

Fig. S1. Sample quality control and single cell characterization of the first trimester human maternal-fetal interface

(A) Summary of workflow used for collection, processing, and single cell analysis of placental villous cells. MFI, Maternal-Fetal Interface, 7AAD: 7-Aminoactinomycin D, FSC-A: Forward Scatter Area. Violin plots showing the distribution of the raw number of (B) features, (C) counts, and (D) the percent mitochondrial DNA content remaining post quality control and data pre-processing captured per cell during sample sequencing. Samples 1-6 (teal) represent the placental tissue samples sequenced by our lab, samples 7-11 (red) represent the placental tissue samples acquired from ArrayExpress, and samples 12-15 (grey) represent the decidual tissue samples acquired from ArrayExpress (E-MTAB-6701). (E) Scatter plot showing the distribution of captured cells across principal components 1 and 2 following dimensional reduction in principal component space. Cell positions are calculated based on cell embeddings calculated during dimensional reduction by principal component analysis. Cells are color coordinated first by tissue type and second by data source, with teal representing those cells originating from placental tissue samples sequenced in this study (GSE174481) and red representing those cells originating from placental and decidual tissue samples acquired from Vento-Tormo et al (E-MTAB-6701). (F) UMAP projection of the captured maternal fetal interface cells, stratified by gestational age. Early (left panel): 9803 early first trimester cells from each identified cluster captured between 6-7 weeks gestational age, before the onset of maternal placental vascularization. Late (right panel): 40,987 late first trimester cells from each identified cluster captured between 9-12 weeks gestational age, after the onset of maternal placental vascularization. Cell cluster names and colors correspond to those in Figure 1B. (G) Heatmap showing the average expression of known immune, non-immune, and trophoblast gene markers, as well as proliferation gene markers in each identified cluster. Cell cluster names and colors correspond to those in Figure 1B. (H) Stacked bar plot showing the percent distribution of cells in each cluster with gene signatures representing a cell in the G₁ phase of proliferation, the G₂ or mitotic (G₂M) phases of proliferation, or the synthesis (S) phase of the cell cycle. Cell cluster names, colors, and abbreviations correspond to those in Figure 1B. (I) UMAP plots showing gene expression in red of a mesenchymal cell gene marker (*VIM*), an immune cell gene marker (*PTPRC*, encoding the CD45 protein), endothelial cell gene markers (*CD34* and *VCAMI*), and a fibroblast gene marker (*ACTA2*). (J) UMAP plots showing gene expression in red of known trophoblast gene markers used to subset and purify first trimester human trophoblasts from other cells within the maternal fetal interface.

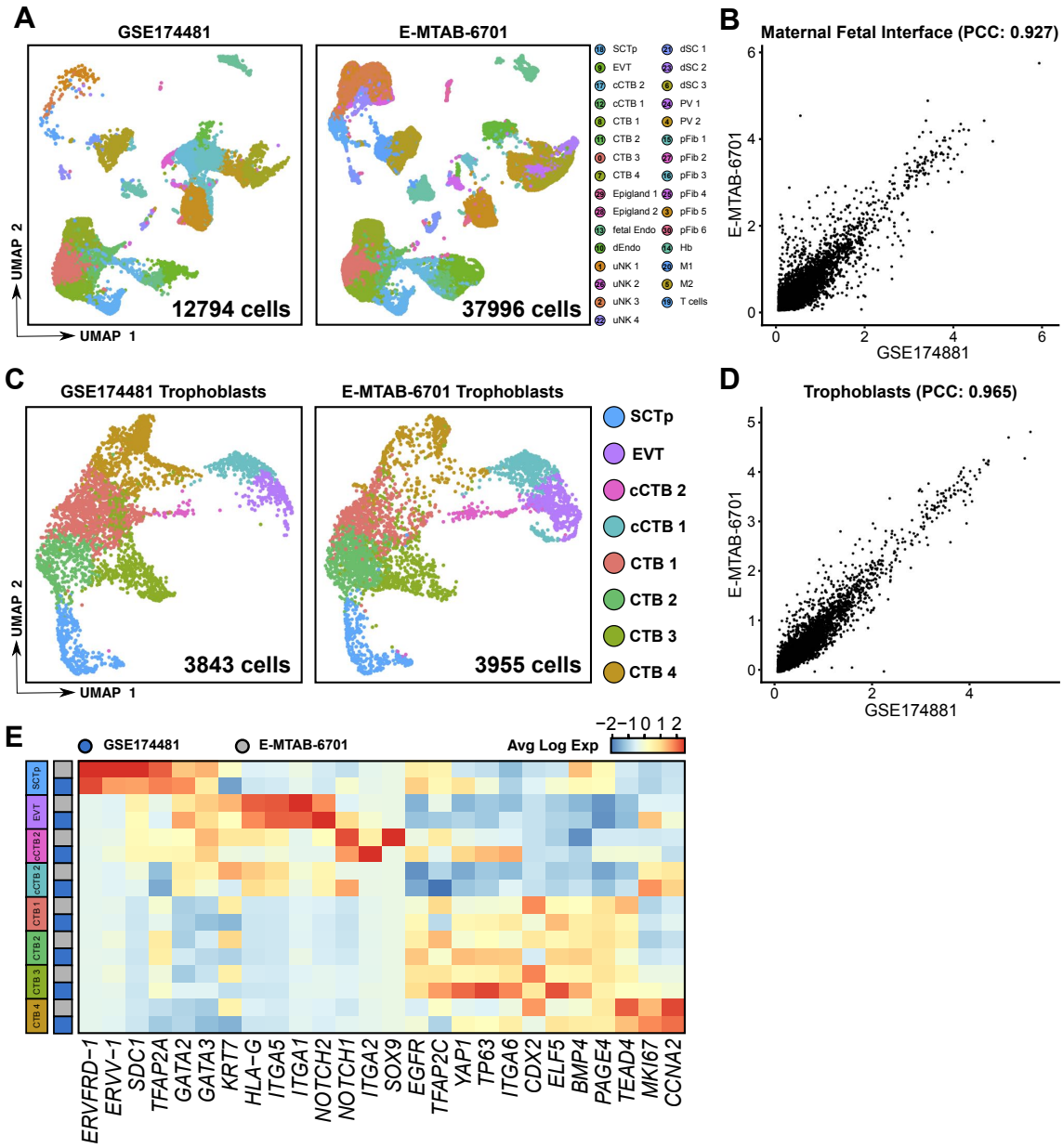


Fig. S2. GSE174481 and E-MTAB-6701 dataset comparison

(A) UMAP projection of the captured maternal fetal interface cells, stratified by data source. GSE174481 (left panel): scRNA-seq data of 12,794 first trimester cells. E-MTAB-6701 (right panel): scRNA-seq data of 37,996 first trimester cells. Cells are color coordinated by identity and abbreviations are identical to those in Figure 1A.

(B) Scatter plot comparing GSE174481 pseudo-bulk gene expression data (x-axis) against E-MTAB-6701 pseudo-bulk gene expression data (y-axis) generated from all captured maternal fetal interface cells. PCC: Pearson's correlation coefficient.

(C) UMAP projection of the captured human chorionic villi trophoblasts, stratified by data source. GSE174481 (left panel): scRNA-seq data of 3,843 first trimester trophoblasts. E-MTAB-6701 (right panel): scRNA-seq data of 3,955 first trimester trophoblasts. Cells are color coordinated by identity and abbreviations are identical to those in Figure 1A.

(D) Scatter plot comparing GSE174481 pseudo-bulk gene expression data (x-axis) against E-MTAB-6701 pseudo-bulk gene expression data (y-axis) generated from all captured human chorionic villi trophoblasts. PCC: Pearson's correlation coefficient.

(E) Heatmap showing the average expression of known trophoblast and proliferation gene markers in all trophoblast states, split by data source. Cell cluster names and colors correspond to those in (C).

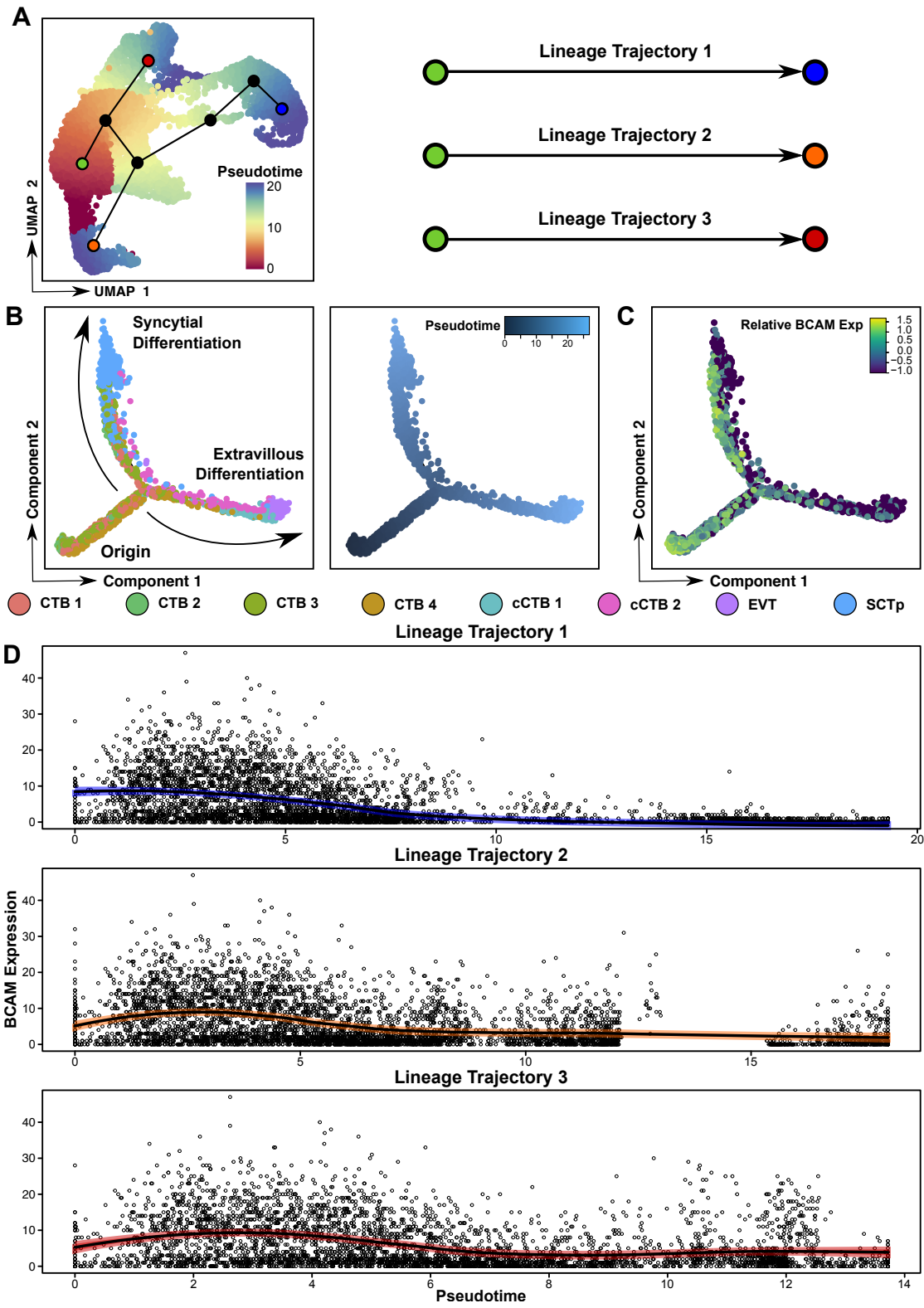


Fig. S3. Slingshot and Monocle2 pseudotime quantification supports identification of a trophoblast progenitor cell origin with the CTB2 state

(A) Lineage trajectories, calculated using Slingshot, projected over the human trophoblast UMAP plot. Cells are color coordinated by estimated Slingshot pseudotime values. Three lineage trajectories are defined with trajectory 1 extending from a CTB2 origin state towards the EVT state, trajectory 2 extending from the CTB2 origin towards the SCTp state, and trajectory 3 extending from the CTB2 origin towards the proliferative CTB4 state. Each trajectory is summarized to the right of the UMAP plot. **(B)** Branched pseudotime ordering of human chorionic villi trophoblasts along the villous and extravillous paths of differentiation, constructed using Monocle2. Cells are first color coordinated by state (left panel) with abbreviations identical to those in Figure 1A, and second colour coordinated by estimated Monocle2 pseudotime values (right panel). **(C)** Branched pseudotime ordering of human chorionic villi trophoblasts, as shown in (B), now color coordinated by relative *BCAM* expression. **(D)** Scatter plot demonstrating raw *BCAM* expression levels in cells ordered along each Slingshot lineage trajectory. Progress along the x-axis indicates increased pseudotime. Loess curves demonstrate the average *BCAM* expression value across pseudotime for all three trajectories.

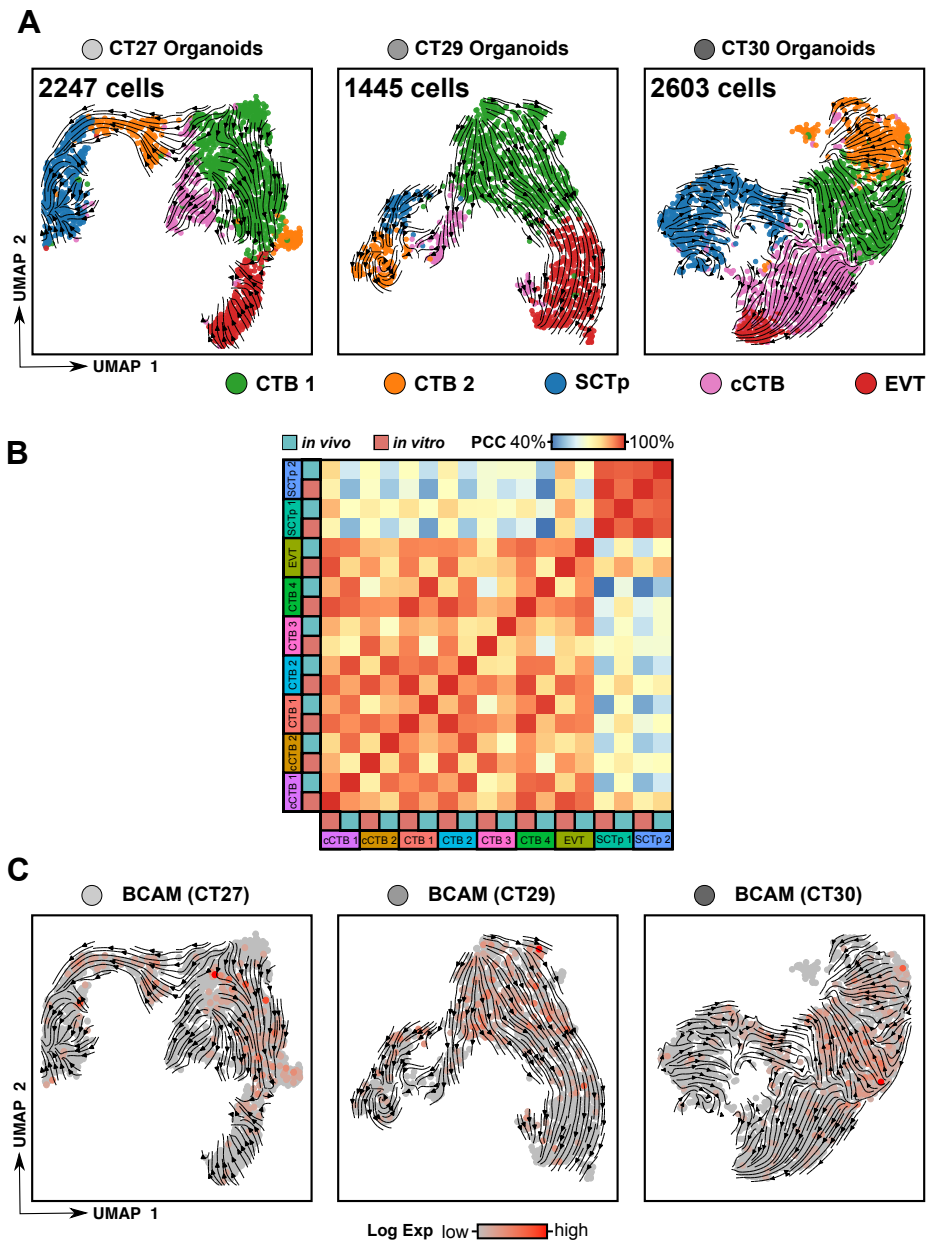


Fig. S4. Individual hTSC lines are comparable and highly correlated to placental villi data

(A) UMAP plots for individual hTSC lines (CT27; 2247 cells, CT29; 1445 cells, CT30; 2603 cells) subjected to scRNA-seq analyses. Individual hTSC lines were subset from the combined UMAP plot shown in Figure 3E and re-clustered, demonstrating 5 distinct trophoblast clusters representing 2 states of CTB (CTB1, CTB2), one cCTB state, one EVT-like state (EVT) and one SCT precursor state (SCTp). Individual cell line UMAP plots are overlain with their projected RNA velocity (scVelo) vector field, cells are color coordinated by state in each UMAP, and abbreviations are as in Figure 1A. **(B)** Correlation heatmap showing the Pearson's correlation coefficient between individual *in vivo* and *in vitro* states identified from integrating all scRNAseq-informed placental villi and hTSC-derived organoid data. Individual *in vivo* or *in vitro* states were treated as pseudo-bulk samples during correlation and abbreviations are as in Figure 1A. **(C)** UMAP plots for individual hTSC lines (CT27, CT29, CT30) subjected to scRNA-seq analyses. Cells in each UMAP are color coordinated by their respective log *BCAM* expression value. Individual cell line UMAP plots are overlain with their projected RNA velocity (scVelo) vector field and abbreviations are as in Figure 1A.

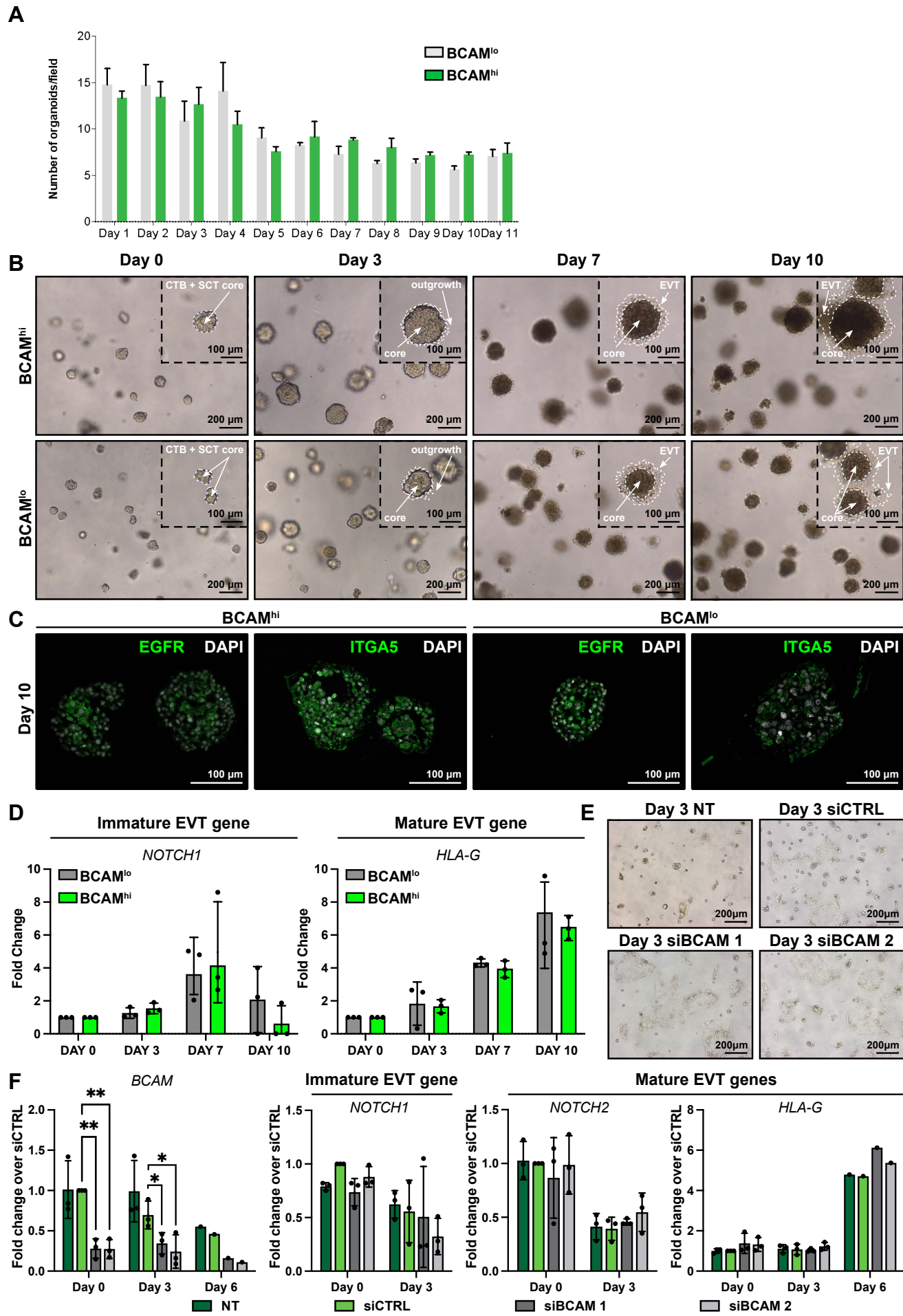
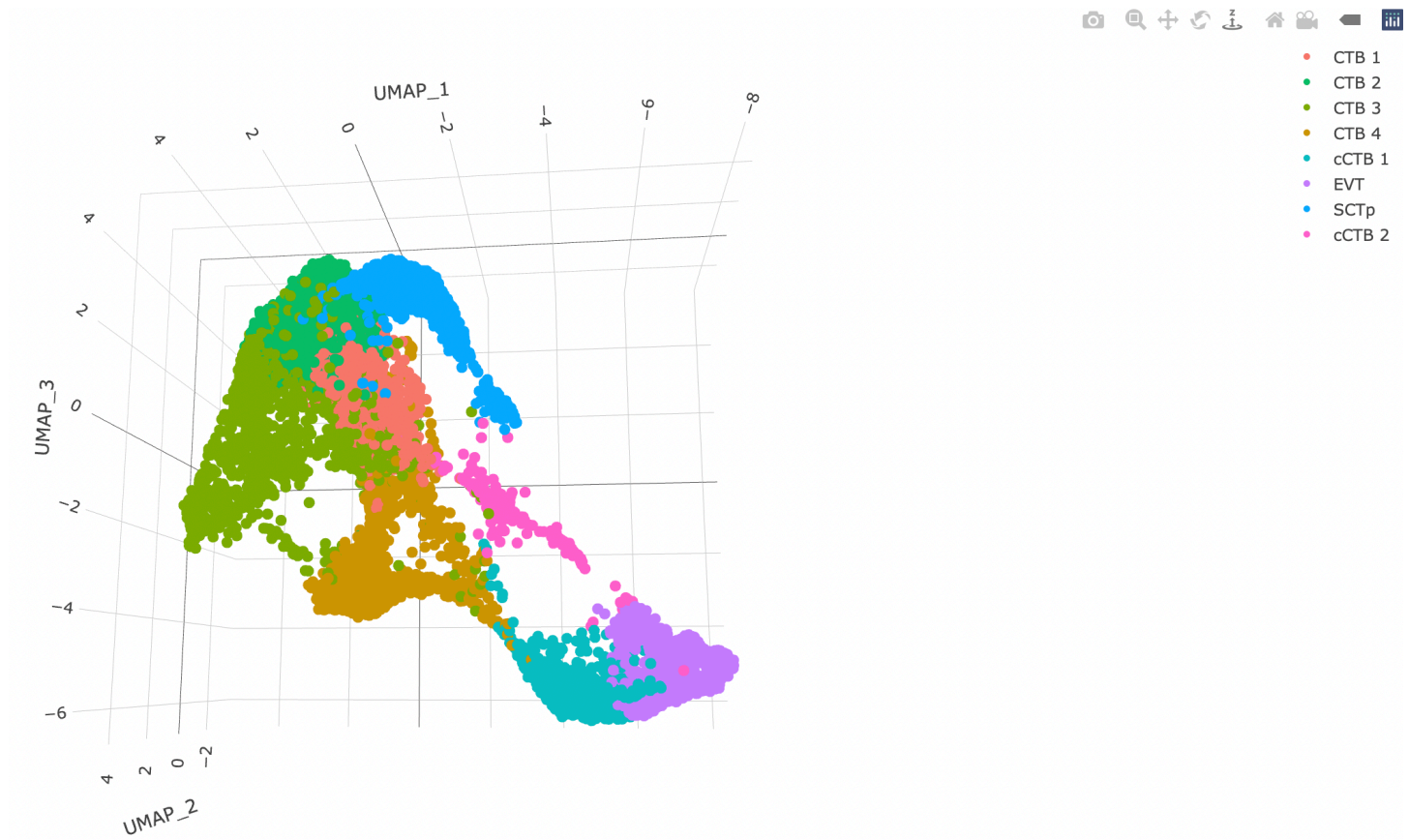
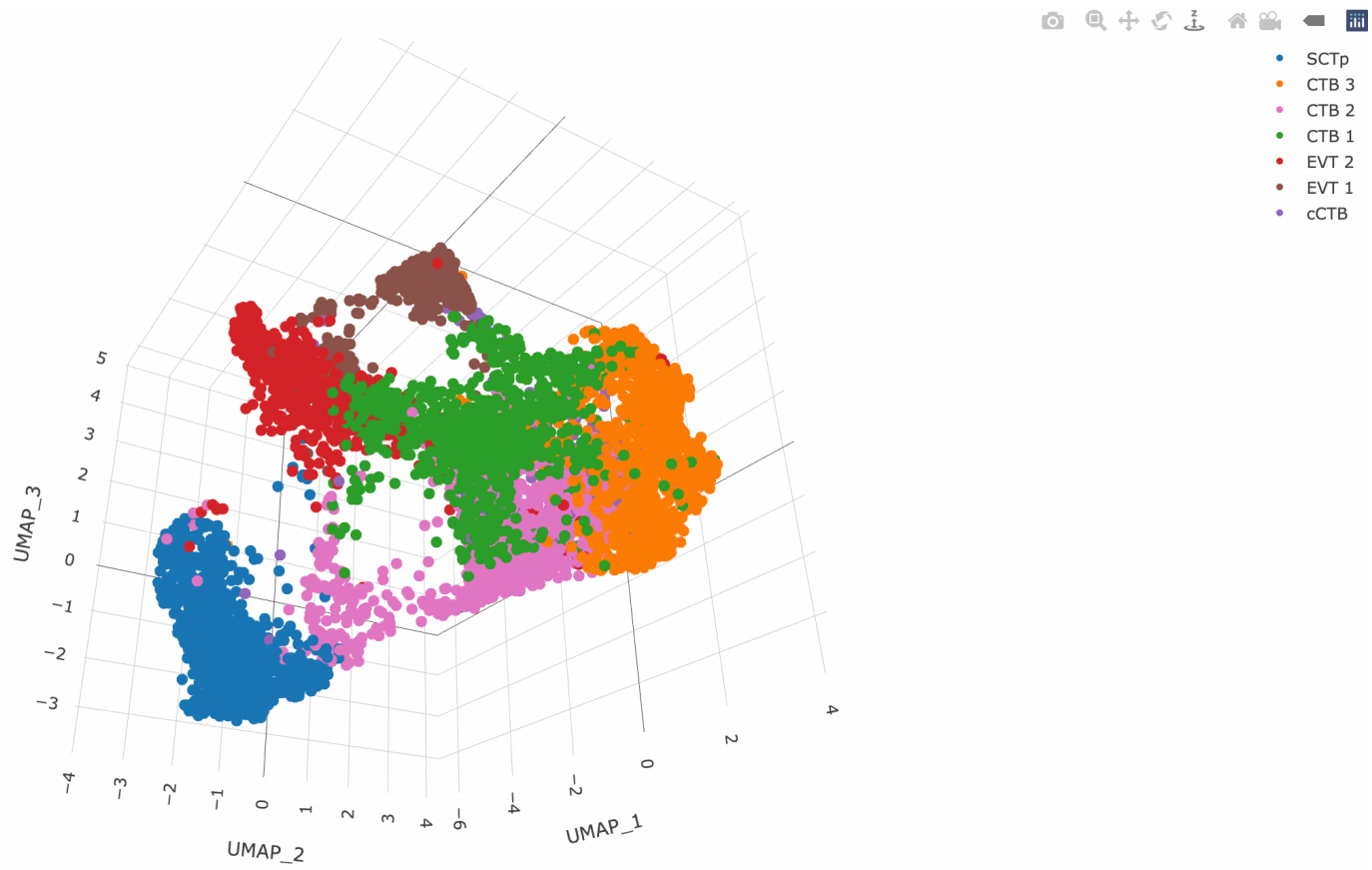


Fig. S5. BCAM does not affect organoid establishment or EVT differentiation (A) Bar graph shows the average of numbers of organoid colonies from BCAM^{hi} and BCAM^{lo} –sorted CT29 organoids cultured over 11 days. Grey bars = BCAM^{lo}; Green bars = BCAM^{hi}. (B) Bright field images of hTSC (CT29) BCAM^{hi}- and BCAM^{lo}-derived organoids cultured over 10 days in EVT Wnt- media. Hashed lines indicate EVT outgrowth. (C) IF images of EGFR and ITGA5 (α 5 integrin) in BCAM^{lo} and BCAM^{hi} sorted hTSC- organoids (CT29) cultured in EVT (Wnt-) media for 10 days. Nuclei are labelled with DAPI (white). (D) Relative measurement of *NOTCH1* and *HLA-G* mRNA levels by qPCR in EVT-differentiated hTSC-derived organoids (CT29) established from cells sorted on low (BCAM^{lo}) and high (BCAM^{hi}) of BCAM. Differentiation was performed on 3 independent times (n=3). (E) Representative bright field images of non-treated (NT), control siRNA-transfected (siCTRL), and BCAM siRNA transfected (siBCAM#1, siBCAM#1) 2D-cultured hTSCs (CT29) on day 2 of culture. (F) Relative measurement of *BCAM*, *NOTCH1*, *NOTCH2*, and *HLA-G* mRNA levels by qPCR in control (NT, CTRL) and BCAM-silenced (siBCAM#1, siBCAM#2) hTSC (CT29) EVT- differentiated cells over 6 days. BCAM-sorting and knock-down experiments were performed on one (n=1; day 6 of BCAM silencing) or three (n=3) independent occasions. * = $p < 0.05$; ** = $p < 0.01$.



Supplementary file 1. 3D-rendering of UMAP-informed *in vivo* trophoblast clusters

Chorionic villi scRNA-seq clusters showing state-specific colors corresponding to CTB1-4, SCTp, cCTB1, cCTB2, and EVT.



Supplementary file 2. 3D-rendering of UMAP-informed *in vitro* trophoblast organoid clusters Trophoblast organoid scRNA-seq clusters showing state-specific colors corresponding to CTB1-3, SCTp, cCTB, EVT1, and EVT2.

Table S1. Single cell sample input preprocessing metrics and metadata

[Click here to download Table S1](#)

Table S2. Genes differentially expressed in placental villous and organoid trophoblast clusters

[Click here to download Table S2](#)

Table S3. Gene signature similarities and differences in CTB and trophoblast organoid states

[Click here to download Table S3](#)

Table S4. List of differentially expressed genes between villous and organoid progenitor CTB

[Click here to download Table S4](#)