial elongation. Cell *186*, 497–512.e23. https://doi.org/10.1016/j.cell.2022.12.043.
69. Zhao, L., Song, W., and Chen, Y.-G. (2022). Mesenchymal-epithelial interaction regulates gastrointestinal tract development in mouse embryos. Cell Rep. *40*, 111053. https://doi.org/10.1016/j.celrep.2022.111053.
70. Zhang, Z., Zwick, S., Loew, E., Grimley, J.S., and Ramanathan, S. (2019). Mouse embryo geometry drives formation of robust signaling gradients through receptor localization. Nat. Commun. *10*, 4516. https://doi.org/10.1038/s41467-019-12533-7.
71. Itoh, N., and Ornitz, D.M. (2004). Evolution of the Fgf and Fgfr gene families. Trends Genet. *20*, 563–569. https://doi.org/10.1016/j.tig.2004.08.007.
72. Ornitz, D.M., and Itoh, N. (2001). Fibroblast growth factors. Genome Biol. *2*, REVIEWS3005. https://doi.org/10.1186/gb-2001-2-3-reviews3005.
73. Niehrs, C. (2012). The complex world of WNT receptor signalling. Nat. Rev. Mol. Cell Biol. *13*, 767–779. https://doi.org/10.1038/nrm3470.
74. Lewis, S.L., Khoo, P.-L., De Young, R.A., Steiner, K., Wilcock, C., Mukhopadhyay, M., Westphal, H., Jamieson, R.V., Robb, L., and Tam, P.P.L. (2008). Dkk1 and Wnt3 interact to control head morphogenesis in the mouse. Development *135*, 1791–1801. https://doi.org/10.1242/dev.018853.
75. Frankenberg, S., Smith, L., Greenfield, A., and Zernicka-Goetz, M. (2007). Novel gene expression patterns along the proximo-distal axis of the mouse embryo before gastrulation. BMC Dev. Biol. *7*, 8. https://doi.org/10.1186/1471-213x-7-8.
76. Bedzhov, I., Bialecka, M., Zielinska, A., Kosalka, J., Antonica, F., Thompson, A.J., Franze, K., and Zernicka-Goetz, M. (2015). Development of the anterior-posterior axis is a self-organizing process in the absence of maternal cues in the mouse embryo. Cell Res. *25*, 1368–1371. https://doi.org/10.1038/cr.2015.104.
77. Kemp, C., Willems, E., Abdo, S., Lambiv, L., and Leyns, L. (2005). Expression of all Wnt genes and their secreted antagonists during mouse blastocyst and postimplantation development. Dev. Dyn. *233*, 1064–1075. https://doi.org/10.1002/dvdy.20408.
78. Ark, A.V., Cao, J., and Li, X. (2018). TGF-β receptors: In and beyond TGF-β signaling. Cell. Signal. *52*, 112–120. https://doi.org/10.1016/j.cellsig.2018.09.002.
79. Schier, A.F., and Shen, M.M. (2000). Nodal signalling in vertebrate development. Nature *403*, 385–389. https://doi.org/10.1038/35000126.
80. Cheng, S.K., Olale, F., Bennett, J.T., Brivanlou, A.H., and Schier, A.F. (2003). EGF-CFC proteins are essential coreceptors for the TGF-β signals Vg1 and GDF1. Genes Dev. *17*, 31–36. https://doi.org/10.1101/gad.1041203.
81. Ghyselinck, N.B., and Duester, G. (2019). Retinoic acid signaling pathways. Development *146*, dev167502. https://doi.org/10.1242/dev.167502.
82. Hernandez, R.E., Putzke, A.P., Myers, J.P., Margaretha, L., and Moens, C.B. (2007). Cyp26 enzymes generate the retinoic acid response pattern

necessary for hindbrain development. Development *134*, 177–187. https://doi.org/10.1242/dev.02706.

83. Yang, C., Qi, Y., and Sun, Z. (2021). The role of sonic hedgehog pathway in the development of the central nervous system and aging-related neurodegenerative diseases. Front. Mol. Biosci. *8*, 711710. https://doi.org/10.3389/fmolb.2021.711710.
84. Daane, J.M., and Downs, K.M. (2011). Hedgehog signaling in the posterior region of the mouse gastrula suggests manifold roles in the fetal-umbilical connection and posterior morphogenesis. Dev. Dyn. *240*, 2175–2193. https://doi.org/10.1002/dvdy.22711.
85. Leitch, H.G., Tang, W.W.C., and Surani, M.A. (2013). Primordial germ-cell development and epigenetic reprogramming in mammals. Curr. Top. Dev. Biol. *104*, 149–187. https://doi.org/10.1016/b978-0-12-416027-9.00005-x.
86. Ohinata, Y., Payer, B., O'Carroll, D., Ancelin, K., Ono, Y., Sano, M., Barton, S.C., Obukhanych, T., Nussenzweig, M., Tarakhovsky, A., et al. (2005). Blimp1 is a critical determinant of the germ cell lineage in mice. Nature *436*, 207–213. https://doi.org/10.1038/nature03813.
87. Esfahani, S.N., Zheng, Y., Arabpour, A., Irizarry, A.M.R., Kobayashi, N., Xue, X., Shao, Y., Zhao, C., Agranonik, N.L., Sparrow, M., et al. (2024). Derivation of human primordial germ cell-like cells in an embryonic-like culture. Nat. Commun. *15*, 167. https://doi.org/10.1038/s41467-023-43871-2.
88. Kobayashi, T., Zhang, H., Tang, W.W.C., Irie, N., Withey, S., Klisch, D., Sybirna, A., Dietmann, S., Contreras, D.A., Webb, R., et al. (2017). Principles of early human development and germ cell program from conserved model systems. Nature *546*, 416–420. https://doi.org/10.1038/nature22812.
89. Hancock, G.V., Wamaitha, S.E., Peretz, L., and Clark, A.T. (2021). Mammalian primordial germ cell specification. Development *148*, dev189217. https://doi.org/10.1242/dev.189217.
90. Sasaki, K., Nakamura, T., Okamoto, I., Yabuta, Y., Iwatani, C., Tsuchiya, H., Seita, Y., Nakamura, S., Shiraki, N., Takakuwa, T., et al. (2016). The germ cell fate of cynomolgus monkeys is specified in the nascent amnion. Dev. Cell *39*, 169–185. https://doi.org/10.1016/j.devcel.2016.09.007.
91. Castillo-Venzor, A., Penfold, C.A., Morgan, M.D., Tang, W.W., Kobayashi, T., Wong, F.C., Bergmann, S., Slatery, E., Boroviak, T.E., Marioni, J.C., and Surani, M.A. (2023). Origin and segregation of the human germline. Life Sci. Alliance *6*, e202201706. https://doi.org/10.26508/lsa.202201706.
92. Petropoulos, S., Edsgärd, D., Reinius, B., Deng, Q., Panula, S.P., Codeluppi, S., Plaza Reyes, A., Linnarsson, S., Sandberg, R., and Lanner, F. (2016). Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. Cell *165*, 1012–1026. https://doi.org/10.1016/j.cell.2016.03.023.
93. Van de Sande, B., Flerin, C., Davie, K., De Waegeneer, M., Hulselmans, G., Aibar, S., Seurinck, R., Saelens, W., Cannoodt, R., Rouchon, Q., et al. (2020). A scalable SCENIC workflow for single-cell gene regulatory network analysis. Nat. Protoc. *15*, 2247–2276. https://doi.org/10.1038/s41596-020-0336-2.
94. Kurtzer, G.M., Sochat, V., and Bauer, M.W. (2017). Singularity: Scientific containers for mobility of compute. PLoS One *12*, e0177459. https://doi.org/10.1371/journal.pone.0177459.
95. Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat. Biotechnol. *32*, 381–386. https://doi.org/10.1038/nbt.2859.
96. Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. *19*, 15. https://doi.org/10.1186/s13059-017-1382-0.
97. Palla, G., Spitzer, H., Klein, M., Fischer, D., Schaar, A.C., Kuemmerle, L.B., Rybakov, S., Ibarra, I.L., Holmberg, O., Virshup, I., et al. (2022). Squidpy: a scalable framework for spatial omics analysis. Nat. Methods *19*, 171–178. https://doi.org/10.1038/s41592-021-01358-2.
98. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive integration of single-cell data. Cell *177*, 1888–1902.e21. https://doi.org/10.1016/j.cell.2019.05.031.
99. Kumar, L., and E Futschik, M. (2007). Mfuzz: A software package for soft clustering of microarray data. Bioinformation *2*, 5–7. https://doi.org/10.6026/97320630002005.
100. Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics *32*, 2847–2849. https://doi.org/10.1093/bioinformatics/btw313.
101. Jin, S., Plikus, M.V., and Nie, Q. (2023). CellChat for systematic analysis of cell-cell communication from single-cell and spatially resolved transcriptomics. Preprint at bioRxiv. https://doi.org/10.1101/2023.11.05.565674.
102. Smedley, D., Haider, S., Durinck, S., Pandini, L., Provero, P., Allen, J., Arnaiz, O., Awedh, M.H., Baldock, R., Barbiera, G., et al. (2015). The BioMart community portal: an innovative alternative to large, centralized data repositories. Nucleic Acids Res. *43*, W589–W598. https://doi.org/10.1093/nar/gkv350.
103. Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat. Commun. *10*, 1523. https://doi.org/10.1038/s41467-019-09234-6.
104. Sherman, B.T., Hao, M., Qiu, J., Jiao, X., Baseler, M.W., Lane, H.C., Imamichi, T., and Chang, W. (2022). DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Res. *50*, W216–W221. https://doi.org/10.1093/nar/gkac194.
105. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation *2*, 100141. https://doi.org/10.1016/j.xinn.2021.100141.
106. Bergen, V., Lange, M., Peidli, S., Wolf, F.A., and Theis, F.J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. Nat. Biotechnol. *38*, 1408–1414. https://doi.org/10.1038/s41587-020-0591-3.
107. Chen, T., Chen, X., Zhang, S., Zhu, J., Tang, B., Wang, A., Dong, L., Zhang, Z., Yu, C., Sun, Y., et al. (2021). The genome sequence archive family: toward explosive data growth and diverse data types. Genom. Proteom. Bioinform. *19*, 578–583. https://doi.org/10.1016/j.gpb.2021.08.001.
108. Xue, Y., Bao, Y., Zhang, Z., Zhao, W., Xiao, J., He, S., Zhang, G., Li, Y., Zhao, G., Chen, R., et al. (2021). Database resources of the national genomics data center, china national center for bioinformation in 2022. Nucleic Acids Res. *50*, D27–D38. https://doi.org/10.1093/nar/gkab951.
109. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| **Antibodies** | | |
| anti-TFAP2C | Santa Cruz | Cat # sc-12762; RRID: AB_667770 |
| anti-SOX2 | Santa Cruz | Cat # sc-3658s23; RRID: AB_10842165 |
| anti-GABBR2 | Bioss | Cat # bs-5357R; RRID: AB_11082892 |
| anti-IGFBP3 | Proteintech | Cat # 10189-2-AP; RRID: AB_2123233 |
| anti-NPNT | Bioss | Cat # bs-19202R; |
| anti-Brachyury (T) | R&D | Cat # AF2085; RRID: AB_2200235 |
| anti-CDX2 | Abcam | Cat # AP-0008; RRID: AB_10705337 |
| anti-Wnt5b | Invitrogen | Cat # PA5-116129; RRID: AB_2900763 |
| anti-PRDM1/BLIMP1 | Abcam | Cat # ab198287; RRID: AB_3065084 |
| anti-Sonic Hedgehog (SHH) | Abcam | Cat # ab53281; RRID: AB_882648 |
| anti-SOX17 | R&D | Cat # AF1924; RRID: AB_355060 |
| **Deposited data** | | |
| Human CS7 embryo | Tyser et al.[28] | ArrayExpress: E-MTAB-9388 |
| Mouse embryos | Pijuan-Sala et al.[4] | ArrayExpress: E-MTAB-6967 |
| Cynomolgus monkey embryos | Zhai et al.[29] | GEO: GSE193007 |
| Human blastocyst *in vitro* culture | Xiang et al.[15] | GEO: GSE136447 |
| Cynomolgus monkey embryos *in vitro* culture | Ma et al.[12] | GEO: GSE130114 |
| Human preimplantation embryos | Petropoulos et al.[92] | ArrayExpress: E-MTAB-3929 |
| Stereo-seq | This paper | GSA-human: HRA005567 |
| **Software and algorithms** | | |
| SAW (v5.4.0) | N/A | https://github.com/STOmics/SAW |
| pySCENIC (v0.12.1) | Van de Sande et al.[93] | https://github.com/aertslab/pySCENIC/ |
| Singularity (v3.10.0) | Kurtzer et al.[94] | https://github.com/sylabs/singularity |
| PASTE (v1.3.0) | Zeira et al.[34] | https://github.com/raphael-group/paste |
| Maya (v2023) | Autodesk | https://www.autodesk.com.cn/products/maya |
| ImageStudio (V1.1.0) | N/A | https://www.stomics.tech/col311/list |
| Hotspot (v0.9.1) | DeTomaso et al.[35] | http://www.github.com/Yoseflab/Hotspot |
| Monocle (v2.26.0) | Trapnell et al.[95] | https://github.com/cole-trapnell-lab/monocle-release |
| Scanpy (v1.9.3) | Wolf et al.[96] | https://github.com/theislab/scanpy |
| Squidpy (v1.2.3) | Palla et al.[97] | https://github.com/theislab/squidpy |
| Seurat (v4.3.0) | Stuart et al.[98] | https://satijalab.org/seurat/ |
| Mfuzz (v2.58.0) | Kumar et al.[99] | https://github.com/MatthiasFutschik/Mfuzz |
| ComplexHeatmap (v2.14.0) | Gu et al.[100] | https://github.com/jokergoo/ComplexHeatmap |
| Plotly (v4.10.1) | N/A | https://plotly.com/r/ |
| CellChat (v2.0.0) | Jin et al.[101] | https://github.com/sqjin/CellChat |
| biomaRt (v2.54.1) | Smedley et al.[102] | https://github.com/grimbough/biomaRt |
| Metascape | Zhou et al.[103] | https://metascape.org/gp/index.html |
| DAVID | Sherman et al.[104] | https://david.ncifcrf.gov/home.jsp |

(*Continued on next page*)

| Continued | | |
|---|---|---|
| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
| clusterProfiler (v4.6.2) | Wu et al.[105] | https://github.com/YuLab-SMU/clusterProfiler |
| scVelo (v0.2.5) | Bergen et al.[106] | https://github.com/theislab/scvelo |
| Other | | |
| Resource website for the 3D embryo | This paper | cs8.3dembryo.com |
| Custom code | This paper | https://doi.org/10.5281/zenodo.10851179 |

## RESOURCE AVAILABILITY

### Lead contact

Inquiries regarding additional information, resources, and reagents should be addressed to the lead corresponding author, Leqian Yu, at leqianyu@ioz.ac.cn.

### Materials availability

This study did not generate new unique reagents. All the reagents in this study were included in the key resources table.

### Data and code availability

- The raw sequence data generated by Stereo-seq have been deposited in the Genome Sequence Archive,[107] in the National Genomics Data Center,[108] China National Center for Bioinformation, Chinese Academy of Sciences (GSA-Human: HRA005567) that are publicly accessible at https://ngdc.cncb.ac.cn/gsa-human. This paper analyzes publicly available data. These accession numbers for the datasets are listed in the key resources table.
- All code supporting the current study has been deposited at Zenodo and is available at https://doi.org/10.5281/zenodo.10851179
- Additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

## EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

### Human embryo sample collection and preparation

An intact human early embryo and a truncated human early embryo were obtained from healthy pregnant women with appropriate written consent at the First Affiliated Hospital of Zhengzhou University. The Ethics Committee of Scientific Research and Clinical Trials of the First Affiliated Hospital of Zhengzhou University (NO. 2023-KY-0099-001), and the Ethics Committee of the Institute of Zoology, the Chinese Academy of Sciences (2023. NO. 007) approved the use of human embryo tissues from elective terminations for research. All human embryo tissue samples used were discarded material from elective terminations that were obtained following informed consent. In both cases, donors gave informed consent for the use of fetal material in research. The decision to terminate the pregnancy occurred before the decision to donate tissue. No payments were made to donors and the donors knowingly and willingly consented to provide research materials.

The obtained embryo is intact, male, and at the 5-week gestation stage (~E17-19) from elective termination of pregnancy by a healthy donor. The copy number variation analysis of the chorionic villi tissue confirmed a normal karyotype. The length of the embryonic disc is approximately 2 mm, and the primitive streak extends to about one-third of the anterior-posterior axis, which allows us to stage the embryo as CS8.

The samples were obtained and transported in the cold PBS solution. After washed in PBS to remove surface blood, samples were micro-dissected using scalpels and syringes. The intact CS8 embryo was subsequently placed in a low-temperature mold and embedded in an optimal cutting temperature (OCT) compound. Transverse slices, 10μm in thickness, were obtained using the Leica CM1950 cryostat. After sectioning, every other slice was placed on Stereo-seq chips with space acquisition probes. The adjacent slices were placed on a glass slide for further staining experiments. The truncated embryo was immediately fixed in 4% paraformaldehyde overnight at 4°C and subsequently embedded in paraffin.

## METHOD DETAILS

### Stereo-seq sequencing

Stereo-seq captures tissue RNA based on DNA nanoballs (DNBs) with a 500-750 nm distance.[32] Spatial transcriptomics capture was performed following the previously described protocol.[32,33] The sample sections on stereo-seq chips were fixed by methanol and

permeabilized. DNBs captured released RNA which were to synthesize cDNA by reverse transcription. After cDNA synthesis, sequencing libraries were constructed for the subsequent sequencing process.

### Processing of stereo-seq raw data

The SAW pipeline (https://github.com/BGIResearch/SAW) implemented in Singularity[94] was employed to process the raw sequencing data. In brief, the pipeline utilized coordinate identity (CID) sequences to map reads to specific chip coordinates, allowing for a maximum of 1 base mismatch. Reads with Molecular identity (MID) containing N bases or having more than 2 bases with a quality score below 10 were filtered out. The remaining reads were aligned to the human genome (hg38) by STAR[109] and counted using a MAPQ > 10 threshold. Gene annotation was performed by a gene transfer format (GTF) ref. [32]. Following mapping and gene annotation, the embryo data was registered and segmented based on ssDNA using ImageStudio on Windows. Finally, a CID-containing expression profile matrix was generated based on the aforementioned procedures.

### Section alignment

In order to construct a 3D model, we employed PASTE,[34] which leverages both transcriptional similarity and physical distances between spots in adjacent slices, to align 62 slices. PASTE computes probabilistic pairwise alignment of adjacent slices based on transcriptional and spatial similarity using fused Gromov-Wasserstein optimal transport. To align all the 2D slices, we sequentially applied the pairwise_align function in PASTE along the z axis. This process allowed us to assign x-y-z three-dimensional coordinates to each binned (bin50) spot.

### Spatial constrained clustering (SCC)

The expression profile matrix of the CS8 dataset was partitioned into bins consisting of 50 × 50 DNBs. On average, each bin50 spot contained 3,283 unique molecular identifiers (UMIs) and 1,233 genes were detected per spot. Following this, the gene expression profile matrix, along with its corresponding coordinates from 62 slices, underwent a series of preprocessing steps using the Seurat R package.[98] These steps included quality control (percent.mt < 5, min_gene = 100, min_cells = 10), data splitting, rotation, and merging. The processed gene profile matrix was then normalized, and principal component analysis (PCA) was performed using the SCANPY Python package.[96] In the process of performing spatial clustering of the expression profile matrix, we performed the k-nearest neighbor graph ($G_{expression}$) of transcriptome data and the spatial k-nearest neighbor graph ($G_{spatial}$) by Scanpy[96] and Squidpy[97] after normalization. The spatial k-nearest neighbor graph was calculated through sq.gr.spatial_neighbors where the spatial information is the x-y coordinates, while the k-nearest neighbor graph was calculated from sc.pp.neighbors. The combined neighbor graph was derived through union ($G_{combined} = G_{spatial} \cup G_{expression}$),[32,33] and as input for Leiden clustering to get the final clusters.

### Cell cycle analysis

We obtained the gene list from Seurat and calculated the S_score and G2M_score for each cluster using the 'CellCycleScoring' function. These scores were then stored in the metadata. Finally, we visualized the distribution of clusters across different cell cycle phases.

### Cluster annotation

To identify clusters, we performed the rank_genes_groups function of Scanpy. The genes were computed and ranked in each cluster based on t-test. The markers were screened when $p < 0.05$ and score are ranking in the top50. Additionally, to validate marker accuracy, we employed FindAllMarkers() to identify markers meeting the criteria of $p$ value < 0.05 and Log2FC > 0.25. At the same time, we performed Gene Ontology (GO) analysis for top 100 markers in each cluster. Significant ($p$ value < 0.05) cluster-specific pathways were filtered to annotate clusters. Furthermore, cell type identification was also conducted by leveraging established marker gene sets, which included the CS7 human embryo,[28] E6.5-E8.5 mouse embryo,[4] and CS8-CS11 primate embryo.[29]

### 3D modeling

We utilized the x-y information of each slice obtained from Stereo-seq for slice alignment and derived the x-y-z spatial coordinate information using PASTE based on transcriptional and spatial similarity, employing the functions pst.pairwise_align and pst.stack_slices_pairwise. By integrating the coordinate information with cluster information, we generated a 3D point cloud representing cell types with x-y-z spatial coordinate information. Finally, we employed Maya (Autodesk, San Rafael, CA, USA;), a 3D animation software, for visualization of the 3D embryo models, and then exported it in obj. format, and subsequently visualized by R package plotly (https://cran.r-project.org/web/packages/plotly).

### 3D gene expression patterns

Following the completion of 3D modeling, we obtained colored cluster surfaces and a cloud of spots. Barcodes containing normalized gene expression or gene module scores were extracted using FetchData function through Seurat. The expression of selected genes was filtered out when normalized gene expression value is lower than 0. The patterns of gene expression and gene modules were visualized using plotly with color indicate the value. However, for Figure 6, since we displayed the expression patterns of more

than one gene in each embryo, we colored points in red/orange/blue when normalized gene expression is more than 0 for better visualization.

### Cell type mapping and datasets integrated

In order to assess the accuracy of our cell type annotations in comparison to published data, we utilized SMART-seq2 data from CS7[28] as a reference to annotate our CS8 data. The data from both Stereo-seq and SMART-seq2 were subjected to normalization, PCA, FindTransferAnchors, and TransferData steps in Seurat. Subsequently, the data were embedded into a low-dimensional space using Stereo-seq UMAP. By calculating and visualizing cell proportions, we observed substantial consistency between our CS8 dataset and the published CS7 dataset.

Additionally, the CS7 and CS8 data were integrated based on 2000 highly variable genes (HVG) through FindIntegrationAnchors and IntegrateData functions by Seurat.

### SCENIC analysis

Following cluster annotation, we performed pySCENIC analysis[93] using a three-step approach based on cisTarget resources of human. Firstly, we converted the binned (bin50) gene expression matrix to a loom file using Python. Next, we utilized the 'pyscenic grn', 'pyscenic ctx', and 'pyscenic aucell' functions to infer the gene regulatory network, predict regulon activity, and identify the gene regulatory networks associated with major transcription factors and their target genes for each cell type. Regulon specificity scores (RSS) were calculated within each cluster. To identify cell type-specific regulons, we selected the top-ranked regulons based on the RSS scores, which can be visualized using a dot plot. The gene regulatory network was visualized using the igraph R package (http://igraph.org).

### Hotspot analysis

To delineate specific regions within the distinct tissue clusters produced through Stereo-seq and the transcription factors governing the spatial distribution of cell types, we performed SCENIC analysis and extracted the regulon activity matrix. Subsequently, we employed the x-y-z coordinates of each spot and the regulon activity matrix to construct a spot similarity graph by create_knn_graph function through Hotspot[35] analysis. Subsequently, features that exhibit highly non-random expression patterns in the similarity graph are selected based on significant ($p < 0.05$) local autocorrelation. Finally, pairwise evaluations are conducted to construct a $Z$ score matrix, analogous to a correlation matrix, which is subsequently clustered to form gene/regulon modules based on local correlation. To visualize the results, we computed summary per-cell regulon module scores using the hs.calculate_module_scores function, and subsequently extracted the module scores for each spot. These module scores were then integrated into our dataset.

Three primary visualization methods were employed. First, we extracted each module's autocorrelations for visualization using ComplexHeatmap.[100] Additionally, we extracted the spatial information and module score to visualize each module with colors representing module scores using Plotly. Finally, we utilized FeaturePlot to display the scores of the modules in each slice according to spatial coordinates.

### Functional enrichment analysis

The genes obtained from marker gene lists, using default parameters in SCANPY/Seurat or Hotspot, were subjected to Gene Ontology and pathway enrichment analysis. This analysis was performed using the R package clusterProfiler,[105] as well as online platforms such as DAVID[104] and Metascape.[103] We focused on the most statistically significant pathways ($p$-value < 0.05) and visualized them using the R package ggplot2.

### Cell-cell communication analysis

After cluster annotation, CellChat[40,101] was performed to explore cell-cell communications. The binned (bin50) gene expression matrix and spatial information, which had undergone normalization and annotation, served as the input for CellChat analysis using default settings. We utilized the CellChatDB.human database to infer the probability of cell-cell communication and enrich significant ($p < 0.05$) ligand-receptor pairs involved in various pathways, including FGF, WNT, and others. Finally, we integrated gene expression data with cell-cell communication probability and strength to visualize and infer the communication patterns.

### Pseudotime analysis

In order to elucidate the pseudotime dynamics of mesoderm development, we employed Monocle[95] to construct a pseudotime trajectory for mesoderm formation. To achieve this, we integrated marker gene lists obtained from Seurat with genes identified through an unsupervised procedure known as 'dpFeature'. Then order individual cells along the trajectory, thereby capturing the progression of mesoderm formation.

### RNA velocity analysis

Unspliced and spliced RNA were extracted from spots (50 × 50 DNBs) in the CS8 using the SAW method. The resulting data matrix was subjected to preprocessing using scVelo,[106] which involved normalization, feature gene selection, and PCA dimension reduction. Subsequently, default parameters were applied to estimate the dynamical parameters and gene-wise RNA velocity vectors.

These velocity vectors were then projected onto the UMAP space for visualization and analysis. To illustrate the dynamic gene expression along the velocity/latent time, differential kinetic test was conducted by scv.tl.differential_kinetic_test function.

### Mfuzz analysis

To identify spatially differentially expressed transcription factors specific to the mesoderm lineage in the CS8 embryo, we employed a ranking approach to order the mesoderm cells along the anterior-posterior axis. Subsequently, we utilized the Seurat package to extract the average gene expression profiles. Following this, we applied a filtering step to select transcription factors and employed the Mfuzz package[99] in R to cluster their expression into 10 modules based on fuzzy c-means algorithm. Lastly, we filtered out transcription factors that exhibited significant correlation with the anterior-posterior axis and visualized them by ComplexHeatmap.

### Cross-species comparative analysis

To assess the comparison among human, cynomolgus monkeys, and mice, we obtained single-cell transcriptomic data from CS8-CS11 cynomolgus monkey embryos and E6.5-E8.5 mouse embryos, along with their annotation information. Initially, we utilized biomaRt to convert cynomolgus monkey and mouse genes to their human orthologues[102] (Table S5). Principal Component Analysis (PCA) was performed to illustrate the correlation between cell types using the expression of one-to-one orthologous (refer to Figure 7A). To achieve this, we projected monkey and mouse clusters to the UMAP of human CS8 embryo through the function MapQuery from Seurat. Subsequently, we identified and analyzed the enrichment of correlation genes and pathways to gauge species conservation and divergence.

### Immunofluorescent staining

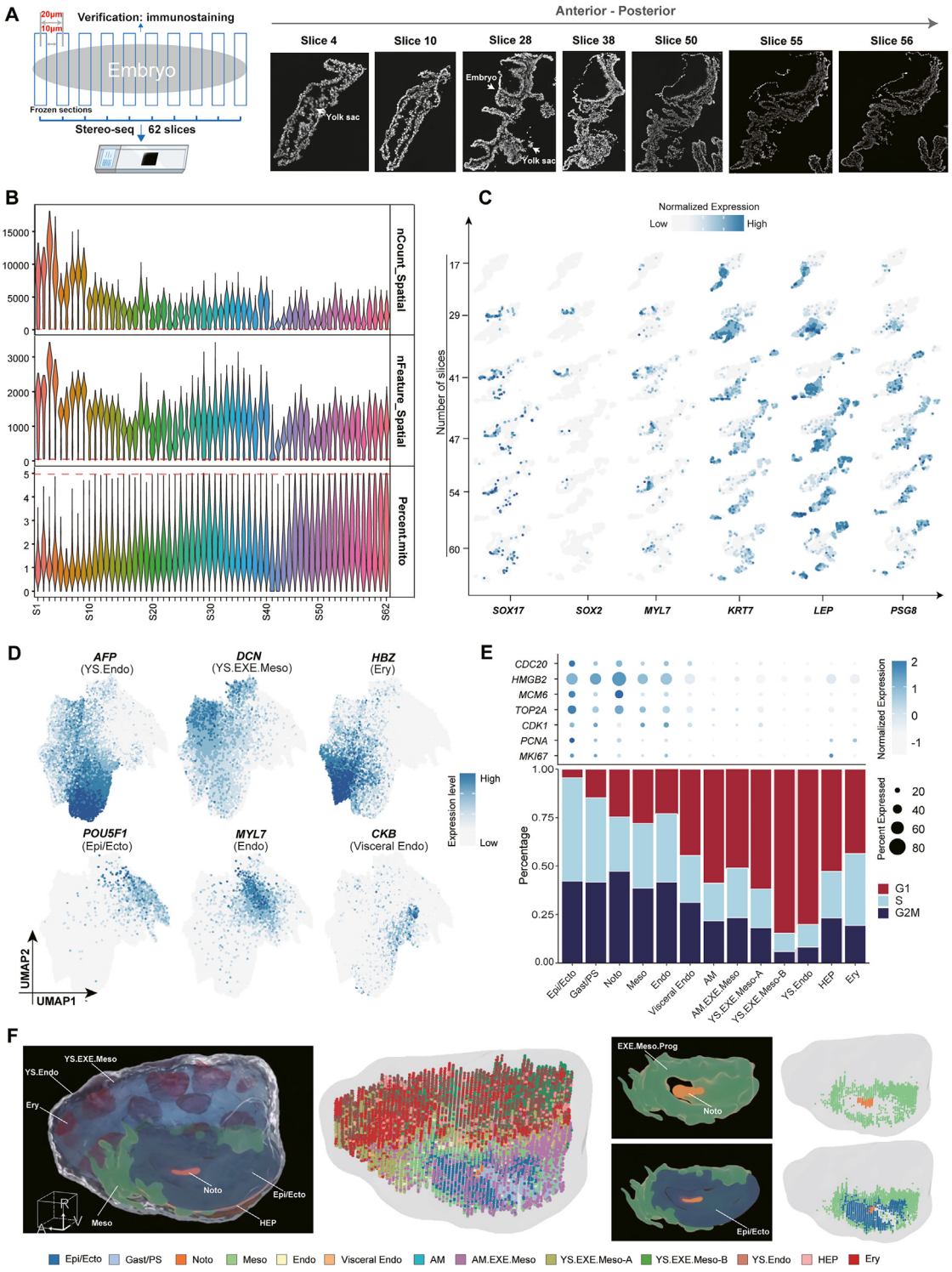
Sections (5 μm) on slides were dewaxed and rehydrated with xylene and ethanol gradients. Slides were immersed in 1mM Tris-EDTA antigen repair buffer solution (pH 9.0) and heated in a microwave oven at 92°C–98°C for 30 min for antigen retrieval. After cooling to room temperature, the slides were washed three times with 1× PBS (5 min each), incubated with 1% Triton X-100 for 30 min and blocked with 2% bovine serum albumin (BSA) for 30 min at room temperature. Next, the slides were incubated with primary antibodies, diluted with blocking solution overnight at 4°C and washed three times with PBST (1× PBS with 0.05% Tween 20, 5 min each). The slides were then incubated with secondary antibodies diluted with blocking solution and 1 mg/mL DAPI (Invitrogen, no. D3571) for 1 h. Finally, after washing three times with PBST (5 min each) the slides were mounted with anti-fade mounting medium (Gibco). IF images were captured by laser-scanning confocal microscope LSM 880 (Carl Zeiss) and processed with Imaris 9.0.2 software (Bitplane) and Zen 7.0 (Carl Zeiss).

## QUANTIFICATION AND STATISTICAL ANALYSIS

Due to the difficulty in the acquisition of human early embryos, analysis in this work was conducted on only one embryo. No methods were used to determine whether the data met the assumptions of the statistical approaches utilized for the biological repeat. In the computational analysis, the statistical analysis outlined in the specific packages was utilized.

CellPress

# Supplemental figures




*(legend on next page)*

**Figure S1. Performance and quality-control-related analyses for Stereo-seq, related to Figure 1**

(A) Left: schematic overview of experimental methodology. For Stereo-seq, the embryo was subjected to serial slices at 10 μm sections and every other sectioned slice was transcriptomic profiled with the carrier of Stereo-seq chip. The average interval distance between consecutive slices was determined to be 20 μm, measured from the center of one slice to the center of the next. Right: spatial visualization of the segmented cells (nucleic acid staining) in slices 4, 10, 28, 38, 50, 55, and 56 (from the anterior to the posterior end of the embryo).

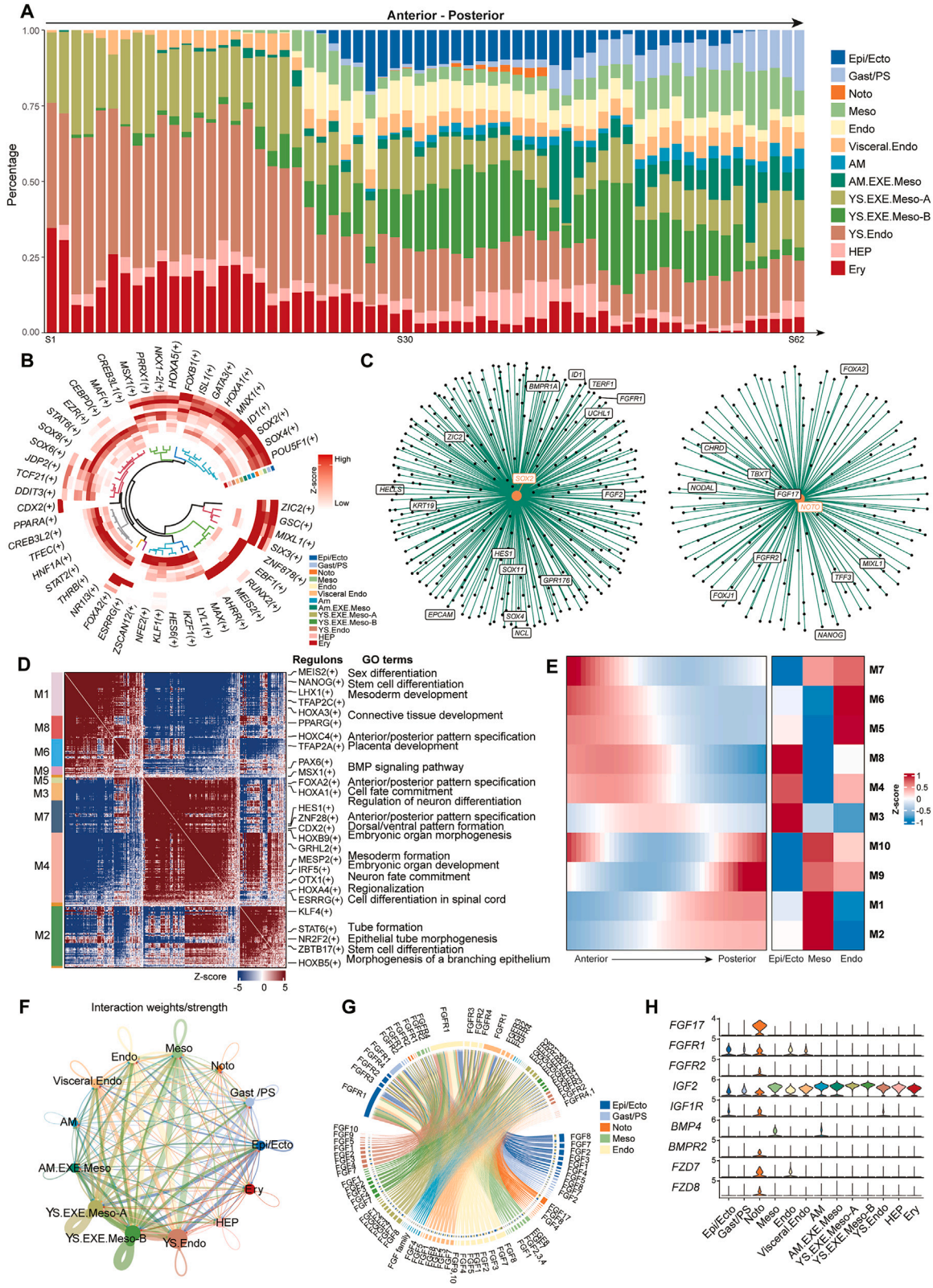
(B) Violin plots of read counts, gene numbers, and percentage of mitochondrial reads detected in each slice. The red dashed line representing the threshold criteria of quality control.

(C) Expression patterns of selected markers are used to help annotate major embryonic cell types cast on the spatial reduction. Blue (or gray) represents a high (or low) expression level as shown on the color key on the top.

(D) UMAP plots showing the expression of marker genes in YS.Endoderm, Endo, Visceral. Endo, Ery, Epi/Ecto, and EXE.Meso.

(E) Top: dotplot depicting the normalized expression levels of cell cycle-related genes in major cell types. Bottom: bar plot showing the distribution of the cell cycle phase by analyzing the expression levels of S- and G2M-phase genes.

(F) 3D model rendering using Maya, from 3D point cloud with x-y-z coordinates of all cell types (left) or intraembryonic cell types (right).




*(legend on next page)*

**Figure S2. Cell type spatial distribution and molecular regulatory landscape, related to Figure 2**

(A) Bar plot showing the fractions of cell types per slice.

(B) Circular heatmap showing the representative TFs in each cluster.

(C) Gene regulatory networks of SOX2 and NOTO in the human CS8 embryo. Selected representative genes were shown.

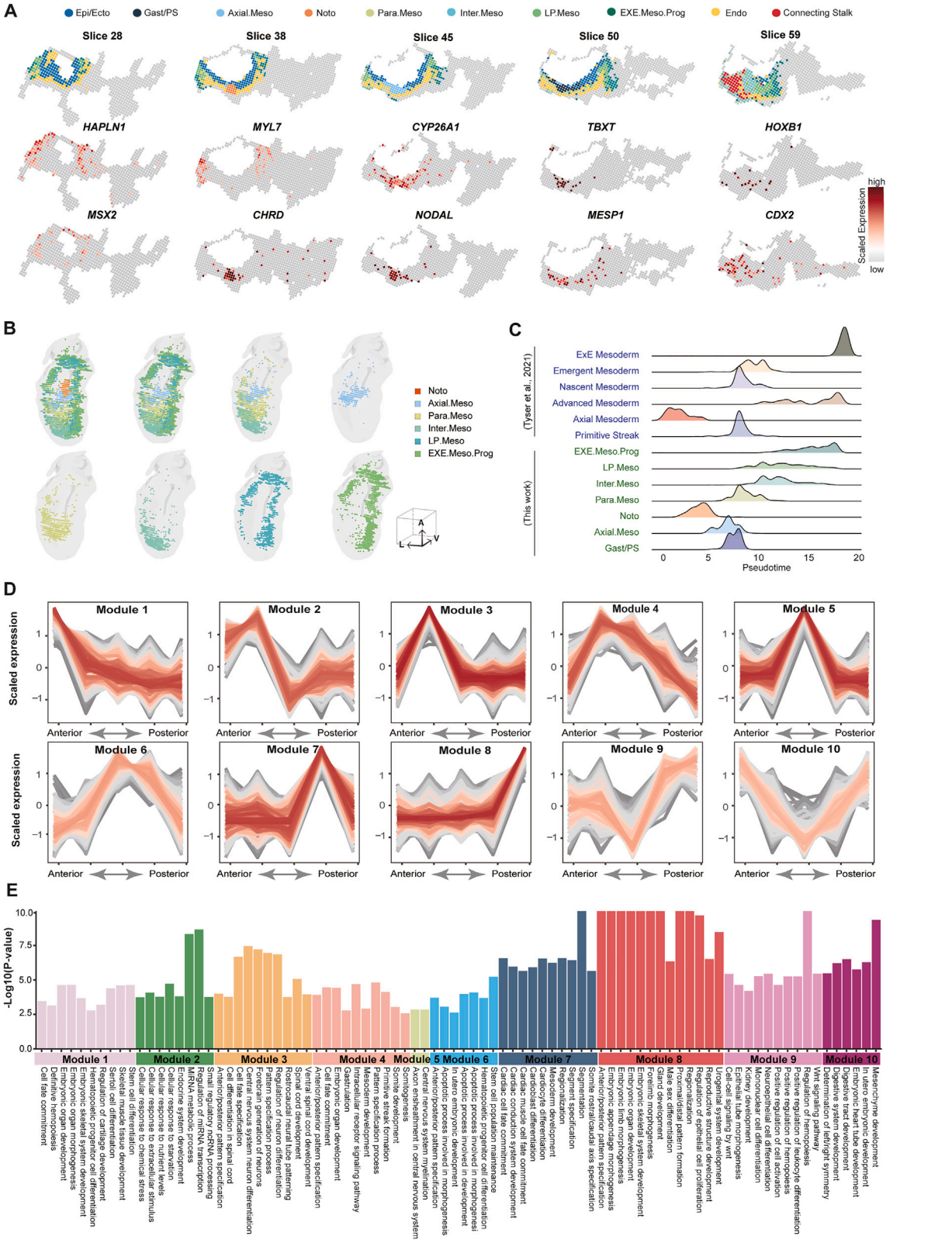
(D) Heatmap showing the correlation of functional regulon modules based on hotspot analysis. Selected GO terms and regulons are highlighted on the right side.

(E) Heatmap plots showing the dynamic expression patterns of modules along A-P or D-V (Epi/Ecto, Meso, and Endo) axes.

(F) The chord diagram showing the cell-cell communication networks within the embryo. The thickness of the line represents the weight of ligand-receptor pairs.

(G) The chord diagram showing the significant interaction pairs involved in the FGF signaling pathways within the embryo.

(H) Violin plots showing the selected ligand-receptor gene expression patterns in the embryo.




*(legend on next page)*

**Figure S3. Spatial distribution of mesoderm and associated gene expression patterns, related to Figure 3**

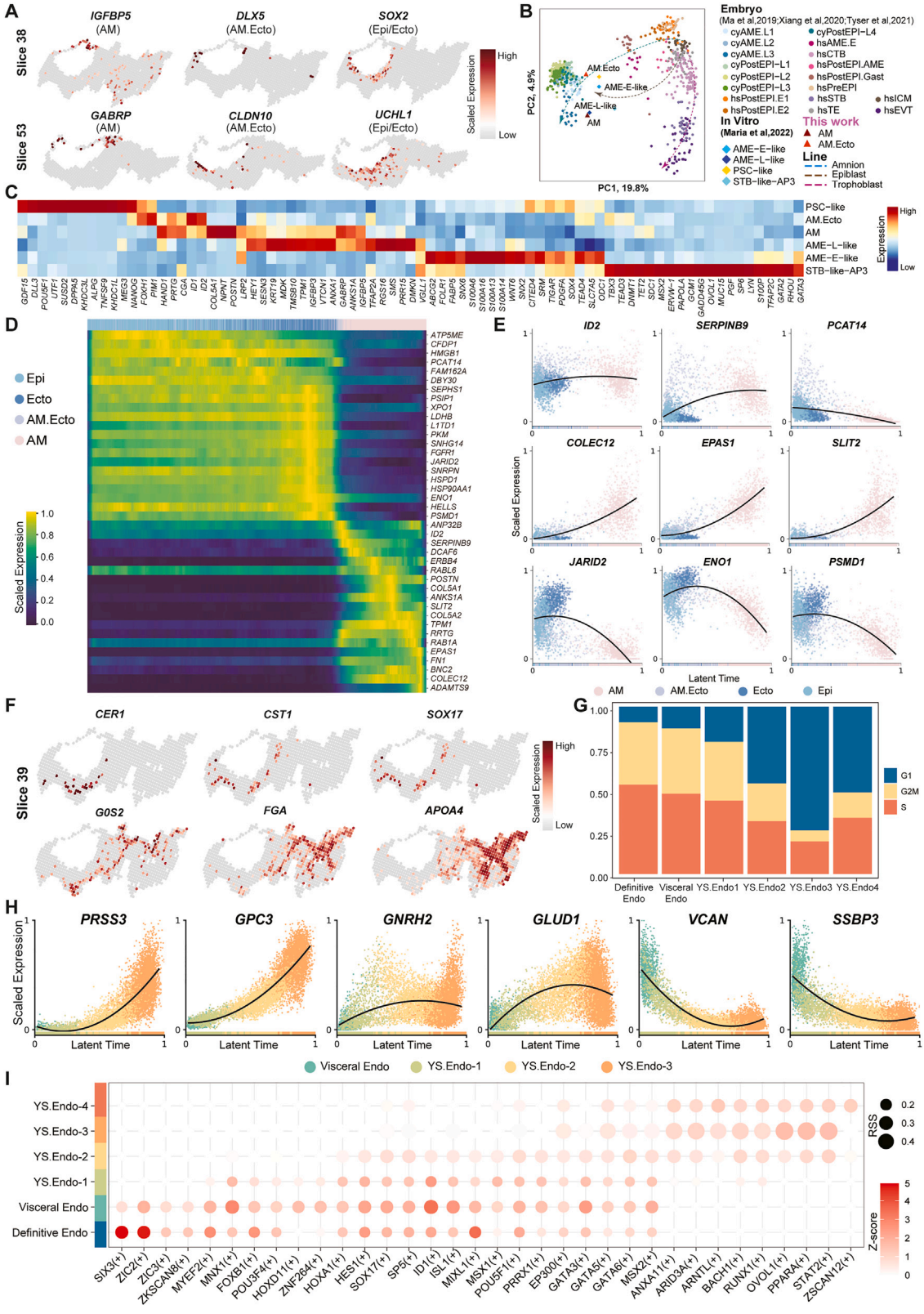
(A) The spatial plots showing the cell type distribution and expression patterns of key representative genes. Red (or gray) represents a high (or low) expression level.

(B) Spatial distribution of mesoderm subtypes in the 3D reconstructed model. The color-code of mesoderm subtypes is consistent with Figure 4A.

(C) Density plot showing the cell distribution along the pseudotime by cell ratio analysis of the mesoderm-related cell clusters in CS7 and CS8 human embryos.

(D) The line plot depicting standardized gene expression levels by fuzzy cluster analysis for ten modules along the A-P axis.

(E) Bar plot displaying the significance ($p$ value < 0.05) of representative GO terms in each module is shown in (D).




*(legend on next page)*

**Figure S4. 3D reconstruction of ectoderm and endoderm, related to Figure 4**

(A) Spatial expression patterns of marker genes of AM, AM.Ecto, Ecto, and Epi in the representative slices (anterior slice 38 and posterior slice 53).

(B) PCA plots showing amniotic cell types from *in vivo*[12,15,28] and *in vitro* [49] datasets. Lines indicating developmental lineages.

(C) Heatmap showing the scaled expression pattern of representative genes in amniotic cell types from *in vivo* and *in vitro* datasets.

(D) Heatmap showing the scaled expression pattern of the top 30 genes identified along the pseudo-developmental trajectory of epiblast and amnion.

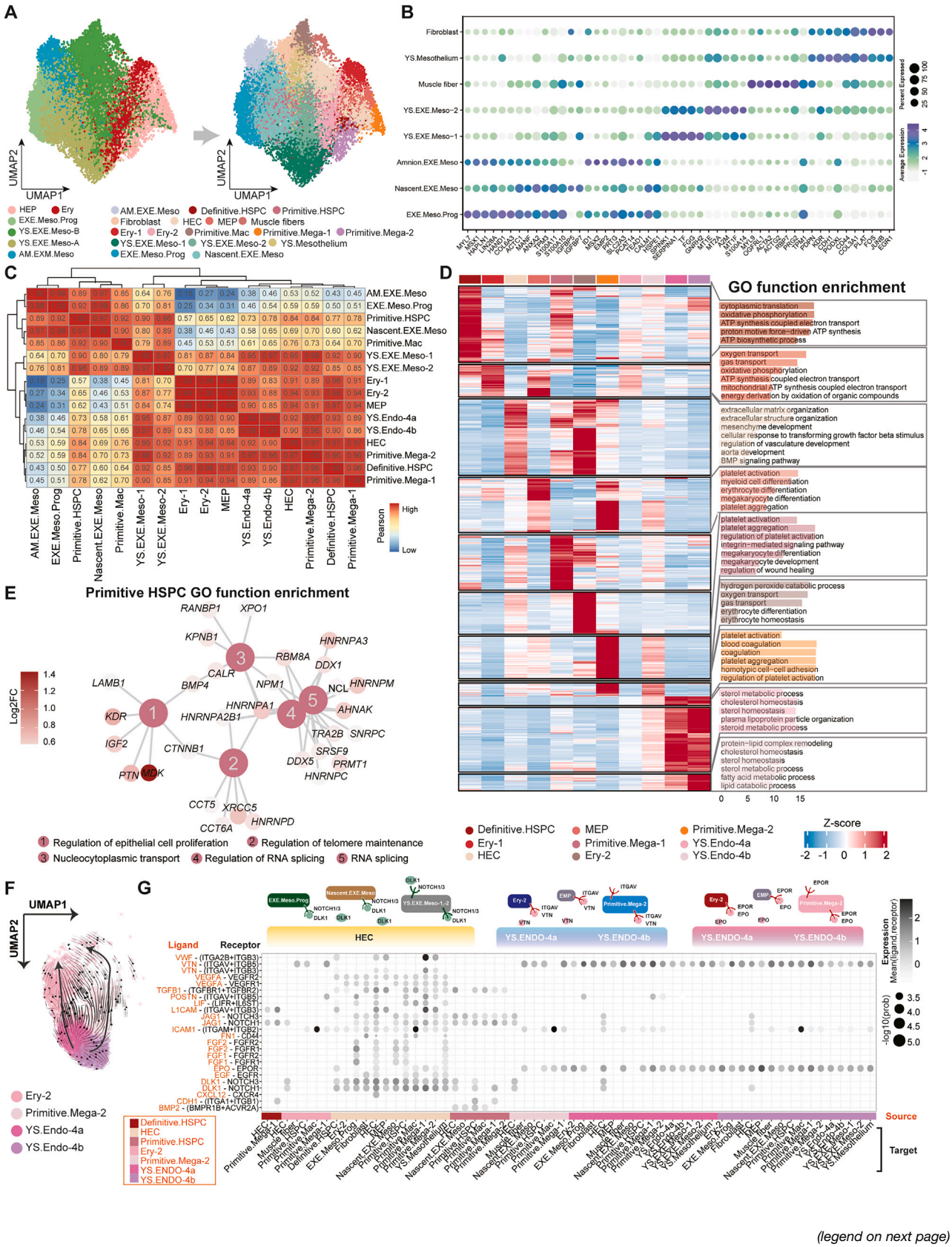
(E) Scatterplots showing expression dynamic of indicated developmental genes along the latent time inferred by RNA velocity analysis of AM, AM.Ecto, Ecto, and Epi.

(F) Spatial expression patterns of endoderm marker genes in the representative slice 39.

(G) Bar plot showing the distribution of the cell cycle phase by analyzing the expression levels of S- and G2M-phase genes.

(H) Scatterplots showing expression dynamics of selected genes along the latent time inferred by RNA velocity analysis of Visceral.Endo, YS.Endo-1, YS.Endo-2, and YS.Endo-3.

(I) Dotplot showing regulatory activity of representative TFs in Definitive.Endo, Visceral.Endo, YS.Endo-1, YS.Endo-2, YS.Endo-3, or YS.Endo-4.




*(legend on next page)*

**Figure S5. Mesoderm spatial distribution and anterior-posterior profiling, related to Figure 5**

(A) UMAP plots showing re-clustering of the selected 4 major extra-embryonic mesoderm cell types and 2 hematopoietic cell types (left) into 17 subclusters (right).

(B) Dotplot depicting the normalized expression levels of representative marker genes within the distinct EXE.Meso clusters.

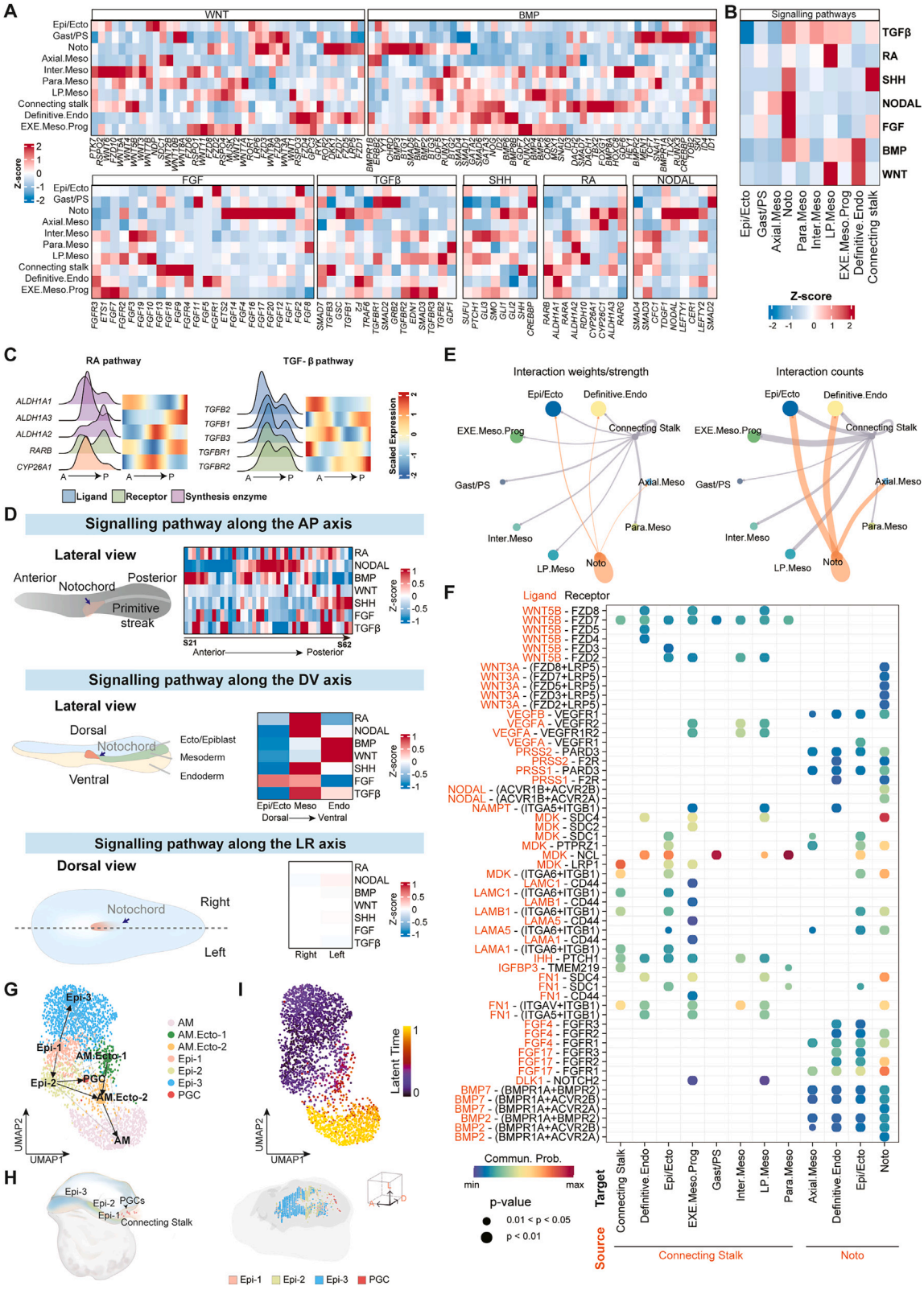
(C) Pearson correlation of EXE.Meso.Prog and its derivatives.

(D) Heatmap showing gene expression programs specific to each hematogenesis-related cell type and the significant enriched GO terms.

(E) Functional networks of the GO enrichment of the primitive HSPC. Node color: fold change of DEGs; edge width: adjusted *p* value. The color bars from grey to red indicate the log2FC from low to high.

(F) RNA velocity analysis overlaid on the UMAP. Dots were colored by cell types.

(G) Bubble plot showing the selected ligand-receptor interactions. The means of the average expression level of interacting molecule 1 in cluster 1 and interacting molecule 2 in cluster 2 are indicated by color.




*(legend on next page)*

**Figure S6. Spatial analysis of key signaling pathways, related to Figure 6**

(A) Heatmap showing the scaled expression pattern of genes involved in WNT, TGF-β, SHH, RA, NODAL, FGF, and BMP pathways in major cell types.

(B) Heatmap showing the activity of WNT, TGF-β, SHH, RA, NODAL, FGF, and BMP pathways in major cell types.

(C) Ridgeline plots of gene distribution and heatmap of gene expression showing dynamic expression patterns of ligands, receptors, co-factors, inhibitors, synthesis enzymes, and metabolic enzymes from TGF-β and RA pathways.

(D) Left: schematic diagram illustrating the segment of cells based along the A-P axis (top), D-V axis (middle), or L-R axis (bottom). Right: heatmap showing dynamic expression patterns of related pathways along the A-P axis (top), D-V axis (middle), or L-R axis (bottom).

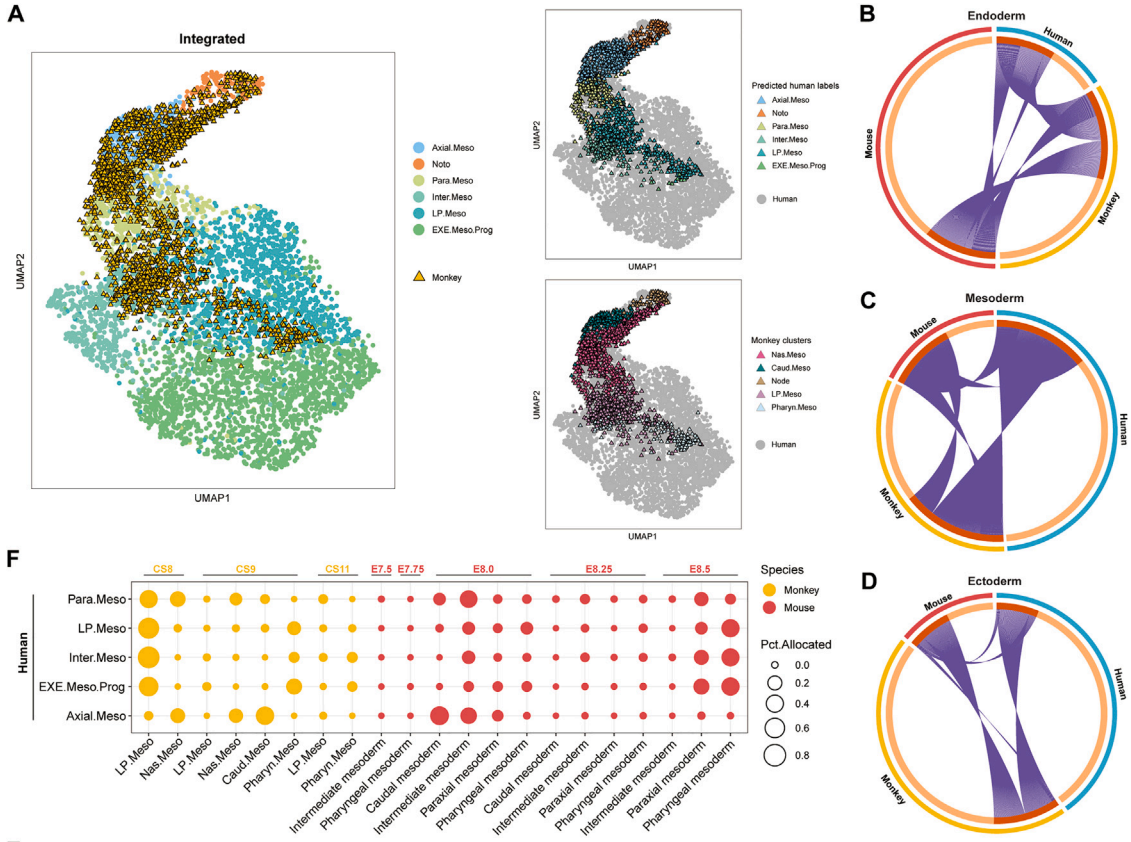
(E) The chord diagram showing the cell-cell communication networks within the embryo. The thickness of the line represents the weight of ligand-receptor pairs (left) or counts of ligand-receptor pairs (right).

(F) Bubble plot showing significant ligand-receptor interaction pairs from connecting stalk and notochord to other cell types (Epi/Ecto, Gast/PS, Definitive Endo, Noto, Axial.Meso, Para.Meso, LP.Meso, Inter.Meso, and EXE.Meso.Prog).

(G) PAGA path of amniotic and epiblast cell types and PGC overlaid on the UMAP.

(H) Spatial distribution patterns of epiblast cell types and PGC in the schematic diagram (left) and the 3D reconstructed CS8 human embryo (right).

(I) Latent time of amniotic and epiblast cell types and PGC overlaid on the UMAP.




*(legend on next page)*

**Figure S7. Cross-species comparison of the gastrula from humans, monkeys, and mice, related to Figure 7**

(A) UMAP visualization of the mesoderm cells from human and monkey CS8 gastrula. Left: colors indicate cell clusters; shapes indicate species. Right: colors indicate cell clusters of humans or monkeys. Nas.Meso, nascent mesoderm; LP.Meso, lateral plate mesoderm; Caud.Meso, caudal mesoderm; Pharyn.Meso, pharyngeal mesoderm (abbreviations for cell types of monkeys, here and after).

(B–D) Chord diagram showing overlaps of marker genes of ectoderm (B), mesoderm (C), and endoderm (D) among humans, monkeys, and mice.

(E) Heatmap of the functionally enriched GO terms by marker genes of three germ layers.

(F) Dot plot showing the percentage of cells from the human mesoderm allocated to monkey and mouse mesoderm cell types.