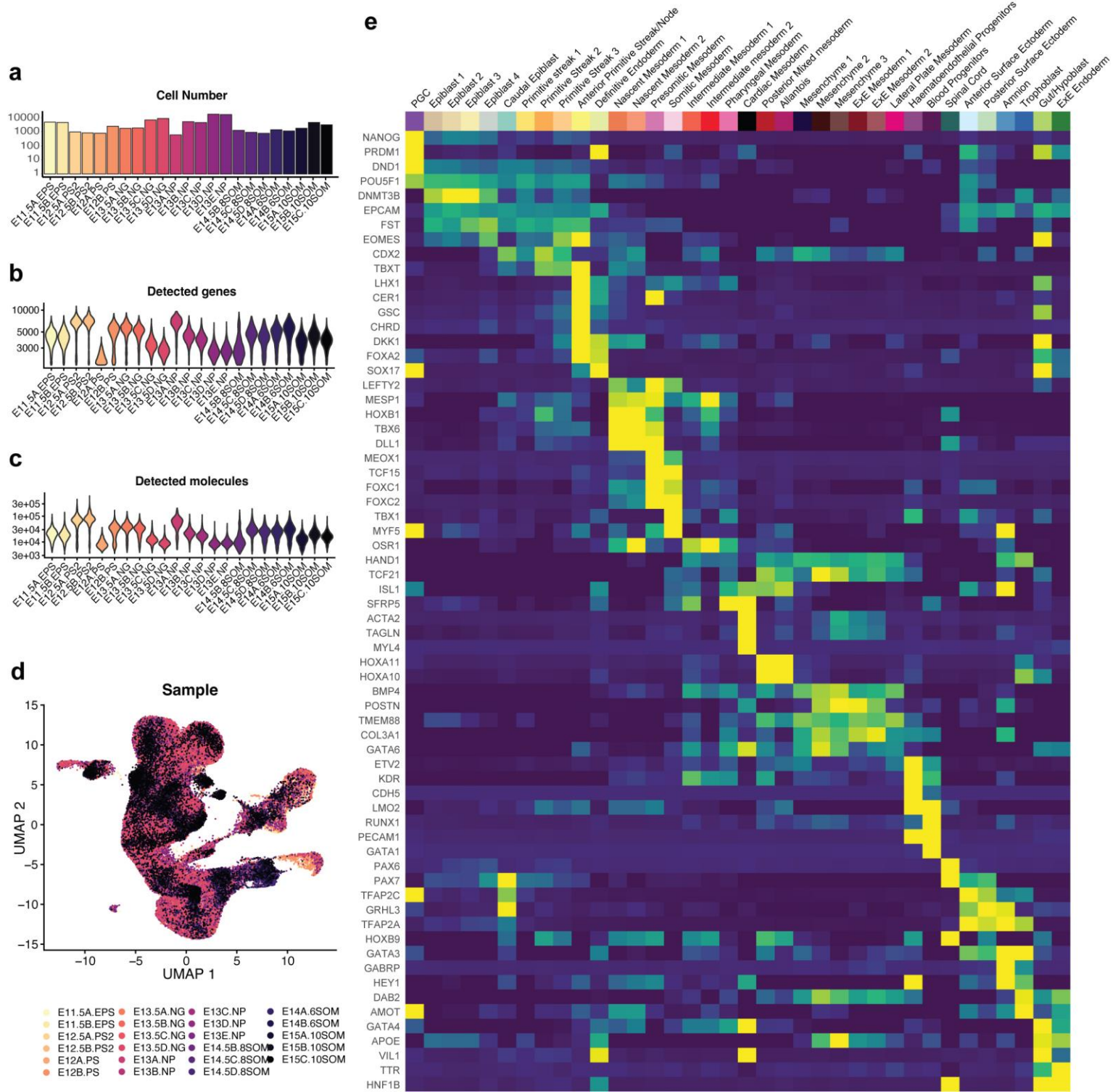
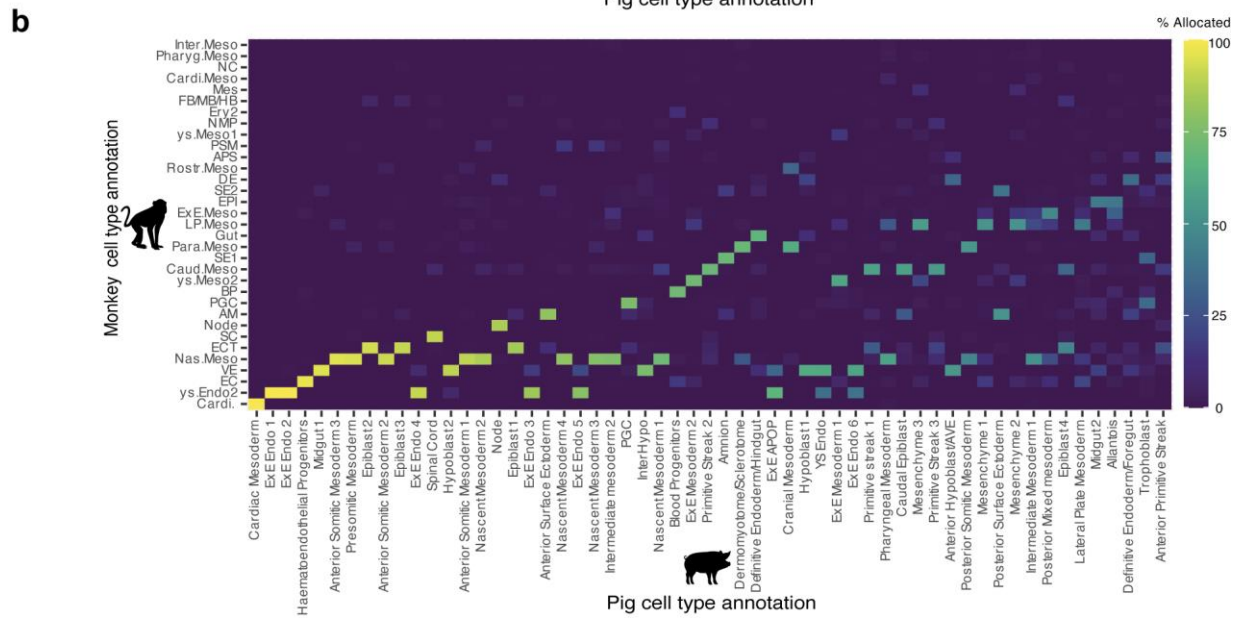
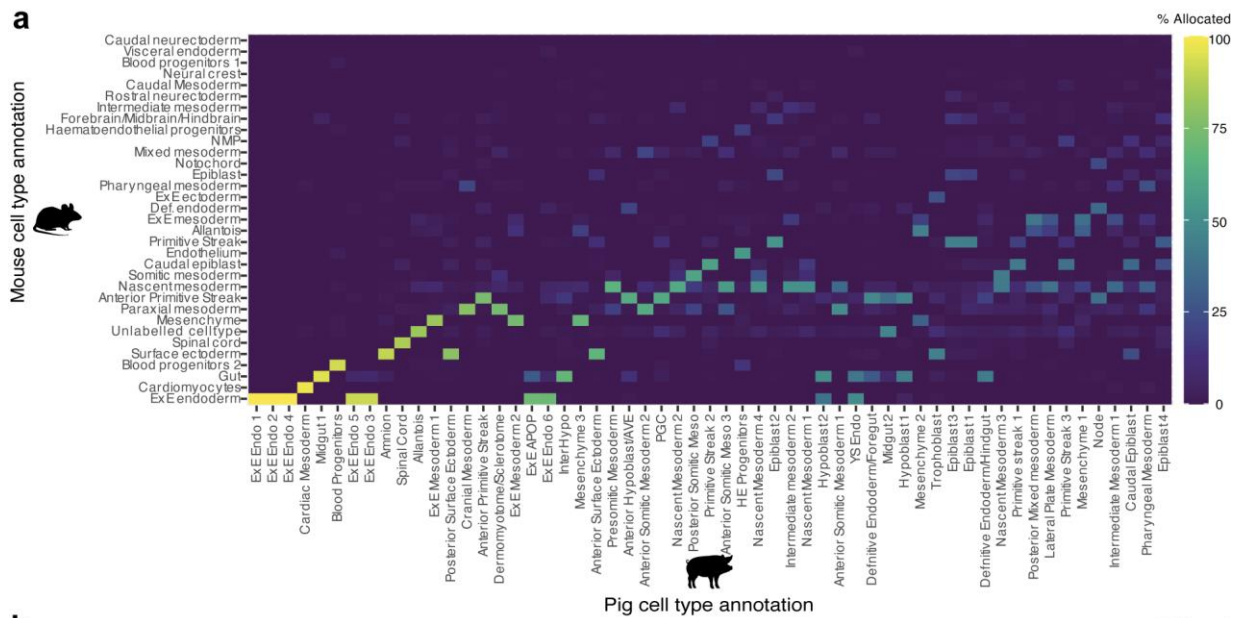


| REAGENT or RESOURCE                   | Antibody dilution used | SOURCE               | IDENTIFIER                         |
|---------------------------------------|------------------------|----------------------|------------------------------------|
| Primary Antibodies                    |                        |                      |                                    |
| TBXT (Goat)                           | 1:500                  | R&D                  | Cat# AF2085, RRID:AB_2200235       |
| SOX17 (Rabbit)                        | 1:500                  | Abcam                | Cat# ab224637, RRID:AB_2801385     |
| SOX17 (Goat)                          | 1:750                  | R&D systems          | Catalog #: AF1924, RRID:AB_355060  |
| FOXA2 (Mouse)                         | 1:750                  | Abnova               | Cat# H00003170-M10, RRID:AB_534871 |
| SOX2 (Rat)                            | 1:5000                 | eBioscience          | Cat# 14-9811-80, RRID:AB_11219070  |
| SNAI1 (Rabbit)                        | 1:200                  | Cell Signaling       | Cat# 3879, RRID:AB_2255011         |
| NANOG (Rabbit)                        | 1:500                  | PeproTech            | Cat# 500-P236, RRID:AB_2929968     |
| Secondary Antibodies                  |                        |                      |                                    |
| Alexa Fluor™ 488 (Donkey Anti-Mouse)  | 1:500                  | Invitrogen           | Cat# A-21206                       |
| Alexa Fluor™ 555 (Donkey Anti-Rabbit) | 1:500                  | Invitrogen           | Cat# A-31572, RRID:AB_162543       |
| Alexa Fluor™ 555 (Donkey Anti-Rat)    | 1:500                  | Invitrogen           | Cat# A78945, RRID:AB_2910652       |
| Alexa Fluor™ 647 (Donkey Anti-Goat)   | 1:500                  | Invitrogen           | Cat# A-21447, RID:AB_2535864       |
| IHC reagents                          |                        |                      |                                    |
| PBS tablets                           |                        | OXOID                | Cat# BR0014G                       |
| Triton-X                              |                        | Sigma-Aldrich        | Cat# 9036-19-5                     |
| Fluoroshield with Dapi                |                        | Sigma-Aldrich        | Cat# F657                          |
| Cell lines                            |                        |                      |                                    |
| EDSCL4 (Porcine)                      |                        | Kinoshita et al 2021 |                                    |
| EDSCL7 (Porcine)                      |                        | Kinoshita et al 2021 |                                    |
| HNES1 (human)                         |                        | Guo et al 2016       |                                    |
| H9 (human)                            |                        | WiCell               |                                    |
| Cell culture reagents and consumables |                        |                      |                                    |
| DMEM/F12 medium                       |                        | Gibco                | Cat# 21103049                      |
| Neurobasal medium                     |                        | Gibco                | Cat# 11320074                      |
| B-27 supplement                       |                        | Gibco                | Cat# 17504001                      |
| N-2 supplement                        |                        | Gibco                | Cat# 17502001                      |
| GlutaMAX                              |                        | Gibco                | Cat# 35050038                      |
| 2-Mercaptoethanol (50 mM)             |                        | Gibco                | Cat# 31350010                      |
| Recombinant human Activin A protein   |                        | Qkine                | Cat# Qk001                         |
| Recombinant zebrafish FGF2 protein    |                        | Qkine                | Cat# Qk002                         |
| XAV939 (WNT inhibitor)                |                        | Sigma-Aldrich        | Cat# X3004                         |
| Y-27632 dihydrochloride (ROCKi)       |                        | Torcris              | Cat# 1254                          |
| LS11-E8-silk (iMatrix-511)            |                        | AMSBIO               | Cat# AMS.892 021                   |
| TRYPLE EXPRESS                        |                        | Gibco                | Cat# 12604013                      |
| CTS™ Synth-a-Freeze™ Medium           |                        | Gibco                | Cat# A1371301                      |
| NutriFreeze D10                       |                        | Sartorius            | Cat# 05-713-1E                     |
| 96-Well CytoOne Plate, TC-Treated     |                        | Starlabs             | Cat# CC7682-7596                   |
| Bovine Serum Albumin solution         |                        | Sigma-Aldrich        | Cat# A8412                         |

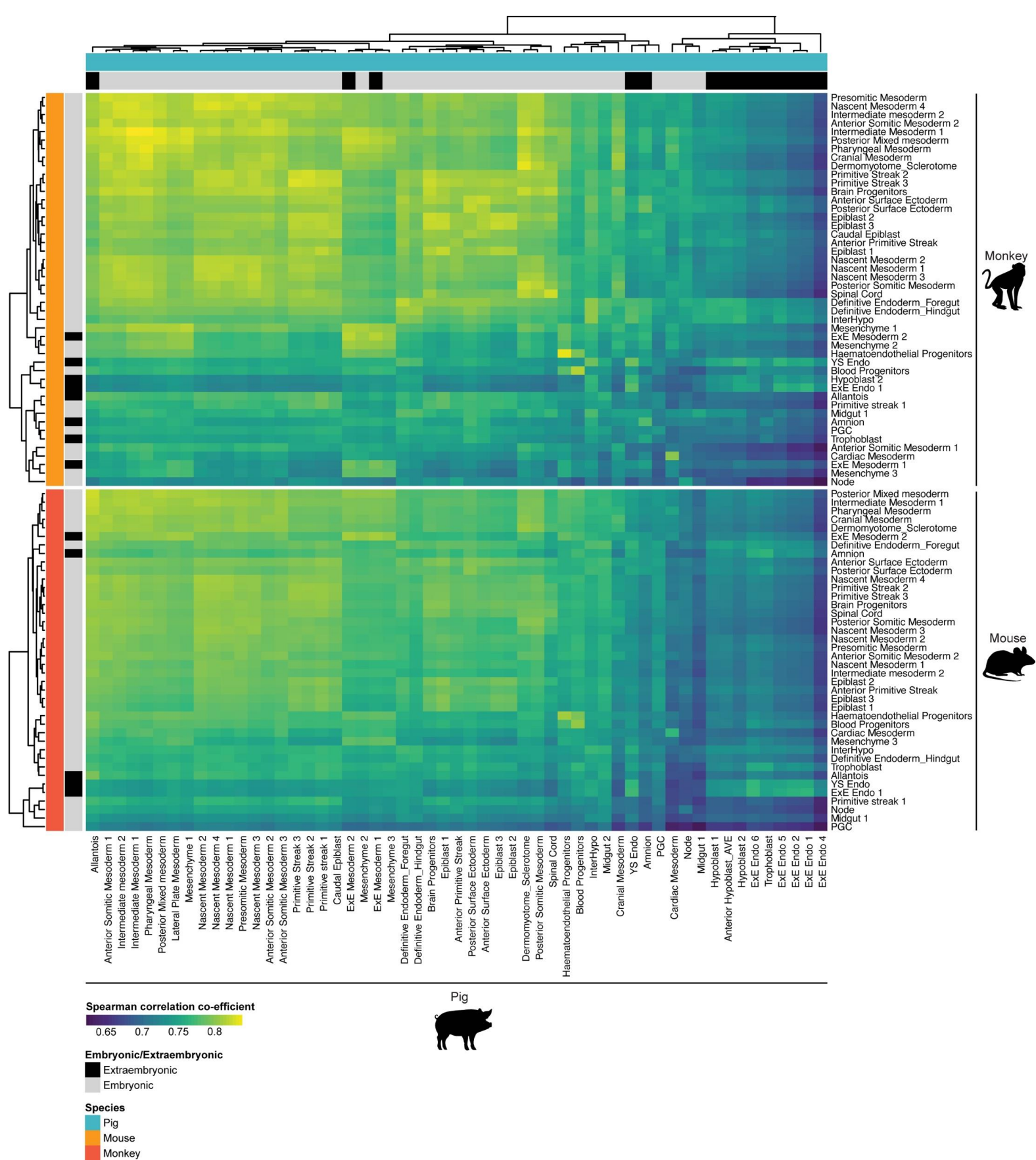
Supplementary Table 1: Summary of reagents used within this study



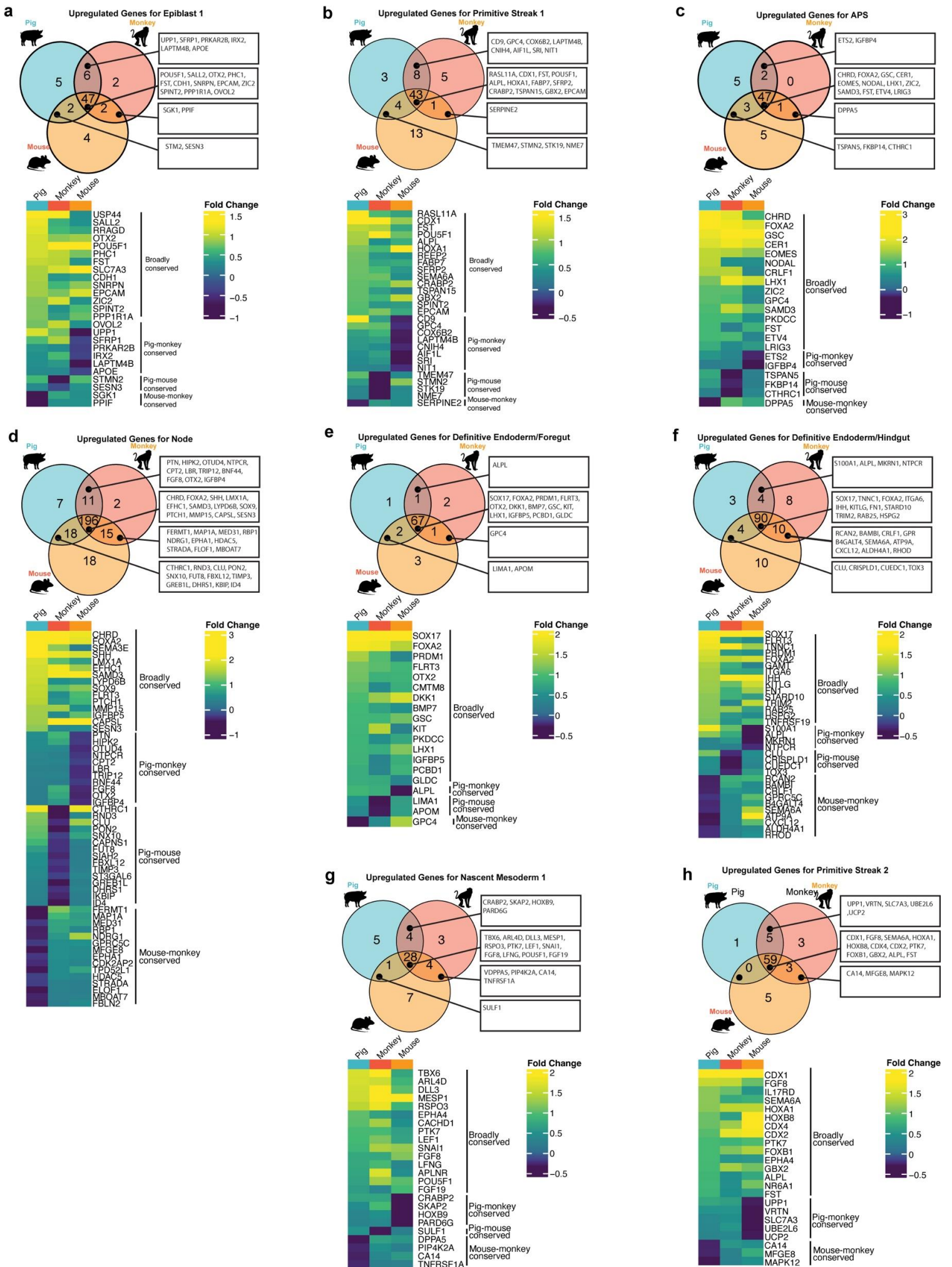
**Supplementary Fig. 1. Cell distribution per sample and gene markers used to define cell identity.** **a**, Number of cells captured in this study in each sample used in this atlas. E, Embryonic day. **b**, Violin plots illustrating the number of detected genes per cell per sample. **c**, The number of molecules detected per cell per sample. **d**, UMAP plot showing atlas cells from Fig.1c) coloured by sample, showing consistency across replicates from the same stage. **e**, Heatmap showing the markers used for cell type identification, a full list of these markers and the source publications are available in Supplementary Table 1. ExE; Extra-embryonic, PGC; Primordial germ cell.



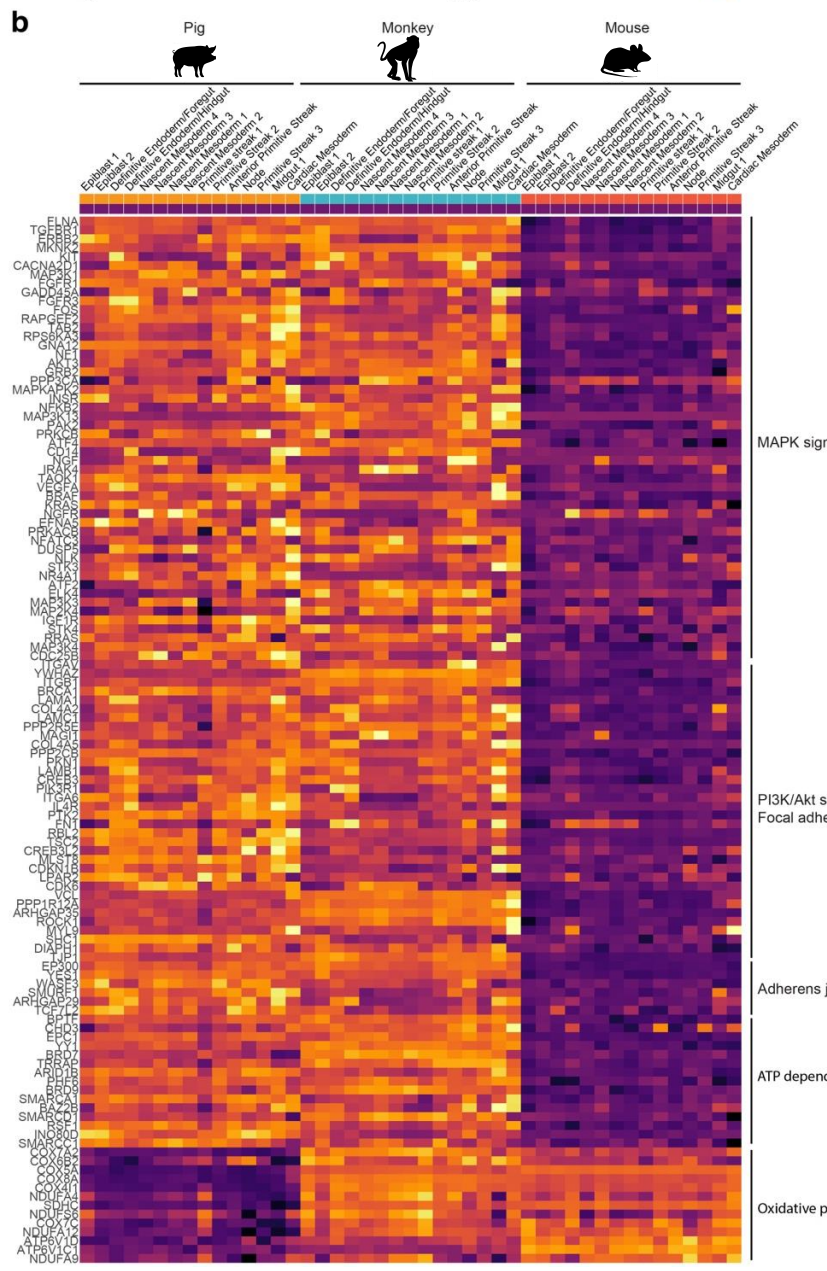
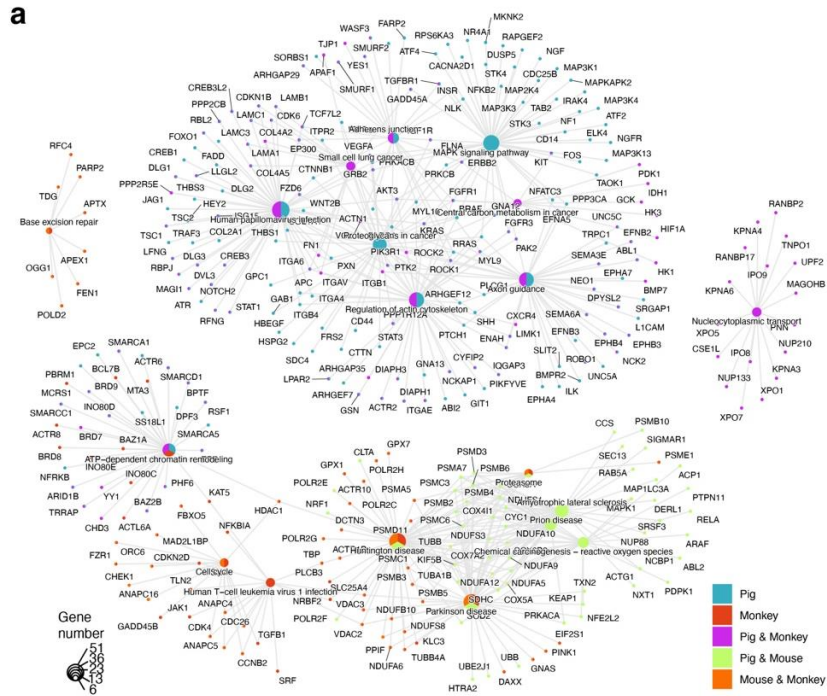
**Supplementary Fig. 2. Comparisons of mouse and monkey cell annotations.** **a**, Heat map showing the percentage of pig cells of a given cell type allocated as a particular mouse cell type after label transfer. **b**, Heat map showing the percentage of a given pig cell type allocated as a particular monkey cell type by label transfer. A percentage of 100 indicates that all of the cells of a given cell type were predicted to be analogous to the cell identity in the queried organism.



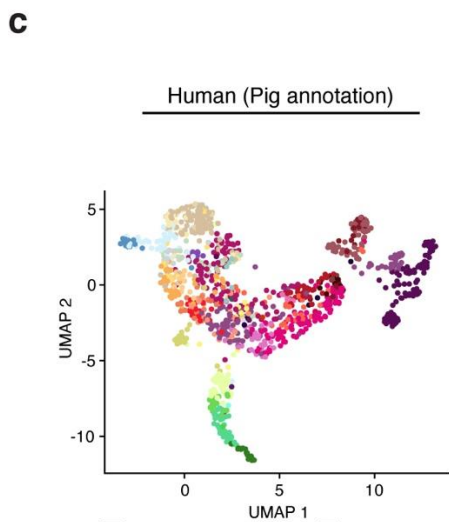
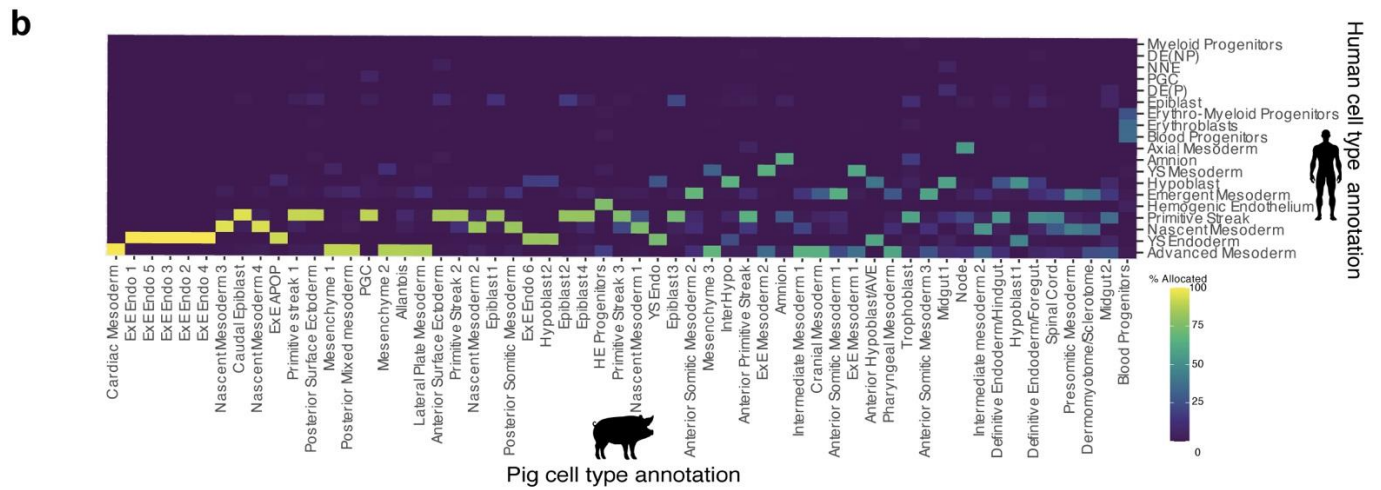
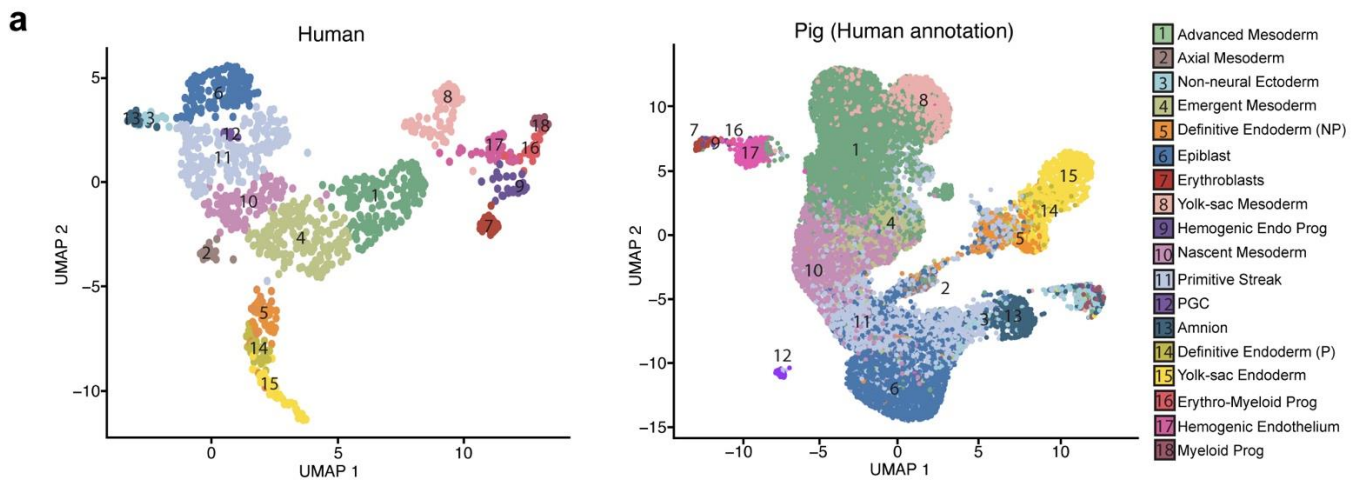
**Supplementary Fig. 3. Extra-embryonic cell types show lower correlation across mammals.** Heatmap showing Spearman correlation coefficient between the transcriptomes of matched cell types in pig, monkey and mouse embryos.



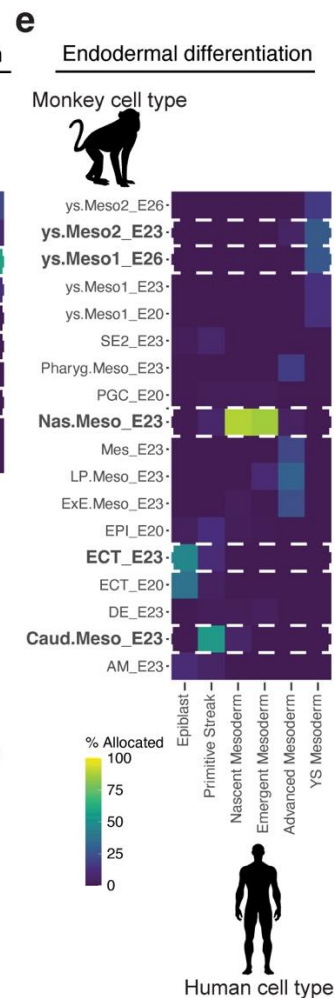
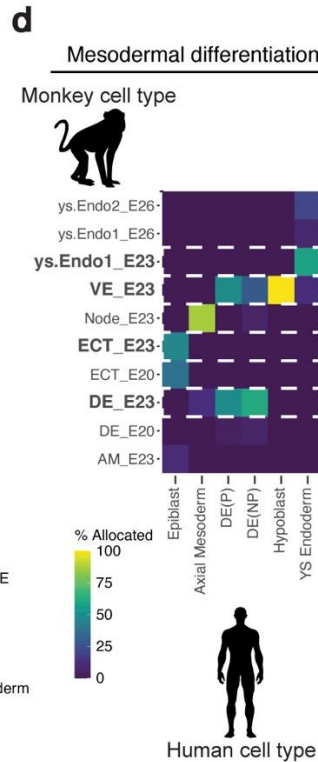
**Supplementary Fig. 4. a, Cell type-specific gene expression programs are highly conserved. a-h, Venn diagrams and heatmaps showing unique and overlapping cell type-specific genes in pig, monkey and mouse. The heatmaps illustrate the average fold change in gene expression within the cell type of interest, compared to the mean fold change across the remaining cell types for a specific species. APS, Anterior primitive streak.**



**Supplementary Fig. 5. Broad differences in early cell behavior are observable across mammals.** a, Plot of Cnet illustrating the enriched KEGG terms and associated genes found in both unique and shared expression profiles among pig, monkey, and mouse epiblast cells. b, Heat map of selected KEGG terms-associated gene expression in early pig, monkey and mouse cell types.

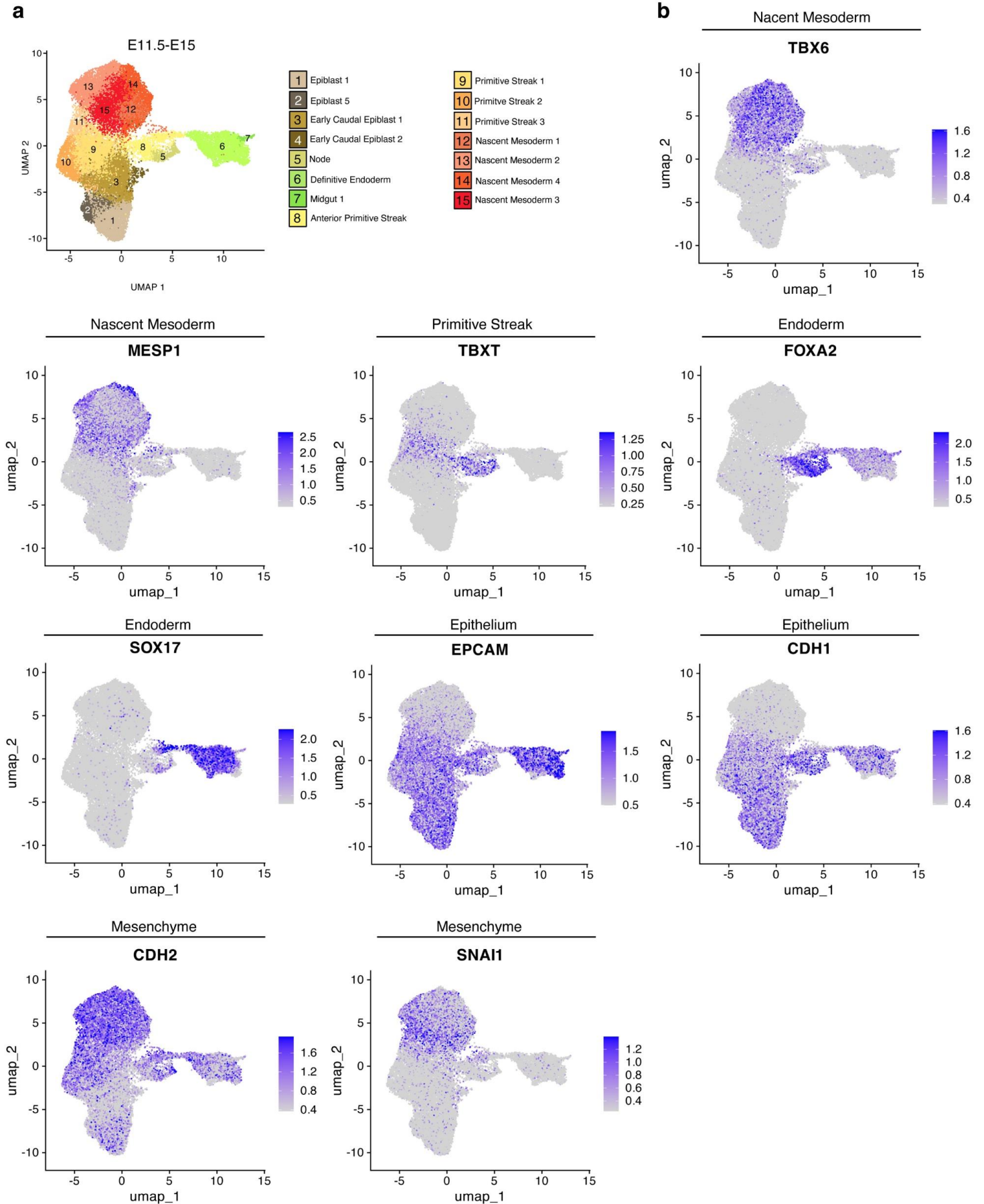


- 1 Cardiac Mesoderm
- 2 Caudal Epiblast
- 3 Mesenchyme 2
- 4 Mesenchyme 3
- 5 Intermediate Mesoderm 1
- 6 Nascent Mesoderm 2
- 7 Nascent Mesoderm 4
- 8 Lateral Plate Mesoderm
- 9 Nascent Mesoderm 1
- 10 Primitive Streak 2
- 11 Primitive Streak 3
- 12 Epiblast 1
- 13 Epiblast 2
- 14 Primitive streak 1
- 15 Epiblast 3
- 16 Anterior Primitive Streak
- 17 Node
- 18 Midgut 2
- 19 Definitive Endoderm/Foregut
- 20 Definitive Endoderm/Hindgut
- 21 Midgut 1
- 22 ExE Endo 3
- 23 ExE Endo 1
- 24 ExE Endo 2
- 25 InterHypo
- 26 Posterior Surface Ectoderm
- 27 ExE Endo 5
- 28 Hypoblast 2
- 29 Hypoblast 1
- 30 ExE Endo 4
- 31 ExE APOF
- 32 Anterior Hypoblast/AVE
- 33 YS Endo
- 34 ExE Endo 6
- 35 Epiblast 4
- 36 Spinal Cord
- 37 Anterior Surface Ectoderm
- 38 Amnion
- 39 Trophoblast
- 40 Cranial Mesoderm
- 41 Mesenchyme 1
- 42 PGC
- 43 Dermomyotome/Sclerotome
- 44 Blood Progenitors
- 45 Haematoendothelial Progenitors
- 46 Presomitic Mesoderm
- 47 Anterior Somitic Mesoderm 3
- 48 Anterior Somitic Mesoderm 2
- 49 Anterior Somitic Mesoderm 1
- 50 Intermediate mesoderm 2
- 51 Pharyngeal Mesoderm
- 52 Posterior Somitic Mesoderm
- 53 Allantois
- 54 ExE Mesoderm 2
- 55 ExE Mesoderm 1
- 56 Posterior Mixed mesoderm
- 57 Nascent Mesoderm 3

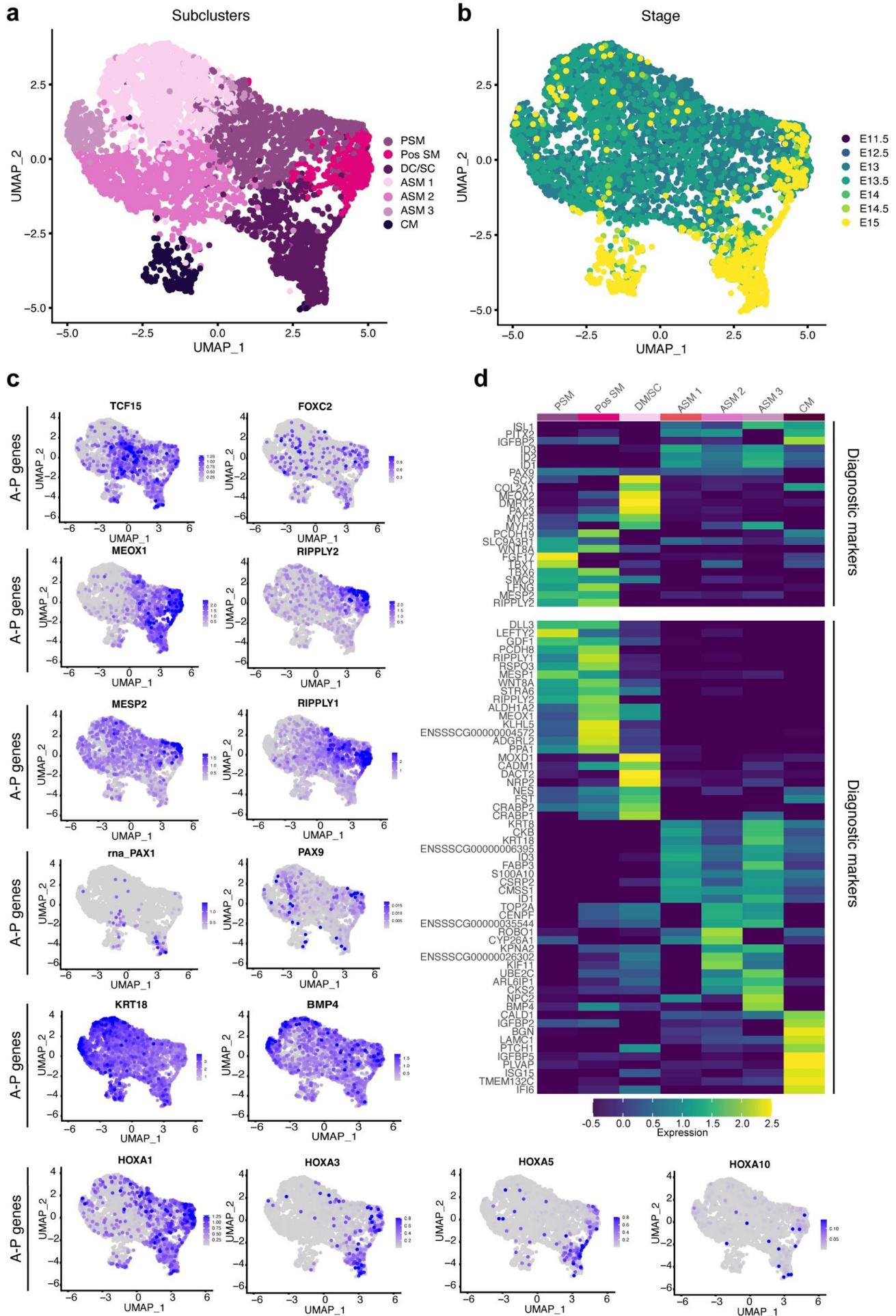


**Supplementary Fig. 6. Homochronicity among monkey and human cell type differentiation.** **a**, UMAPs showing E17-19 human embryo cell types<sup>7</sup> and pig E11.5 to E15 with mouse annotations after reciprocal PCA-based projection onto the human dataset. P, primitive; NP, non-primitive. **b**, Heat map showing the percentage of pig cells of a given cell type allocated as a particular human cell type after label transfer. **c**, As with a with a, pig annotations after reciprocal PCA-based projection onto the human dataset. **d**, Heat maps showing the percentage of human mesodermal cells<sup>7</sup> allocated to a particular monkey cell identity<sup>6</sup> after label transfer. **e**, As with d, except with endodermal cell types. 100% means that all cells of a given cell type were predicted to be analogous to the cell identity in the queried organism.

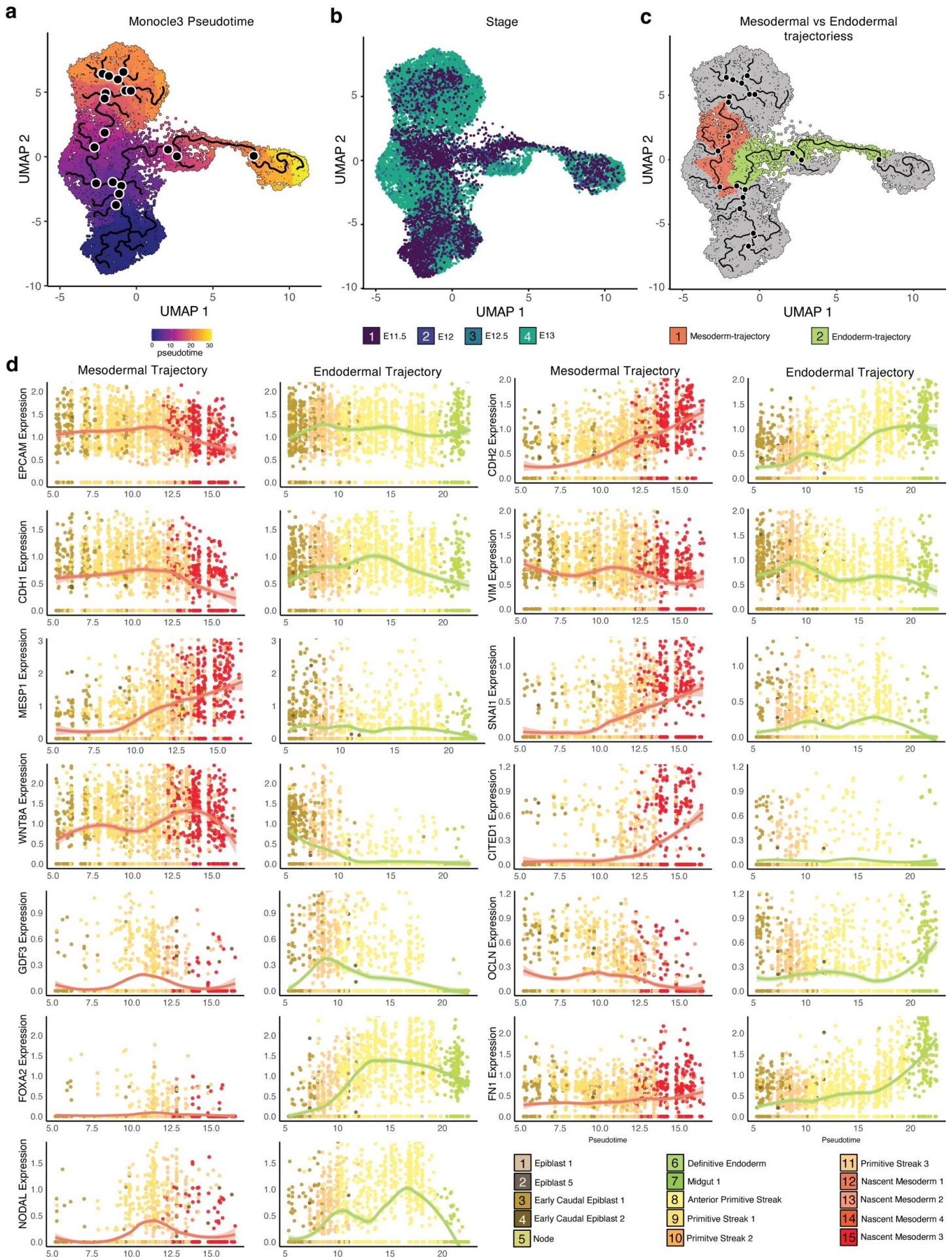




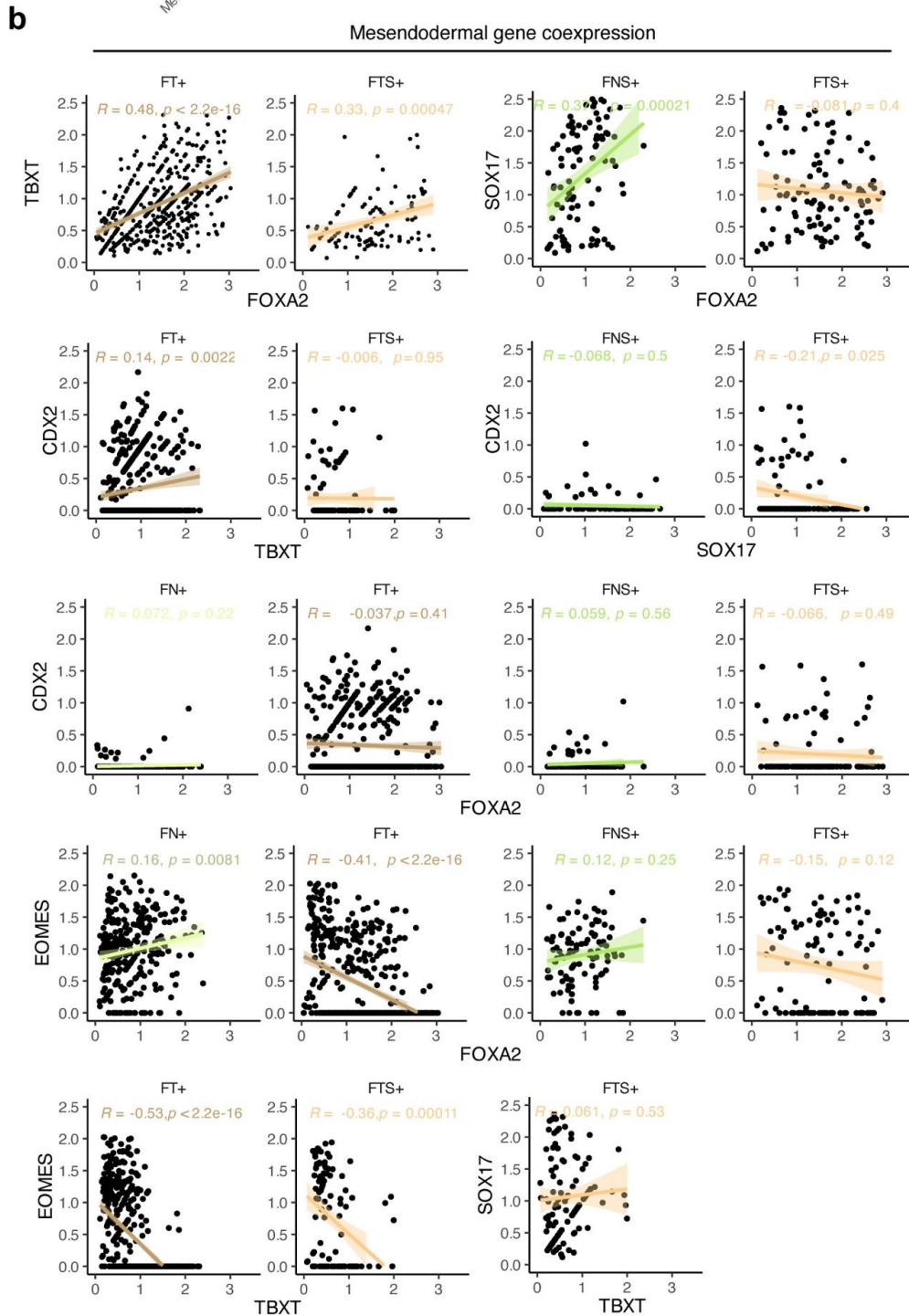
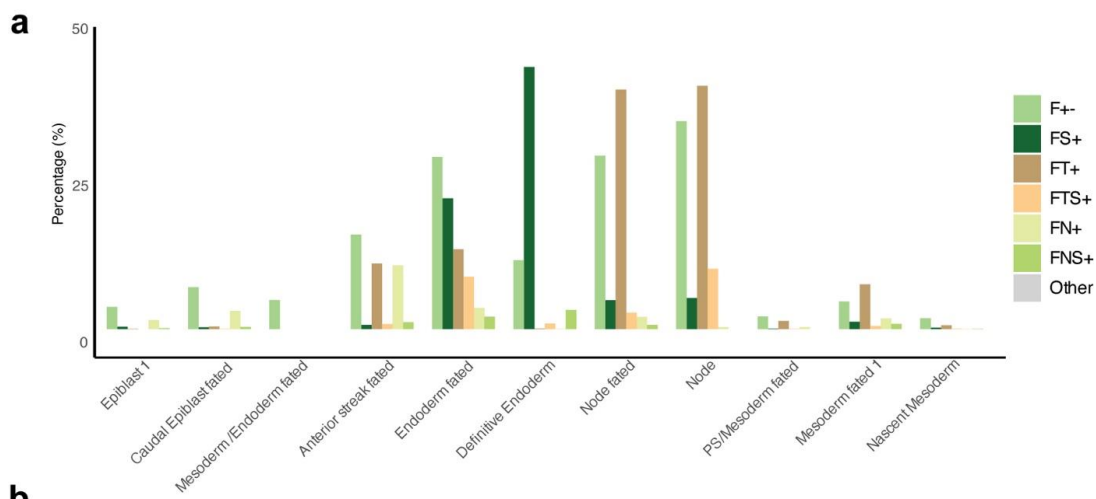
**Supplementary Fig. 7. Feature plots of Mesendodermal, Endodermal and EMT-related genes.** **a**, UMAP plot from Fig. 4 showing Epiblast, Primitive Streak, Anterior Primitive Streak, Node, Nascent Mesoderm and Endoderm sub-clusters coloured by cluster. **b**, Feature plots coloured by normalised gene expression of selected marker genes relating to Fig. 4. Feature plot scales were determined by the 5<sup>th</sup> and 95<sup>th</sup> percentiles.



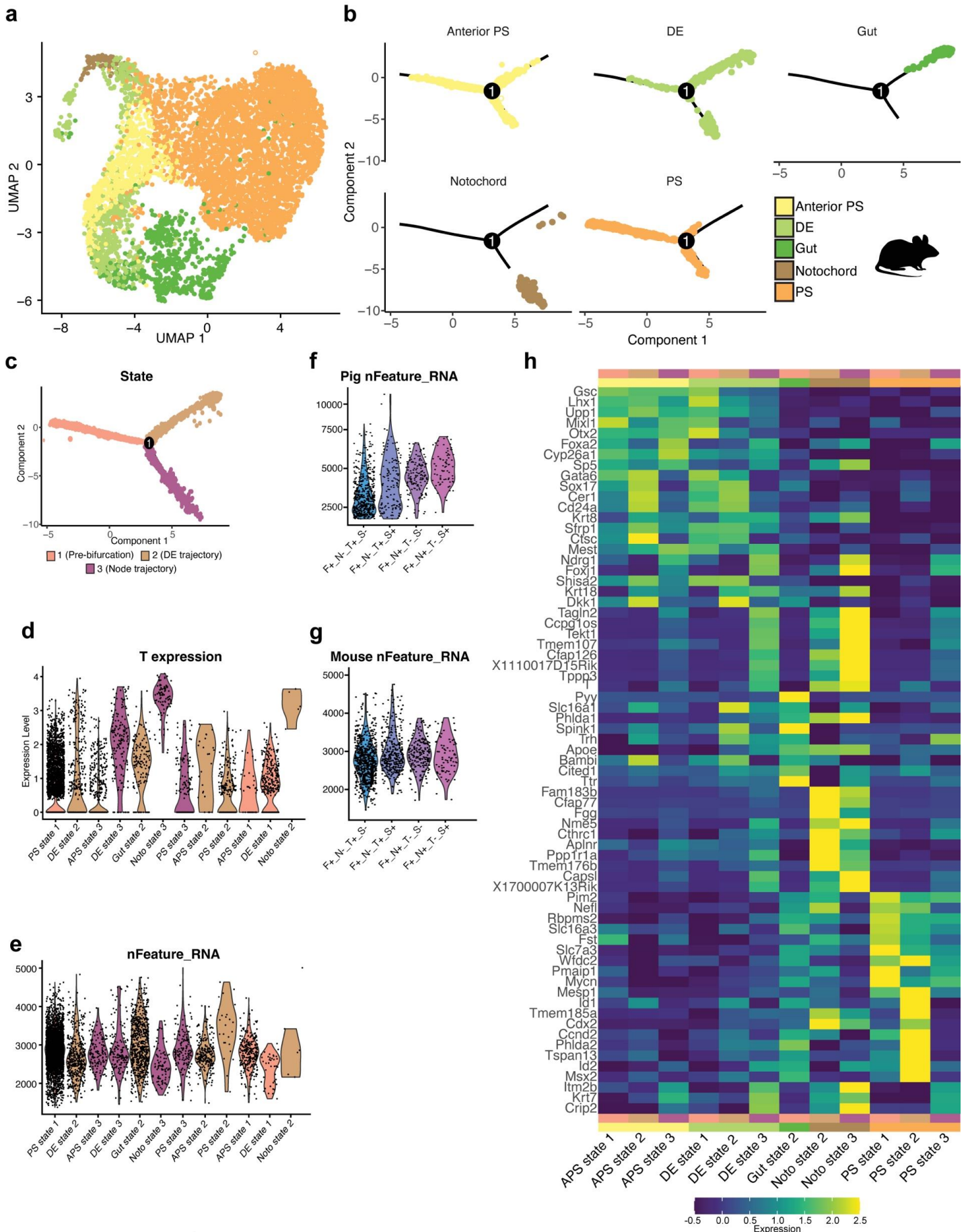
**Supplementary Fig. 8. Somitogenesis in porcine embryos.** **a**, UMAP plot showing sub-clustering of Presomitic and Somitic mesoderm clusters. 7 subclusters were identified: Presomitic mesoderm (PSM), Posterior somitic mesoderm (Pos SM), Dermomyotome/Sclerotome (DC/SC), Anterior Somitic Mesoderm 1, 2 and 3 (ASM1, ASM2, ASM3), as well as Cranial Mesoderm (CM). **b**, UMAP plot showing cells from **a**. Cells are coloured by embryonic day. **c**, UMAP feature plots of key markers of somite development. Feature plots coloured by normalised gene expression. Feature plot scales were determined by the 5<sup>th</sup> and 95<sup>th</sup> percentiles. **d**, Heat map illustrating the scaled average expression of selected genes. Shown are diagnostic marker genes for each subcluster and the top 10 most significant marker genes.



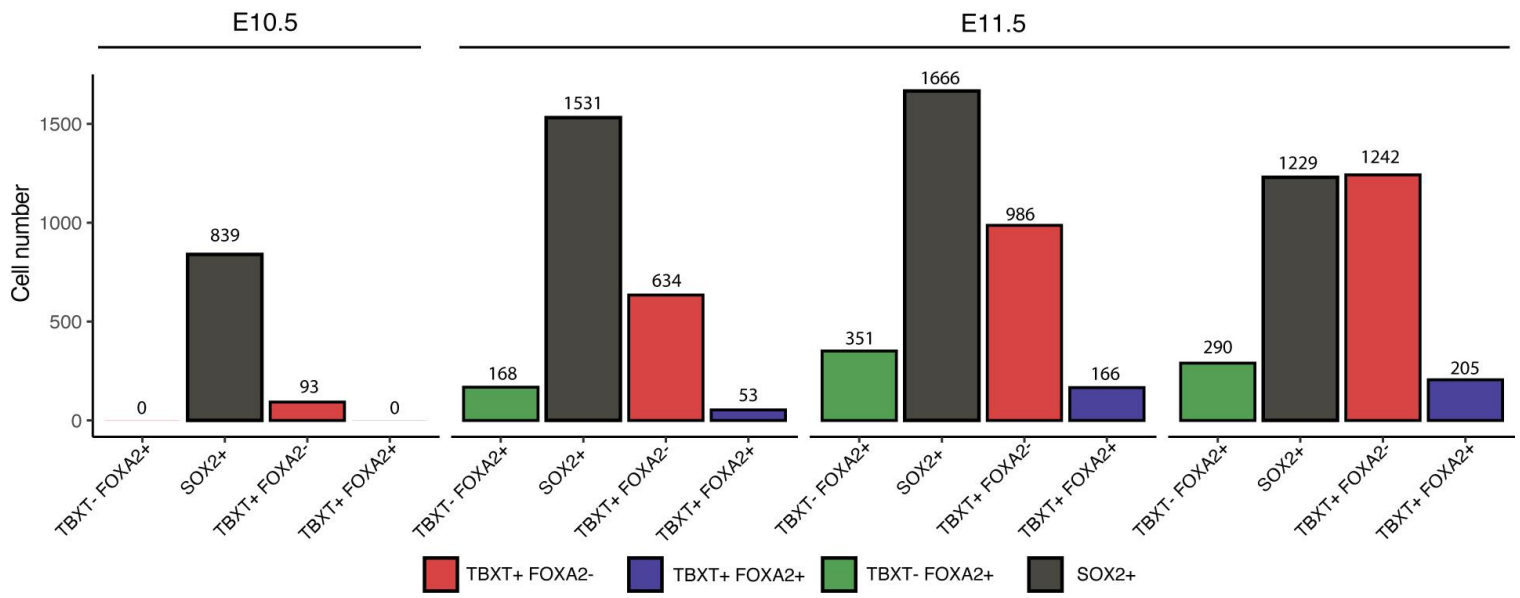
**Supplementary Fig. 9. Dynamics of EMT-related genes in mesendoderm and endoderm fated cells.** **a**, UMAP plot showing Epiblast, Primitive Streak, Anterior Primitive Streak, Node, Nascent Mesoderm and Endoderm sub-clusters from Fig.5a with reversed graph embedding trajectories projected on top using Monocle3, coloured by Monocle3 pseudo time. Black nodes mark trajectory branching points. **b**, UMAP coloured by developmental time point. **c**, UMAP with reversed graph embedding trajectories projected on top, cells associated with nodes going in the nascent mesoderm trajectory are coloured in red, while those associated with the endoderm trajectory are coloured in green. **d**, Scatter plots showing gene expression and Monocle3 pseudo time values in mesodermal (left) vs endodermal (right) fated cells.



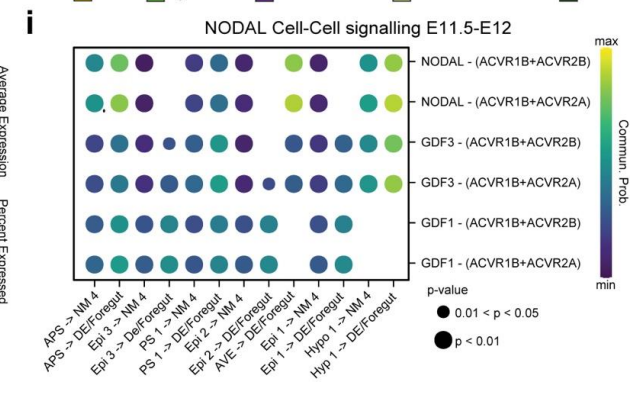
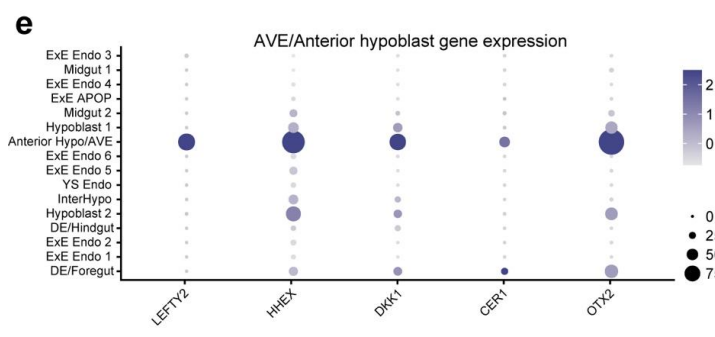
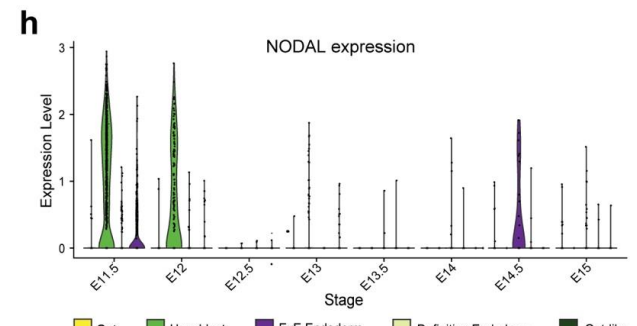
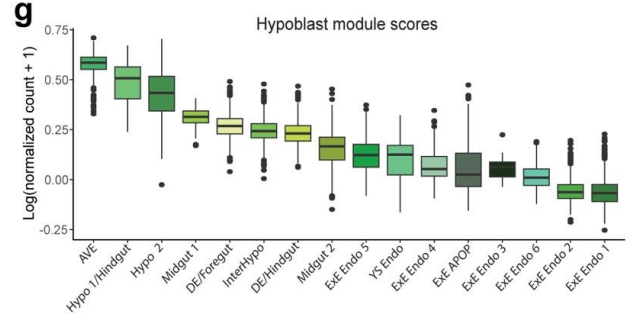
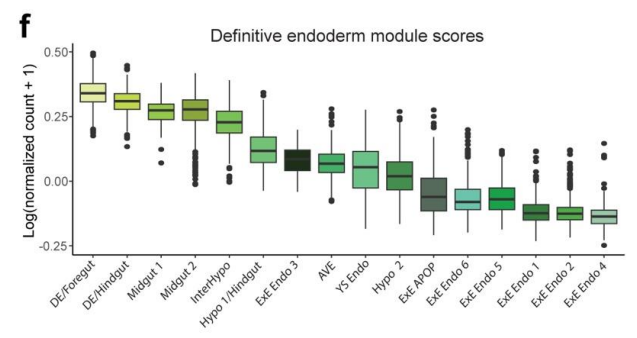
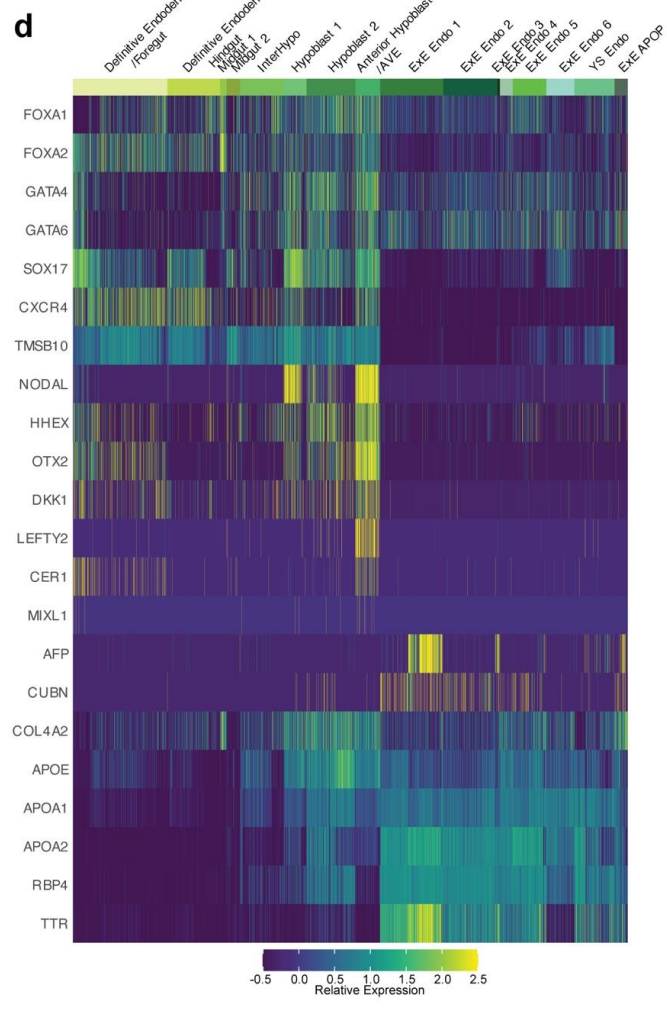
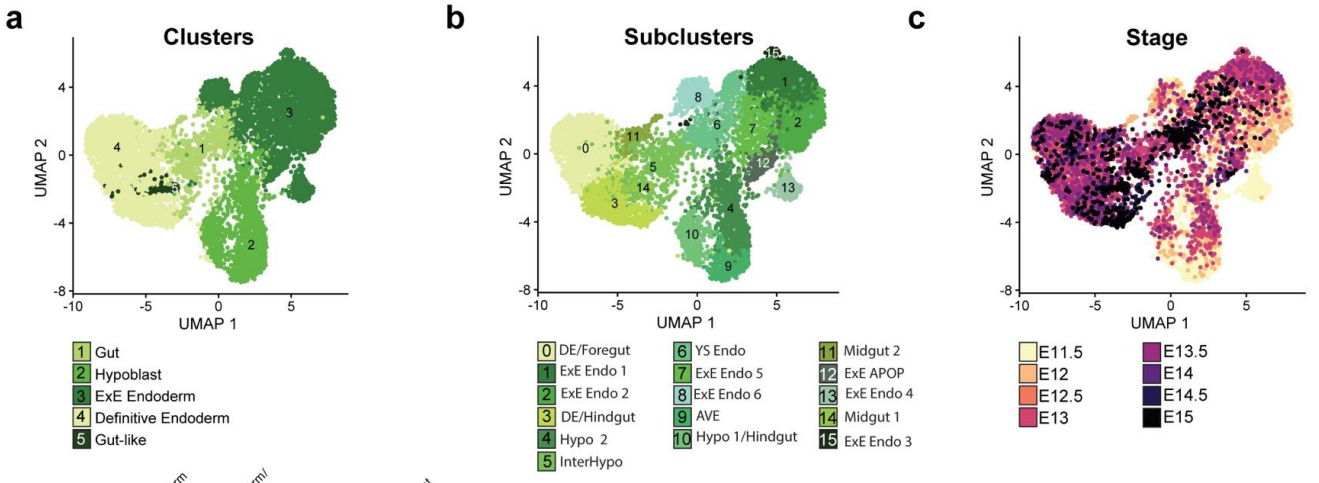
**Supplementary Fig. 10. Mesodermal and endodermal genes are mutually exclusive. a**, Bar graphs illustrating cells categorised by FOXA2, NANOG, TBXT and SOX17 co-expression status and their relative percentage abundance in E11.5-13 epiblast, caudal epiblast, PS, APS, node and DE fated cells. **b**, Scatter plots showing mesoderm and endoderm associated transcription factor co-expression of F/N/T/S categorised cells from Fig 4.a & Extended Data Fig.8a. Correlation coefficient (R) and adjusted p-value (p) following Pearson correlation test indicated.



**Supplementary Fig. 11. Endoderm forms from TBXT-low cells in mice.** **a**, UMAP plot showing PS, APS, Notochord, DE and Gut clusters (5,860 cells) from Pijuan-Sala et al<sup>2</sup>, coloured by cell type. **b**, Force-directed graph of the cells illustrated in **a**, after pseudo-temporal ordering using Monocle2. Force-directed graphs are split by cell type. **c** Force-directed graph of the cells illustrated in **a** and **b** coloured by state. State 1 represents cells that have not committed to notochord or DE/Gut fates. State 2 represents cells moving toward a DE/gut fate. State 3 represents cells moving toward a Notochord fate. **d**, Violin plots of T expression in the cell types defined in **c**. **e**, Violin plots showing the number of mapped genes in each cell type defined in **d**. **f**, Violin plots showing the number of mapped genes in the cell types of interest defined in Figs 5-7. **g**, Violin plots showing the number of mapped genes in the mouse-equivalent cell types shown in **f**. Both mouse and pig DE/Node progenitors show a comparable number of mapped genes. **h**, Heat map illustrating the scaled mean expression of the top 10 most significant DEGs between the cell types from **c-e**.

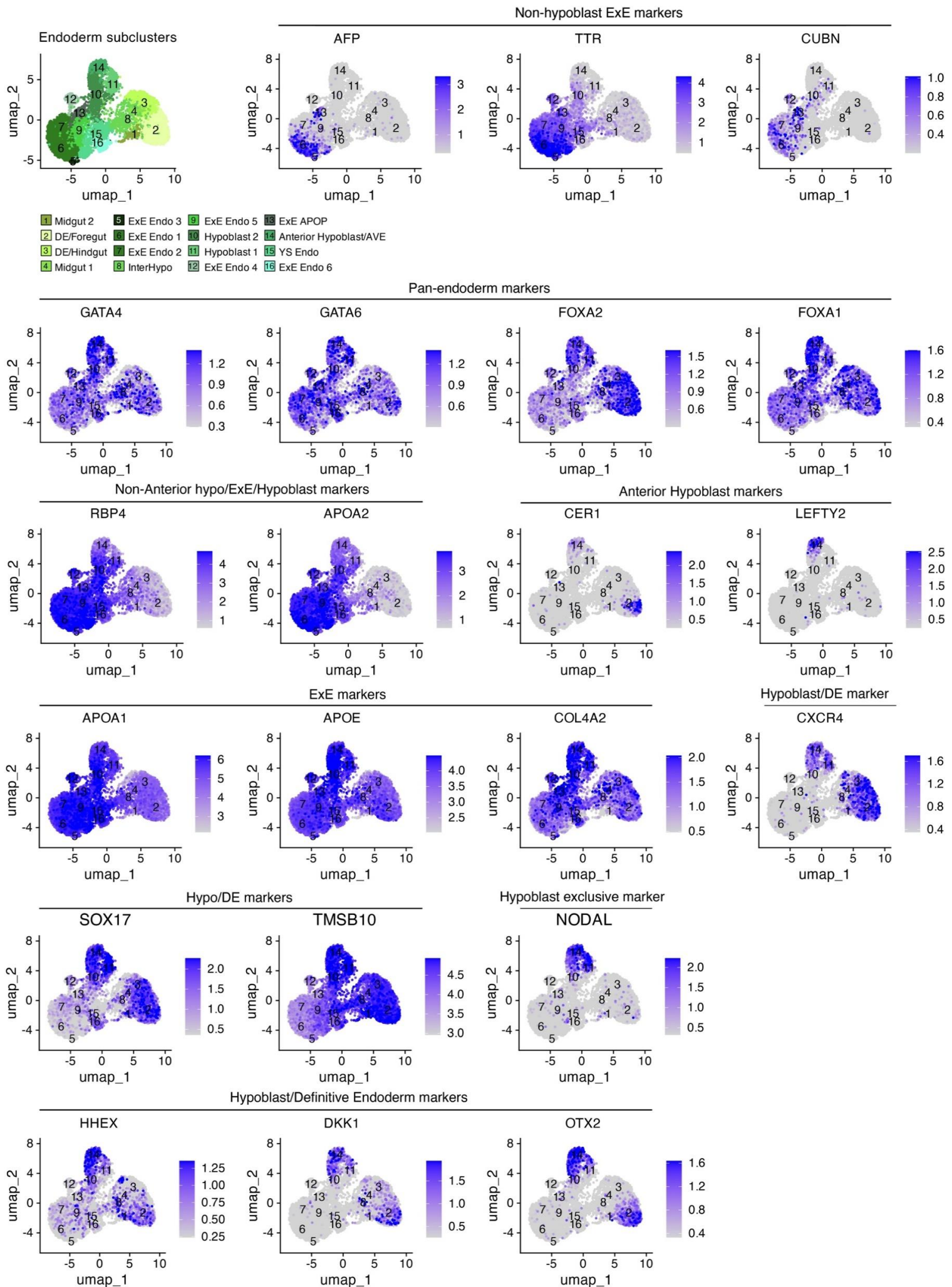


**Supplementary Fig. 12.** Quantification of 3D segmentation of whole mount porcine embryo immuno-staining from figure 5c. Segmentation via Trackmate and Stardist. Thresholding was achieved via Otsu's method within Fiji.

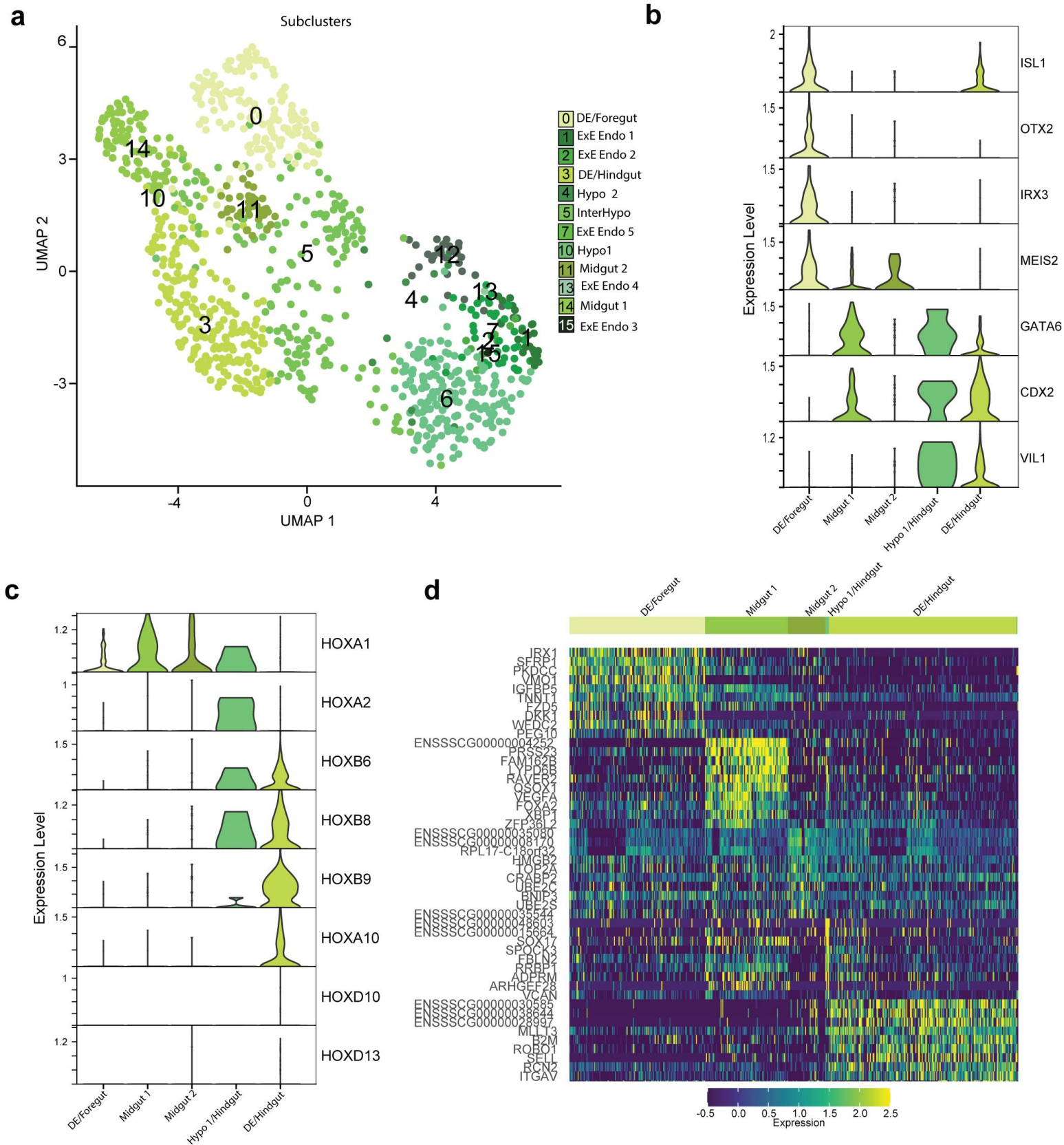


**Supplementary Fig. 13. The hypoblast is a major source of signaling and contributes to DE.** **a**, UMAP plot showing DE, gut-like, gut, hypoblast and ExE Endoderm subclusters coloured by cell-type. **b**, UMAP plot showing embryonic and extraembryonic endodermal sub-clusters coloured by cluster. **c**, UMAP plot showing cells from a&b coloured by timepoint, showing stage specificity of clusters. **d**, Heat map illustrating the cluster-scaled mean expression of the top 10 most significant newly discovered marker genes for each subcluster. **e**, Dot plot showing the average expression and percentage of cells expressing AVE/Anterior hypoblast related genes. **f&g**, Box and whisker plots showing the respective DE and hypoblast module scores in each subcluster. Modules were defined as the top 100 most significant markers of the E11.5 DE or hypoblast (See Methods). **h**, Violin plot showing the expression of *NODAL* at each time point in the clusters from a. **i**, Predicted *NODAL* cell-cell signalling between selected cell types in E11.5-E12 embryos.

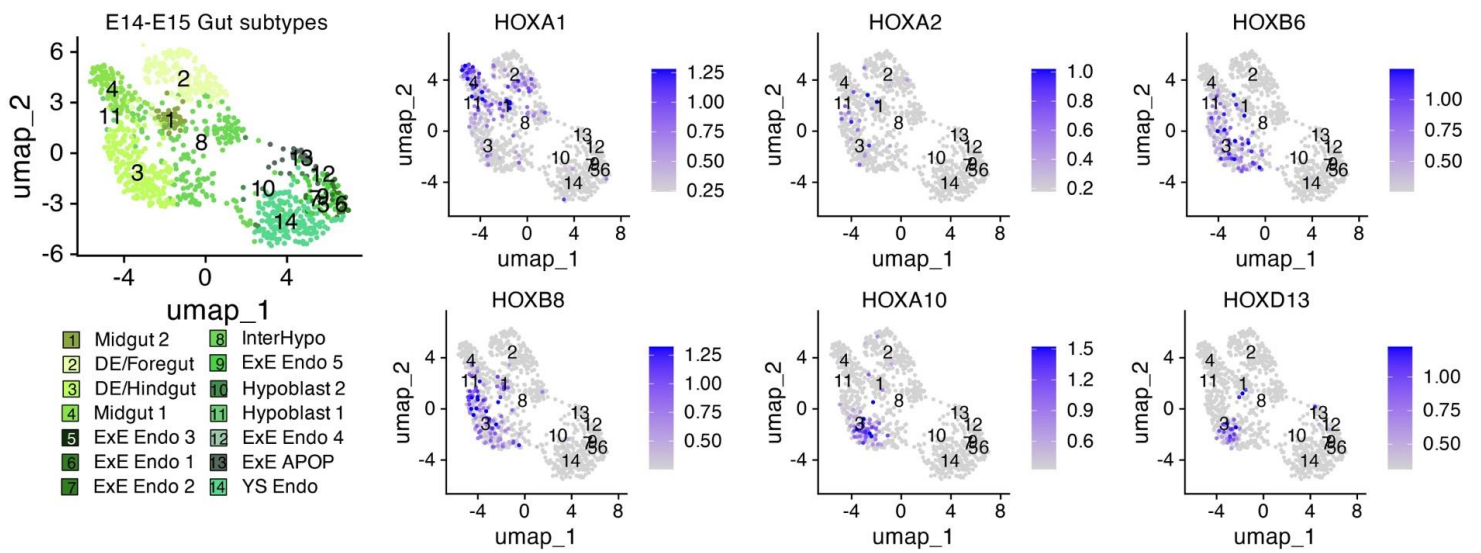




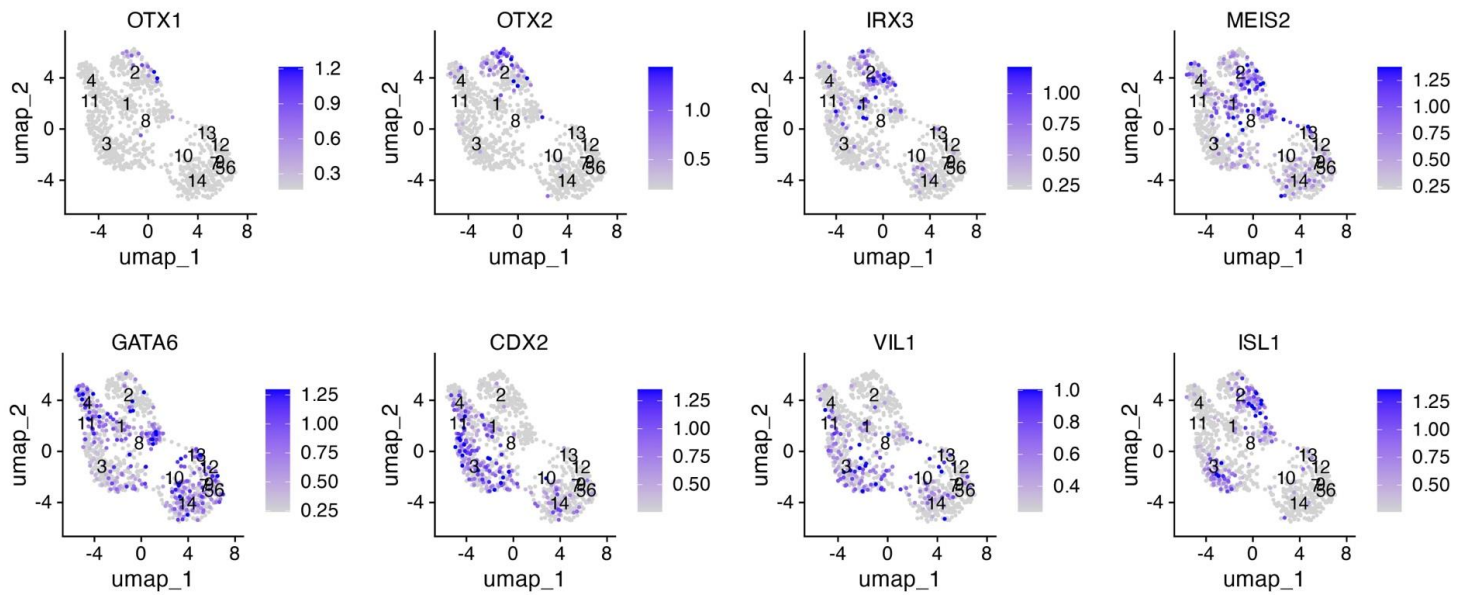
**Supplementary Fig. 14. Feature plots to verify Endodermal subtypes.** UMAP plot (top left) showing embryonic and extraembryonic endoderm sub-clusters coloured by cluster and plots coloured by normalised gene expression of selected marker genes relating to Supplementary Figure 15. Feature plot scales were determined by the 5<sup>th</sup> and 95<sup>th</sup> percentiles.



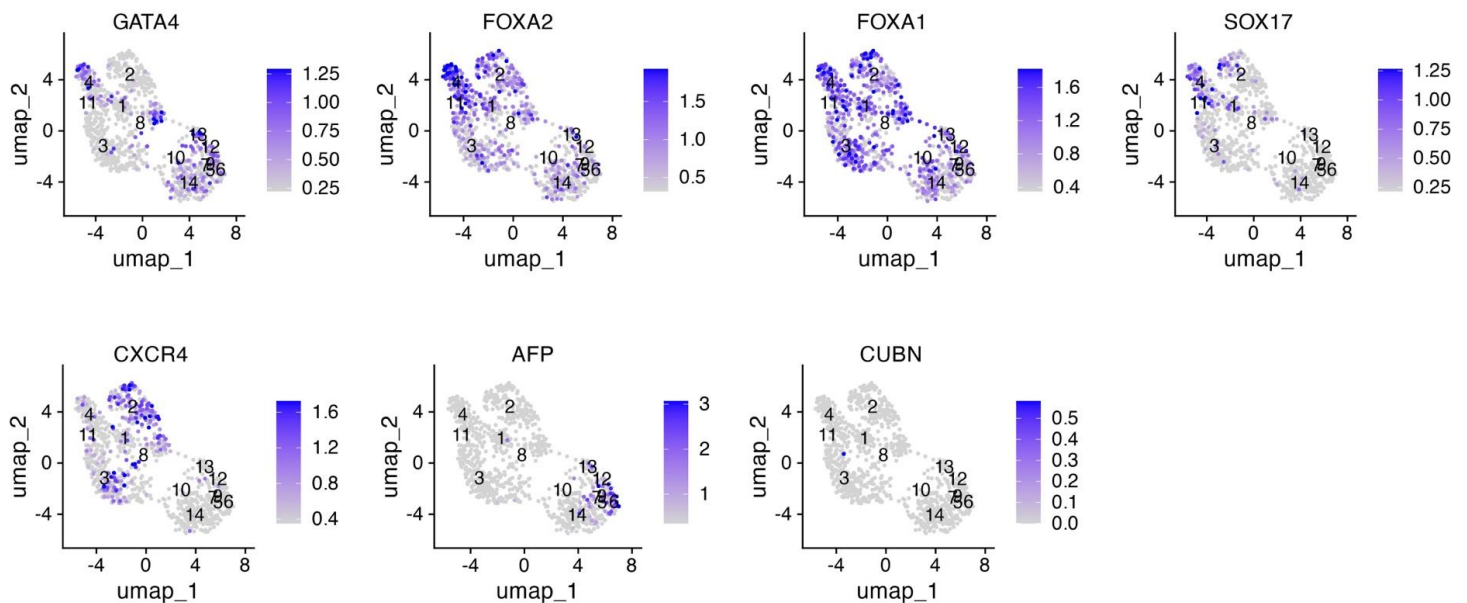
**Supplementary Fig. 15. Mature gut types emerge between E14-E15.** **a**, UMAP plot showing embryonic and extraembryonic sub-clusters from Fig 10 from E14-E15. **b**, Violin plots of gene expression showing late endodermal subclusters show specific expression of gut markers described by Nowotschin et al<sup>42</sup>. **c**, Violin plots of Hox gene expression showing the distinct anterior-posterior positioning of gut subclusters. **d**, Heat map illustrating the scaled mean expression of the top 10 most significant markers of each gut subcluster.



## Foregut/Midgut/Hindgut markers



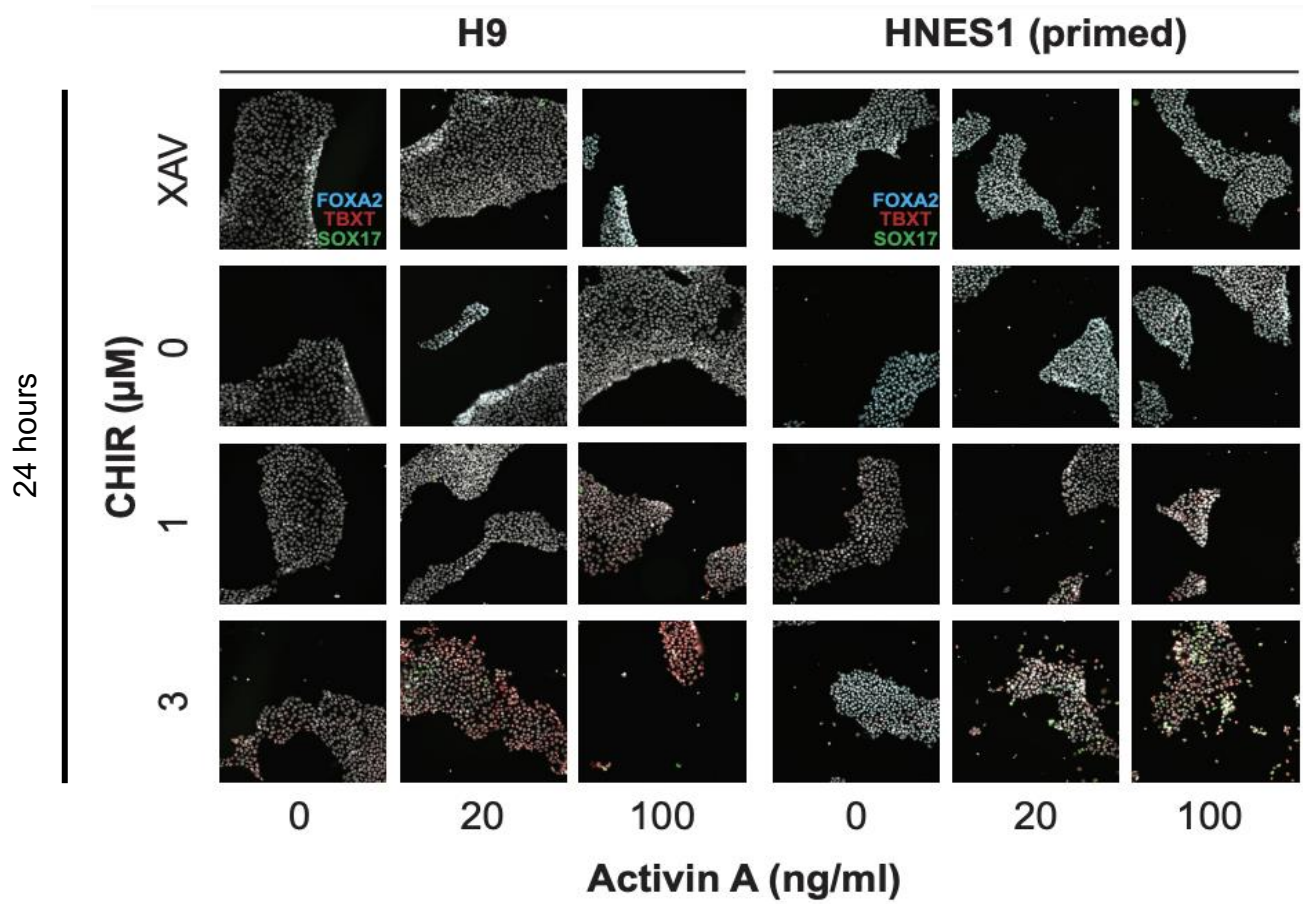
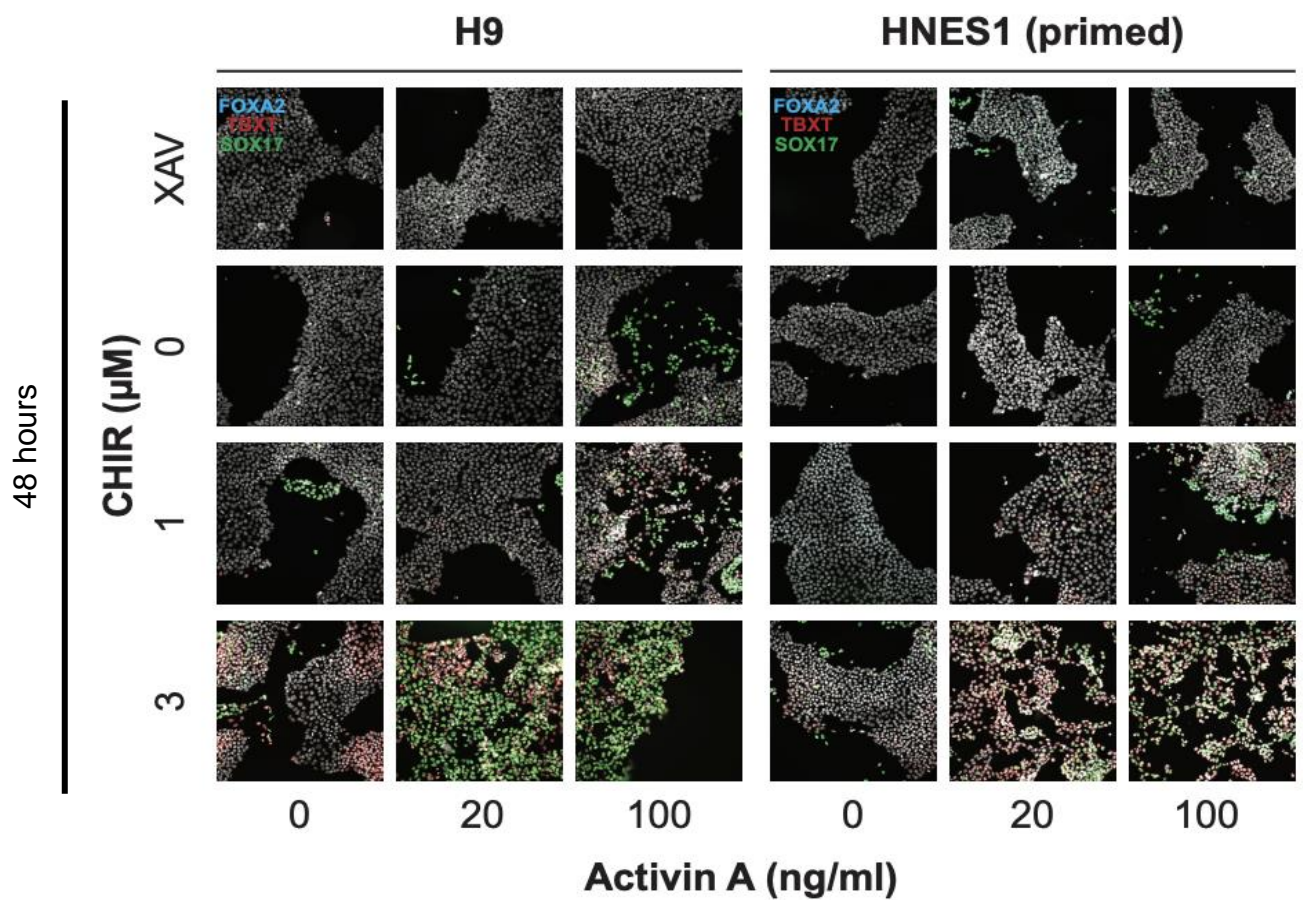
## Endodermal markers



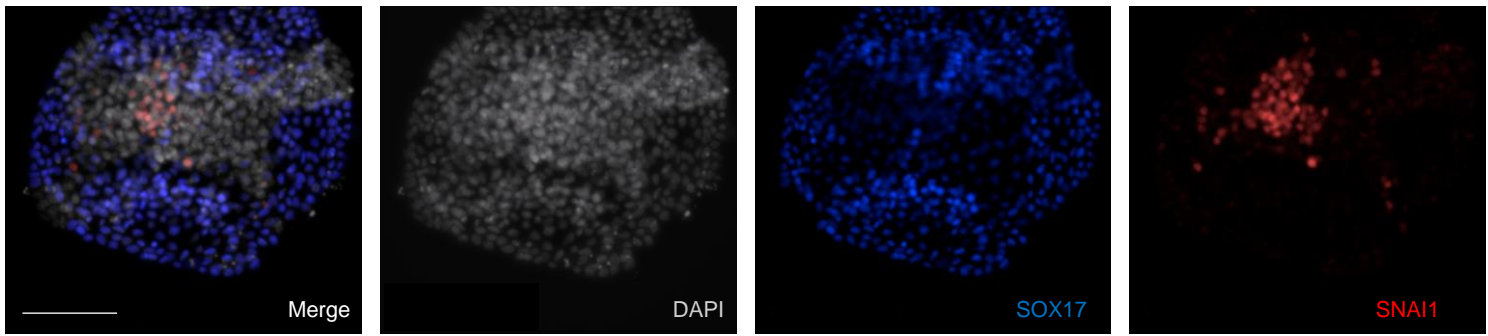
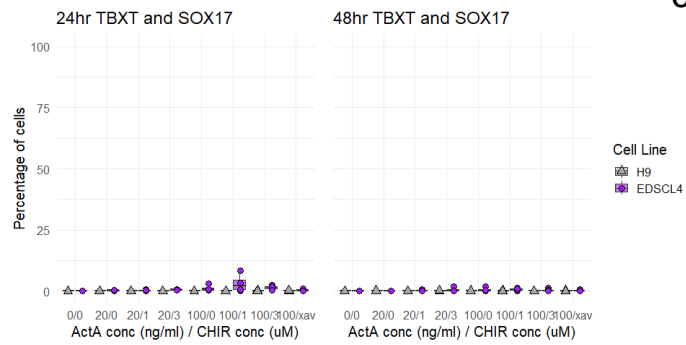
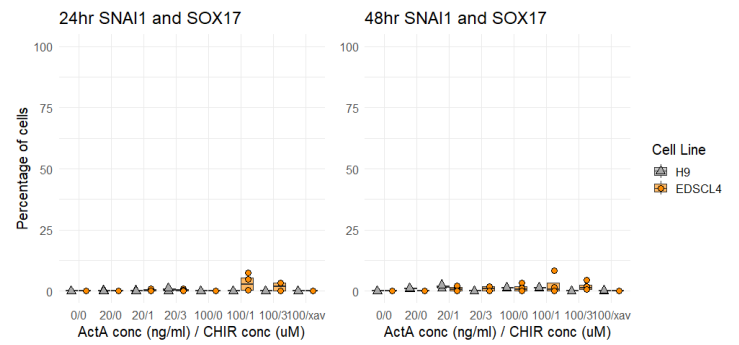
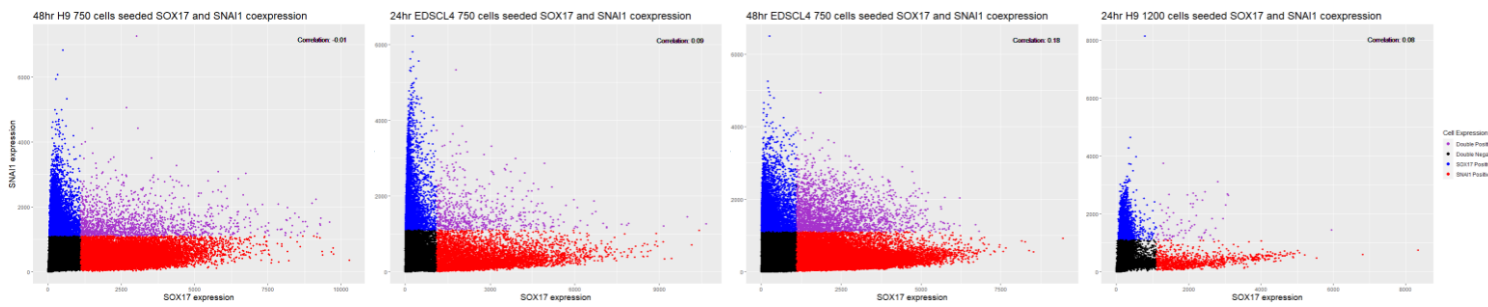
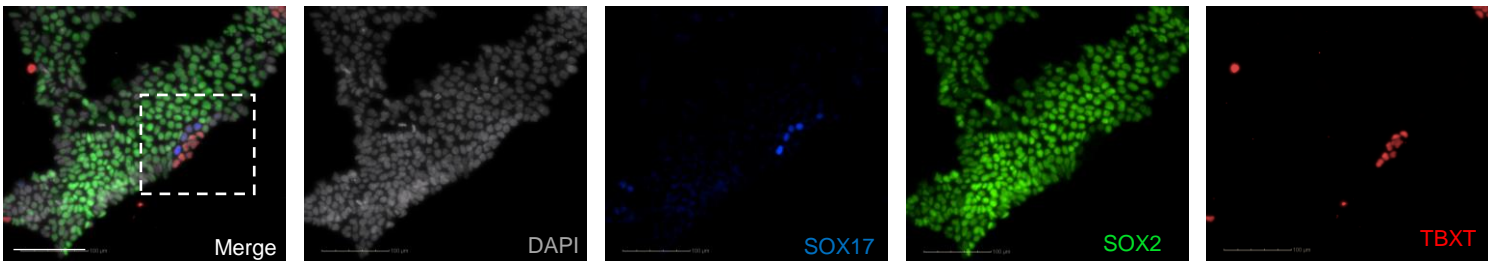
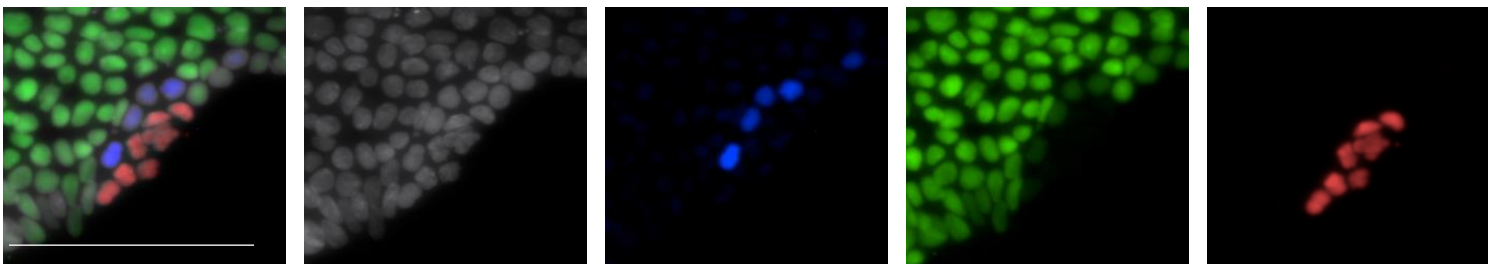
**Supplementary Fig. 16. Feature plots to verify Gut subtypes.** UMAP plot showing E14-E15 gut sub-types identified in Extended Data Fig. 11 and UMAP plots coloured by normalised gene expression of selected marker genes.



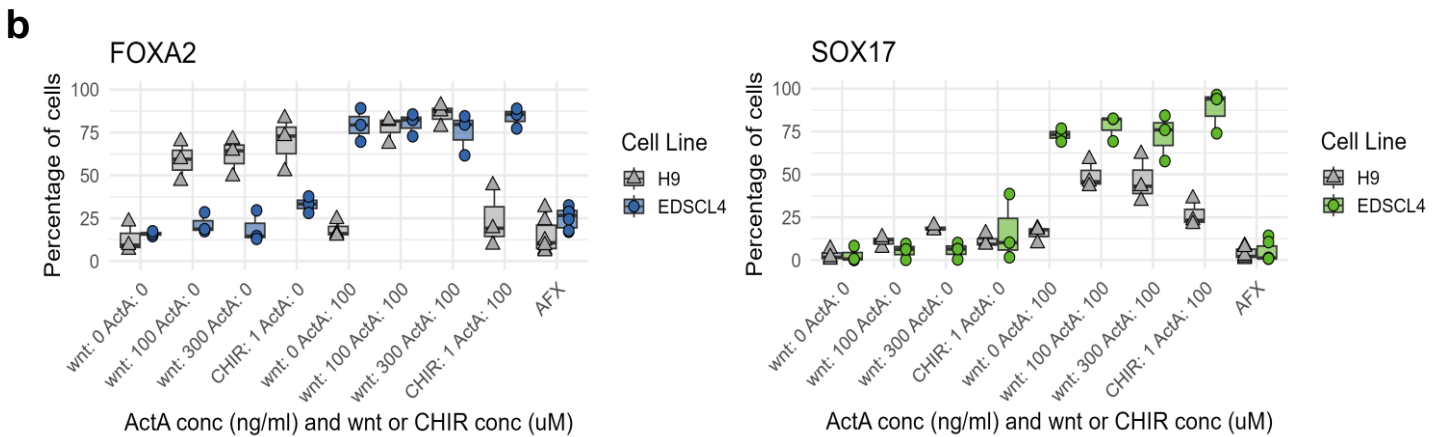
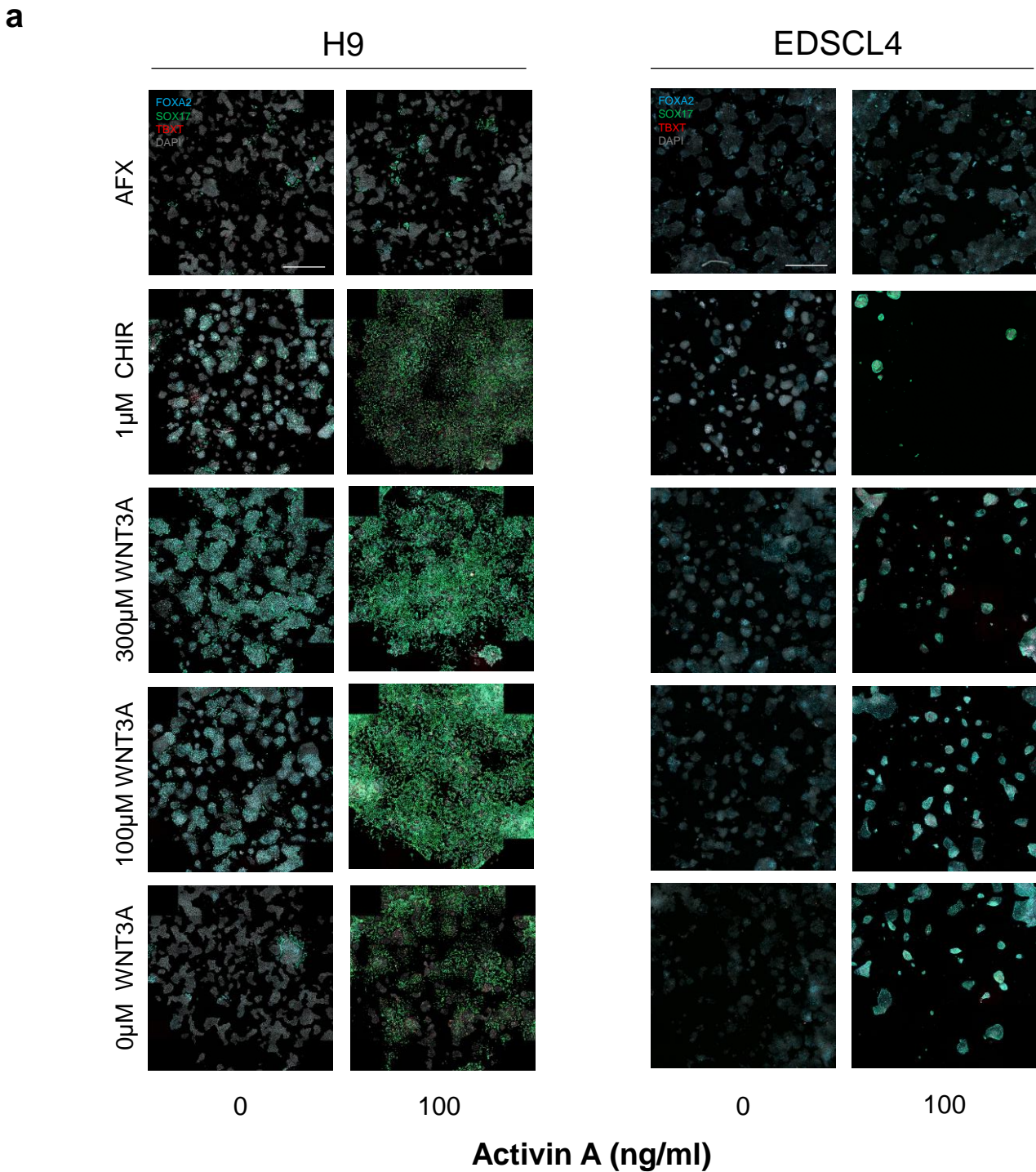
**Supplementary Fig. 17. WNT and Nodal signaling in early pig gastrula embryos.** **a**, Predicted canonical WNT3, WNT3A and WNT8A cell-cell signalling in E11.5-12 Epiblast, Hypoblast and PS clusters. **b**, Schematic summarising **a**. **c**, Predicted Hedgehog cell-cell signalling between selected cell types in E11.5-14 embryos.

**a****b**

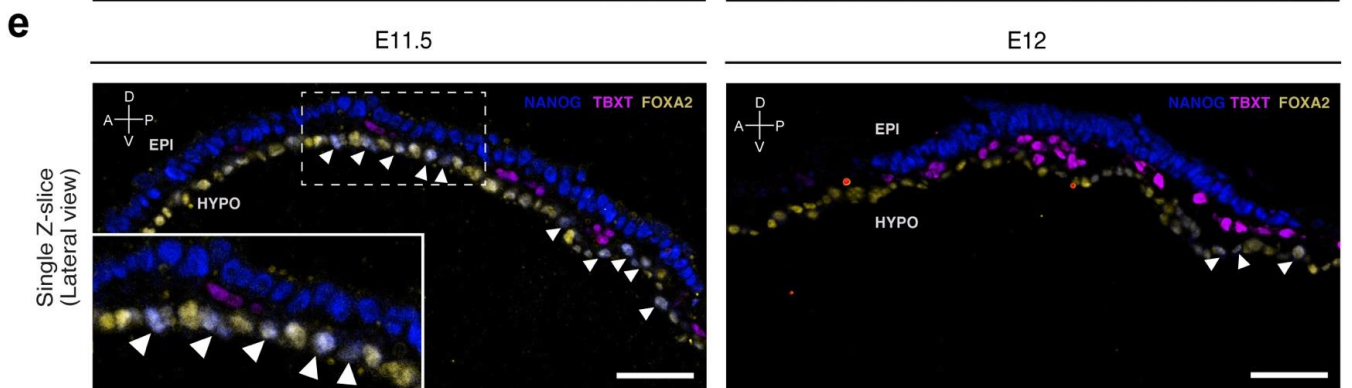
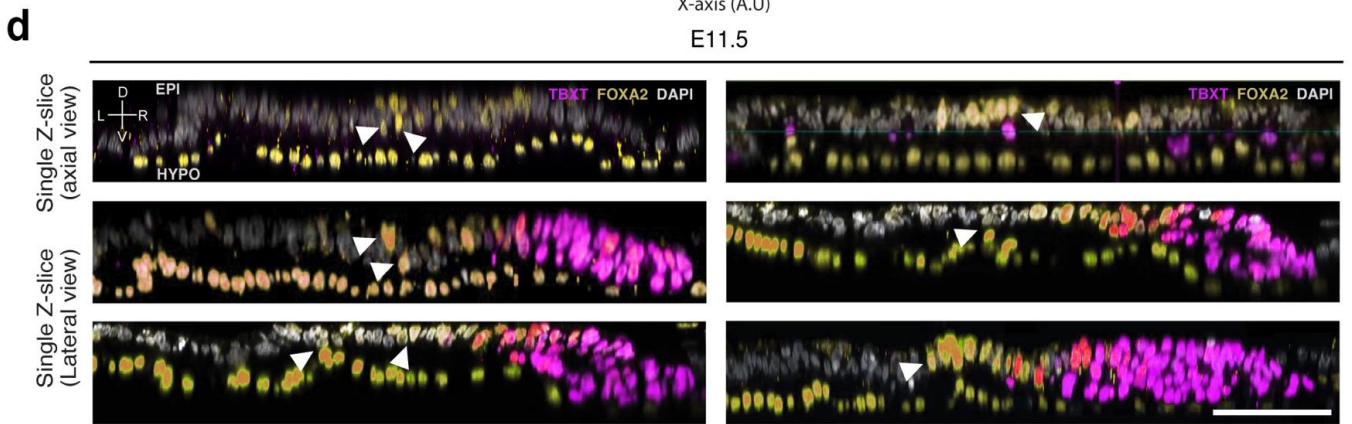
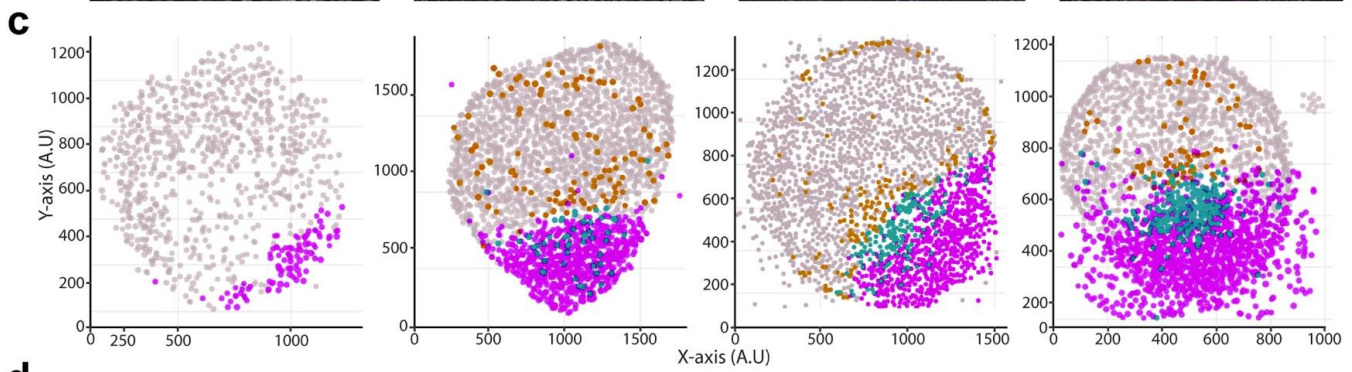
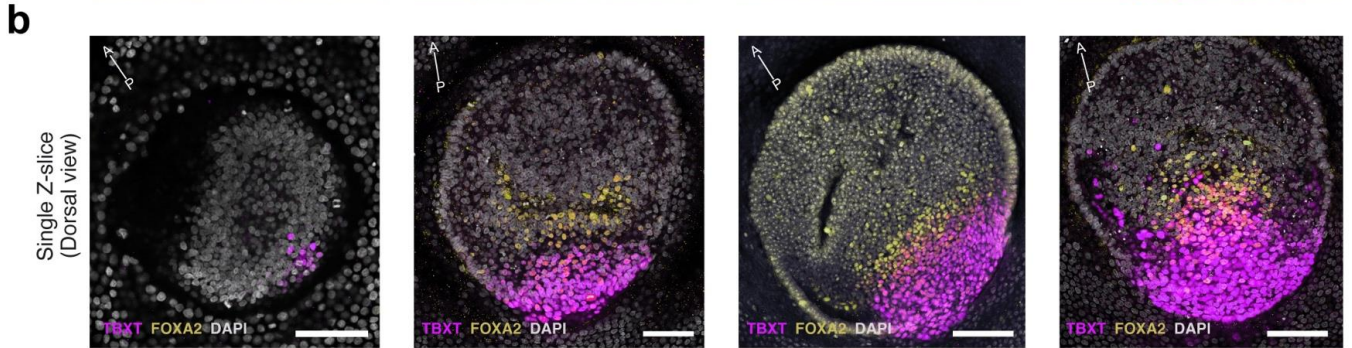
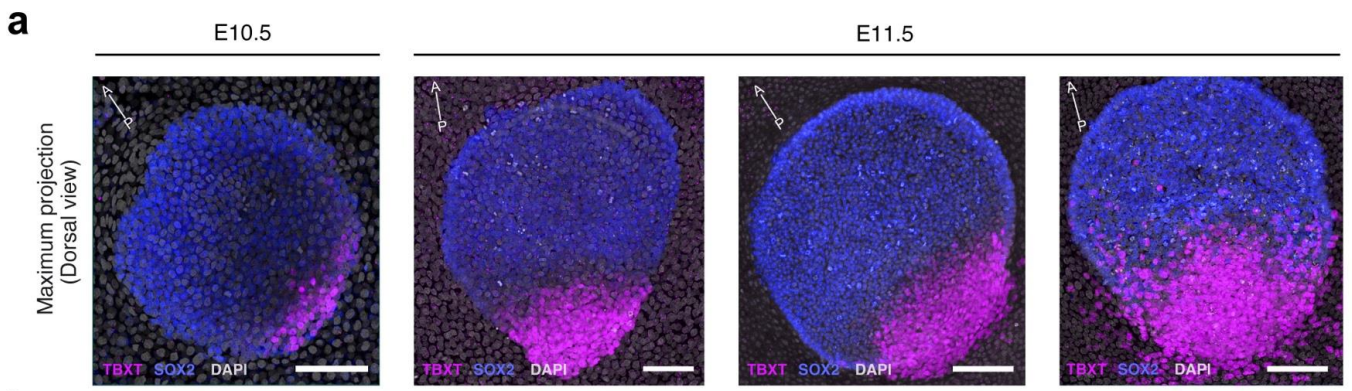
**Supplementary Fig. 18. Co-expression of mesodermal and endoderm markers.** Representative images of H9 and HNES1 hESC lines after treatments with differing amounts of XAV, CHIR and/or Activin A. **a**, 24hrs after treatment **b**, 48hours after treatment. FOXA2 (blue), TBXT (red) and SOX17 (green) expression.

**a****b****c****d****e****f**

**Supplementary Fig. 19. Co-expression of endoderm markers.** **a**, Representative images of EDSC4 after 48 hours of induction with ActA (100ng/ml) and CHIR (1 $\mu$ M). Scale bar: 100 $\mu$ m. **b** Quantification of SOX17 and TBXT co-expression across the 2D experiments. **c**, Quantification of SOX17 and SNAI1 co-expression across the 2D work. **d**, Selection of correlation graphs between SNAI1 and SOX17. Correlation value calculated via Pearson's correlation. **e**, representative images of EDSC4 after 8 hours of induction with ActA (100ng/ml) and CHIR (1 $\mu$ M). Scale bar: 100 $\mu$ m. **f** enlargement of ROI (dotted box in e). Scale bar: 100 $\mu$ m.

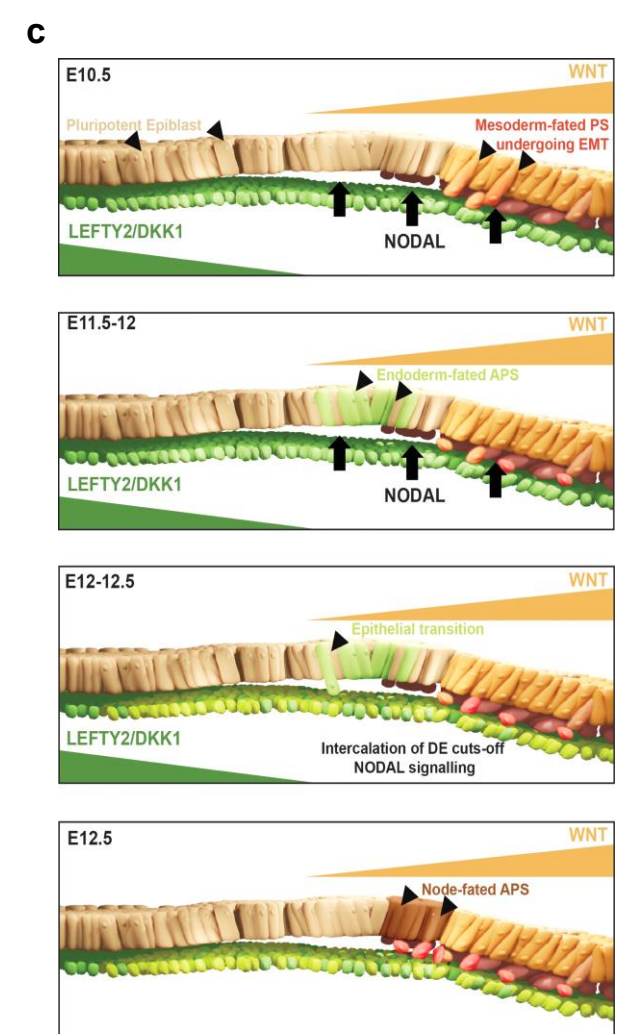
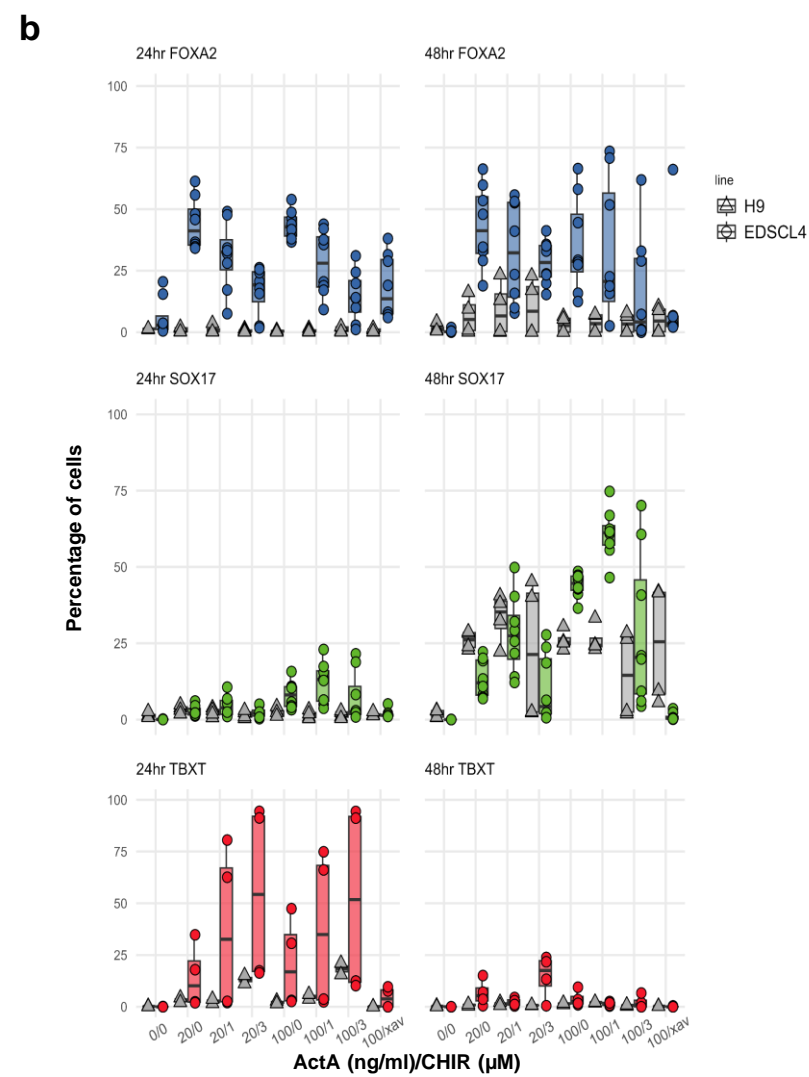
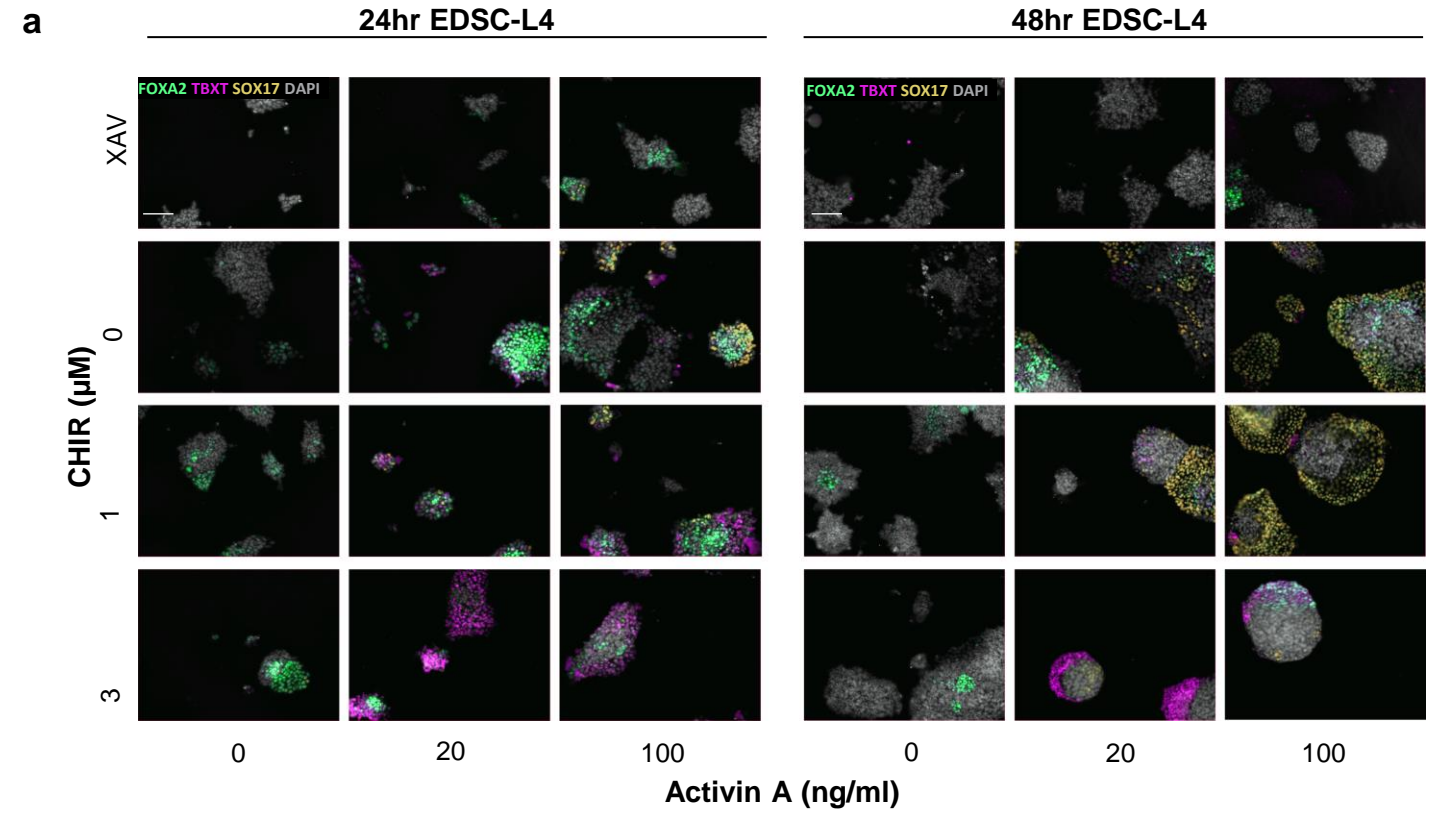


**Supplementary Fig. 20. Effect of Activin A and WNT on endoderm formation in 2D differentiation. a,** Overview of each well within human (H9) and porcine (EDSCL4) 2D WNT3A and Activin A differentiation experiments. Green: SOX17, blue : FOXA2, grey : DAPI, red : TBXT. Scale bar: 1000um. **b,** Quantification of SOX17 and FOXA2 after 48 hrs incubation with WNT3A, CHIR and ActA (n=3).





**Supplementary Fig. 21. FOXA2 and TBXT domains are spatially separated (Colour-blind friendly).** **a**, Maximum intensity projection (Dorsal view) of E10.5 to E11.5 porcine embryos showing TBXT and SOX2 expression. E11.5 embryos are ordered left to right by age. E, Embryonic day. **b**, Single z-slice of the embryos shown in **a**, showing FOXA2 and TBXT expression. Scale bar: 50µm. **c**, *In Silico* representations of embryos following 3D segmentation of embryos from **a** and **b**. **d**, Axial and lateral reconstructed sections of embryos stained for FOXA2 and TBXT. Epiblast layer is oriented above the hypoblast/DE layer. White arrowheads indicate FOXA2+ TBXT- cells that are spatially separated from the TBXT domain. Scale bar: 50µm. **e**, Lateral sections of E11.5 and E12 embryos, showing expression of NANOG, FOXA2 and TBXT. Epiblast layer is oriented above the hypoblast/DE layer. White arrowheads indicate NANOG+ cells. Scale bar: 50µm.



**Supplementary Fig. 22. Effect of Activin A and WNT signaling in pig EpiSC and hESC (Colour-blind friendly).** **a**, Representative images depicting differentiation conditions for EDSCL4. Images were captured using an Operetta CLS high-throughput microplate imager. Scale bar: 200  $\mu\text{m}$ . **b**, Box plots summarizing 2D differentiation experiments. Data is normalised to well background signal and results represent biological triplicates. Centre line represents median, minima and maxima hinges represent the 25<sup>th</sup> and 75<sup>th</sup> percentiles respectively. **c**, Proposed model of epiblast-DE differentiation in pig embryos.