

## HUMAN DEVELOPMENT

## RESEARCH ARTICLE

# Cell trajectory modeling identifies a primitive trophoblast state defined by BCAM enrichment

Matthew J. Shannon<sup>1,2</sup>, Jennet Baltayeva<sup>1,2</sup>, Barbara Castellana<sup>1,2</sup>, Jasmin Wächter<sup>1,2</sup>, Gina L. McNeill<sup>1,2</sup>, Ji Soo Yoon<sup>1,3</sup>, Jenna Treissman<sup>1,2</sup>, Hoa T. Le<sup>1,2</sup>, Pascal M. Lavoie<sup>1,4</sup> and Alexander G. Beristain<sup>1,2,\*</sup>

## ABSTRACT

In early placental development, progenitor cytotrophoblasts (CTB) differentiate along one of two cellular trajectories: the villous or extravillous pathways. CTB committed to the villous pathway fuse with neighboring CTB to form the outer multinucleated syncytiotrophoblast (SCT), whereas CTB committed to the extravillous pathway differentiate into invasive extravillous trophoblasts (EVT). Unfortunately, little is known about the processes controlling human CTB progenitor maintenance and differentiation. To address this, we established a single cell RNA sequencing (scRNA-seq) dataset from first trimester placentas to identify cell states important in trophoblast progenitor establishment, renewal and differentiation. Multiple distinct trophoblast states were identified, representing progenitor CTB, column CTB, SCT precursors and EVT. Lineage trajectory analysis identified a progenitor origin that was reproduced in human trophoblast stem cell organoids. Heightened expression of basal cell adhesion molecule (*BCAM*) defined this primitive state, where *BCAM* enrichment or gene silencing resulted in enhanced or diminished organoid growth, respectively. Together, this work describes at high-resolution trophoblast heterogeneity within the first trimester, resolves gene networks within human CTB progenitors and identifies *BCAM* as a primitive progenitor marker and possible regulator.

**KEY WORDS:** Placenta, Trophoblast, Progenitor, Single cell RNA-sequencing, Differentiation, Organoids, Basal cell adhesion molecule, Human

## INTRODUCTION

The placenta is a blastocyst-derived tissue that shares the genetic identity of the developing fetus. Derived from cells of the trophoblast and extra-embryonic mesoderm, the human placenta matures into a functional organ by 10–12 weeks gestation (Lee et al., 2016; Chang et al., 2018). Through its direct interaction with maternal uterine epithelial, stromal and immune cells, the placenta coordinates the establishment of the maternal-fetal interface and serves as a barrier that is essential for nutrient-waste exchange between maternal and fetal circulations. Trophoblast-derived trophoblasts perform many functions of the placenta. These include, but are not limited to, the facilitation of nutrient and oxygen

transfer between mother and fetus (Burton et al., 2020), the modulation of the maternal immune response towards the semi-allogeneic fetus and placenta (Tilburgs et al., 2015; Turco and Moffett, 2019) and the production of hormones (i.e. chorionic gonadotropin, progesterone, placental lactogen) and other factors required for pregnancy (Fowden et al., 2015).

Of all eutherian mammals, the human placenta is the most invasive. Humans and rodents both share a haemochorial arrangement of trophoblast and uterine tissue, but control of uterine surface epithelial and stromal erosion, as well as artery remodeling by trophoblasts during blastocyst implantation and early placental establishment, is far more extensive in humans (Turco and Moffett, 2019). Importantly, key regulators of trophoblast lineage specification in rodents (i.e. *Cdx2*, *Eomes*, *Esrrb* and *Sox2*) do not appear to play essential (or identical in the case of *Cdx2*) roles in human trophoblasts (Knöfler et al., 2019). These important differences in human trophoblast biology underlie the need to use human placentas and cell systems for generating fundamental knowledge central to human trophoblast and placental development. Although recent advances in regenerative trophoblast culture systems have created the necessary tools required for in-depth cellular and molecular understandings central to trophoblast differentiation (Haider et al., 2018; Okae et al., 2018; Turco et al., 2018), to date, few studies have used these platforms to examine at high resolution trophoblast progenitor and/or stem cell dynamics.

In humans, two major trophoblast differentiation pathways – villous and extravillous – give rise to all trophoblasts in the fetal-maternal interface (Fig. 1A). In the villous pathway, progenitor cytotrophoblasts (CTB) fuse with neighboring CTB to replenish or generate a new syncytiotrophoblast (SCT), the multinucleated outer layer of the placenta that is the site of transport of most substances across the placenta. Cells committed to differentiate along the extravillous pathway are predicted to originate from CTB progenitors proximal to anchoring villi (Knöfler et al., 2019), though populations of villous CTB may already be programmed towards extravillous differentiation or may alternatively harbor a level of bi-potency (Chang et al., 2018; Lee et al., 2018). Nonetheless, progenitors devoted to this pathway give rise to extravillous trophoblasts (EVT) that anchor the placenta to the uterine wall (i.e. proximal and distal column CTB). EVT adopt invasive characteristics hallmarked by interstitial and endovascular EVT subsets that together facilitate uterine tissue and artery remodeling and modulate maternal immune cell activity (Moffett et al., 2017; Pollheimer et al., 2018; Papuchova et al., 2020). Defects in trophoblast differentiation along either pathway associate with impaired placental function that may contribute to, or drive the development of, aberrant conditions of pregnancy (Jauniaux et al., 2020).

In this study, we combined an in-house-generated single cell RNA-sequencing (scRNA-seq) dataset with a publicly available one (Vento-Tormo et al., 2018) to define trophoblast heterogeneity

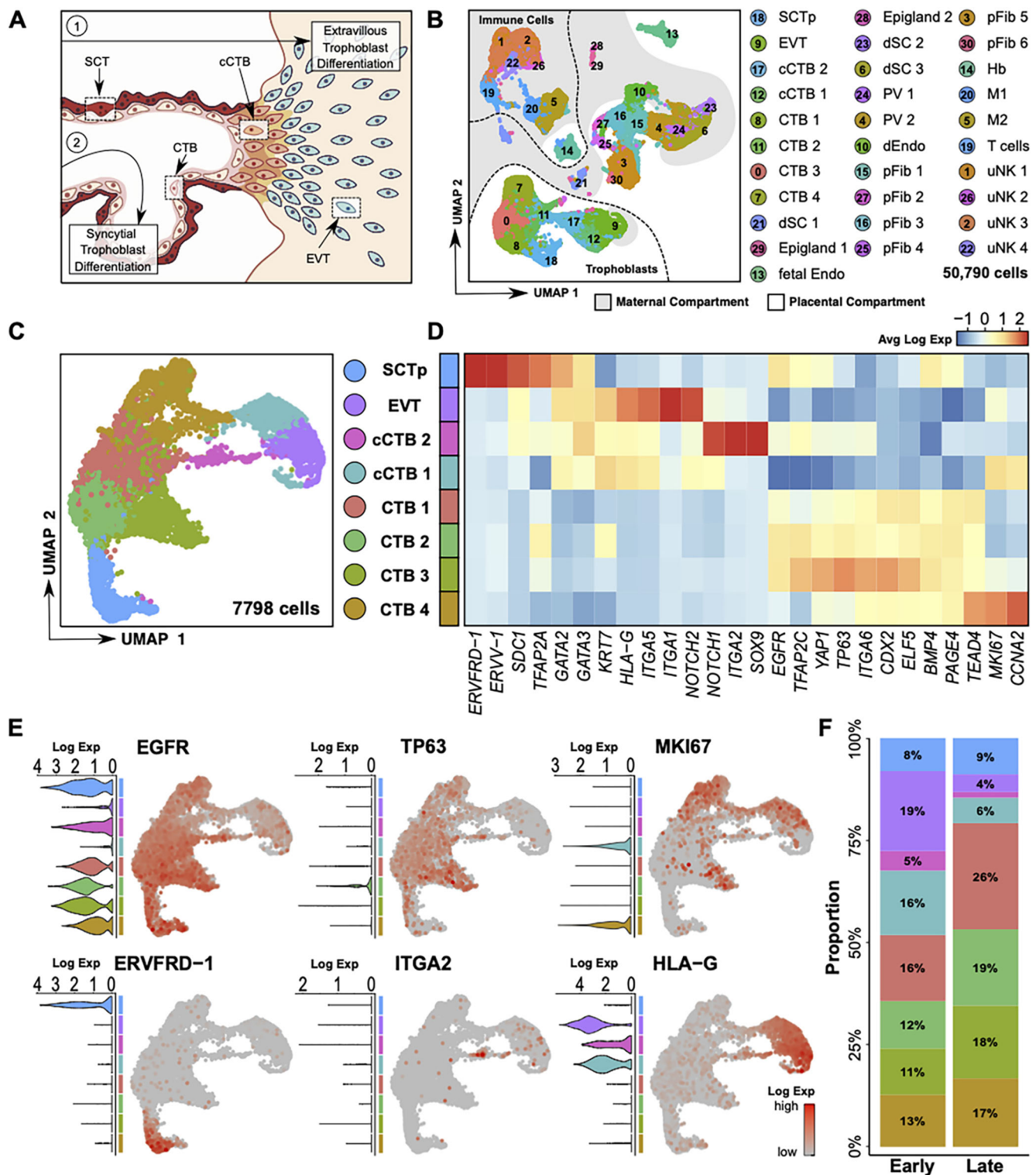
<sup>1</sup>The British Columbia Children's Hospital Research Institute, Vancouver V5Z 4H4, Canada. <sup>2</sup>Department of Obstetrics & Gynecology, The University of British Columbia, Vancouver V6Z 2K8, Canada. <sup>3</sup>Department of Surgery, The University of British Columbia, Vancouver V5Z 1M9, Canada. <sup>4</sup>Department of Pediatrics, The University of British Columbia, Vancouver V6H 3V4, Canada.

\*Author for correspondence (aberista@mail.ubc.ca)

DOI: P.M.L., 0000-0002-2205-0362; A.G.B., 0000-0002-1117-6242

Handling Editor: Samantha Morris

Received 26 May 2021; Accepted 7 December 2021



**Fig. 1. Single cell characterization of first trimester trophoblasts.** (A) Schematic of a first trimester chorionic villus including anatomical representations of both an anchoring villous and floating villous. Highlighted are trophoblast sub-populations and the villous and extravillous cell differentiation trajectories: cytotrophoblast (CTB), column cytotrophoblast (cCTB), syncytiotrophoblast (SCT) and invasive extravillous trophoblast (EVT). (B) UMAP plot of 50,790 captured cells, demonstrating 31 distinct clusters identified within the first trimester human maternal-fetal interface. Dashed lines separate maternal decidua immune cells from non-immune cells and placental trophoblasts, and background color indicates maternal or placental tissue origin of cells. Epigland, epiglandular cell; fetal Endo, fetal endothelial cell; dEndo, decidual endothelial cell; dSC, decidual stromal cell; PV, perivascular cell; pFib, placental fibroblast; Hb, fetal Hofbauer cell; M1, macrophage cell 1; M2, macrophage cell 2; uNK, uterine natural killer cell. (C) UMAP plot of 7798 captured trophoblasts, demonstrating eight distinct trophoblast clusters: CTB1-4, cCTB1/2, SCTp and EVT. (D) Heatmap showing the average expression of known trophoblast and proliferation gene markers in each trophoblast cluster. (E) UMAP plots and violin plots showing gene expression (red) of CTB-, cCTB-, SCT- and EVT-associated genes in all scRNA-seq-informed trophoblast states. Violin plots show cluster averaged sample distribution of normalized Log2 expression levels. (F) Proportion of first trimester trophoblasts in each scRNA-seq-defined state in early (6-7 weeks) and late (9-12 weeks) gestational time-points.

and differentiation kinetics in the first trimester of pregnancy. Cell trajectory modeling in chorionic villi and regenerative organoid trophoblasts identified a primitive cell origin with cell types and pathways conserved across these cell systems, but also identified important differences between *in vivo* and *in vitro* states. Within this upstream progenitor state, we show elevated expression of the gene basal cell adhesion molecule (*BCAM*), where *BCAM* enrichment or silencing enhances or impairs trophoblast organoid growth, respectively. Together, this work reaffirms prior knowledge underlying trophoblast heterogeneity, expands on our understanding of progenitor trophoblast complexity, and resolves new cell differentiation states specific to stage of trophoblast development. Further, this work identifies *BCAM* as a conserved cell-matrix adhesion protein in chorionic villous and trophoblast organoid progenitors important for trophoblast stem cell regeneration and growth.

## RESULTS

### Single cell transcriptomics resolves trophoblast heterogeneity and cell state dynamics

Single cell transcriptomic analysis enables the identification of novel cell types and states in a systematic and quantitative manner. Here, we applied the 10x Genomics Chromium platform to sequence the human placental transcriptome at single cell resolution ( $n=7$  placentas), and merged this data with single cell transcriptomic profiles of human first trimester uterine-placental tissues consisting of four placentas and four decidual specimens (Vento-Tormo et al., 2018) (Fig. S1A). To examine differences in cellular heterogeneity related to stage of development in the first trimester, we grouped individual samples into early (6–7 weeks) and late (9–12 weeks) gestational age (GA) groups. Clinical characteristics and sequencing metrics from each patient and sample are summarized in Table S1. After quality control and data pre-processing, 50,790 cells remained [12,794 cells derived from our in-house dataset; 37,996 cells derived from the Vento-Tormo et al. (2018) dataset] (Fig. S1B–D; Fig. S2A), and principal component analysis (PCA) validated data integration between the two datasets with separation of cells explained by tissue type, not data source (Fig. S1E). Cells clustered into 14 cell types and 31 distinct cell states by transcriptomic similarity, and cell cycle phase proportions per cluster informed the level of proliferation in each cell state (Fig. 1B; Fig. S1F–H). Decidual cells were identified by vimentin expression and decidual immune cells were specifically identified by *PTPRC* transcripts that encode the leukocyte CD45 antigen. Additional markers were used to identify endothelial, stromal and fibroblastic cell compartments (Fig. S1I). Integration of our and Vento-Tormo et al. (2018) datasets showed high concordance of cell states/clusters and overall cellular heterogeneity (Fig. S2A). UMAP clustering of all cells showed strong correlation between datasets (Fig. S2A,B). When trophoblasts were selected and re-clustered separately, inter-dataset correlation was even stronger (Fig. S2C,D).

For trophoblast identification, cells expressing combinations of trophoblast-related transcripts (i.e. *KRT7*, *EGFR*, *TEAD4*, *TP63*, *TFAP2A*, *TFAP2C*, *GATA2*, *GATA3*, *HLA-G*, *ERVFRD-1*) and relatively lacking mesenchymal- and immune-related transcripts (i.e. *VIM*, *WARS1*, *PTPRC*, *CD34*, *CD14*, *CD86*, *CD163*, *NKG7*, *KLRD1*, *HLA-DPA1*, *HLA-DPB1*, *HLA-DRA*, *HLA-DRB1*, *HLA-DRB5*, and *HLA-DQA1*) were subset for further analysis (Fig. S1G, J). Eight trophoblast clusters from 7798 cells were resolved; 3843 and 3955 trophoblasts were contributed from our in-house and Vento-Tormo et al. (2018) datasets, respectively (Fig. 1C;

Fig. S2C). Within these eight clusters, we identified four CTB states (CTB1–4; *TEAD4*<sup>+</sup>, *ELF5*<sup>+</sup>, and *EGFR*<sup>+</sup>), a fifth CTB state expressing SCT-associated genes that we termed syncytiotrophoblast precursor (SCTp; *ERVFRD-1*<sup>+</sup>, *CGA*<sup>+</sup>), two column trophoblast states (cCTB1/2; *NOTCH1*<sup>+</sup>, *ITGA5*<sup>+</sup>) and a state specific to EVT (EVT; *HLA-G*<sup>+</sup>, *ITGA1*<sup>+</sup>) (Fig. 1D,E; Table S2). As expected, CTB1–4 showed varying degrees of enrichment for transcription factors associated with trophoblast progenitor/stemness (i.e. *CDX2*, *TP63*, *ELF5*, *YAP1*), where the CTB3 state showed the greatest level of enrichment of these four genes (Fig. 1D).

Column trophoblasts are precursors to EVT, where cells situated within proximal and distal sites of anchoring columns can be molecularly distinguished by expression of NOTCH and integrin genes (Zhou et al., 1997; Haider et al., 2016). To this end, cCTB2 expresses high levels of *NOTCH1*, *SOX9* and *ITGA2*, the latter of which defines a distinct EVT progenitor population and likely represents a proximal column progenitor (Lee et al., 2018) (Fig. 1D, E). This possibility is strengthened in that cCTB2 expresses, at comparable levels to CTB1, the progenitor CTB genes *TP63*, *ITGA6* and *YAP1* (Fig. 1D,E). Compared with cCTB2, cells in the cCTB1 state showed higher levels of *ITGA5*, *NOTCH2* and *HLA-G*, indicating that this state is developmentally downstream of cCTB2 and cells in this state likely reside within central and distal regions of anchoring columns (Haider et al., 2014) (Fig. 1D,E). Stratifying cells by phases of the cell cycle showed that cCTB1 column cells are predominately in G<sub>2</sub>/M and S phases, whereas ~30% of cCTB2 column cells are in G<sub>1</sub> (Fig. S1H). Terminally differentiated EVT and mature SCTp states are largely G<sub>1</sub> cells, as are progenitor CTB1–3 states (Fig. S1H). This was in contrast to CTB4, in which the vast majority of cells are undergoing DNA synthesis and mitosis (Fig. S1H), express hallmark genes of proliferation (i.e. *MKI67*, *CCNA2*), and show enrichment for the transcription factor *TEAD4* (Fig. 1D,E). Importantly, these distinct trophoblast states largely align with recent single cell transcriptomic characterizations of placental cells from first trimester pregnancies (Liu et al., 2018; Suryawanshi et al., 2018; Vento-Tormo et al., 2018).

Having defined eight trophoblast types/states within first trimester placental tissue, we next examined cell state stability across the first trimester. Relative frequencies of CTB1–4 states increase, whereas proportions of cCTB and EVT states decrease from early (6–7 weeks) to late (9–12 weeks) gestation (Fig. 1F). Notably, an almost 5-fold reduction in the proportion of EVT (19% to 4%) is observed (Fig. 1F), suggesting that villous development and branching contribute to the increase in CTB frequency. Unexpectedly, and unlike CTB1–4, frequencies in SCTp remained relatively constant (8% to 9%) (Fig. 1F). Together, these findings define unique states of trophoblasts in the first trimester placenta and highlight how proportions of these states change between the early and late first trimester, a developmental window during which the human placenta becomes fully functional.

### Single cell trajectory modeling identifies a CTB origin

RNA velocity, the time derivative of the gene expression state, can be used to predict the future transcriptional state of a cell. RNA velocity analysis in combination with single cell trajectory analyses provides us with insights into the transcriptional dynamics of developing cells and enables high-resolution lineage pathway ordering. By applying scVelo to refine progenitor and cell endpoint relationships in our dataset, we identified an origin that overlapped with CTB2; this predicted origin was also proximal to the CTB3 state, indicating that these two states contain a



proportionally high content of primitive trophoblasts (Fig. 2A). Two paths were identified extending from this origin towards the EVT state and one path developing from the origin towards the SCTp state. As well, we found an additional differentiation trajectory, extending from CTB2 into CTB3, representing possible CTB progenitor reserves or progenitor states in transition (Fig. 2A). When shown in a 3D perspective, a linkage between CTB3 and CTB4 becomes evident that is not obvious in 2D clustering, indicating that cells in CTB2 can transition to proliferative CTB4 cells via multiple cell pathways (Supplementary file 1). Monocle3 (Cao et al., 2019) and Slingshot (Street et al., 2018) pseudotime ordering were also applied to increase confidence in trophoblast hierarchical ordering. Similar to RNA velocity, Monocle 3 identified a region overlapping with CTB2 and CTB1 as the origin; increased complexity is observed as CTB progress along one of two trajectories that lead to either EVT or SCT endpoint states (Fig. 2B). Specifically, Monocle3 ordering revealed a looped trajectory within CTB2 that exits towards CTB3, SCTp and EVT states and, consistent with scVelo ordering, also highlights two trajectories extending from CTB1 and CTB4 states into cCTB states (Fig. 2B). Slingshot pseudotime ordering also identified CTB2 as the state overlapping with the origin (Fig. S3A). To summarize, three different cell trajectory computational tools identified trophoblast differentiation pathways having overall strong concordance with each other, in which the CTB2 state was consistently identified as the progenitor origin.

We next compared these refined trajectories with expression patterns of known CTB progenitor/stem cell genes *CDX2*, *TEAD4*, *YAP1*, *TP63*, *ELF5* and *EPCAM*. Predominant expression of *ELF5*, *TEAD4* and *YAP1* was observed in all CTB states, whereas expression of the CTB stem cell transcription factor *CDX2* showed prevalent levels in cells in CTB1 and CTB3, and to a lesser extent CTB2 and CTB4 (Fig. 2C). Average *TP63* expression was highest in CTB3, though a greater proportion of cells within CTB2 expressed *TP63* compared with other CTB states, and *EPCAM* showed greatest enrichment within CTB1 and CTB4 (Fig. 2C). Because the scVelo-predicted origin aligned closely to CTB2 and CTB3, we examined genes both shared and enriched in each of these states (Fig. 2D). Twenty-four genes were shown to overlap between CTB2 and CTB3, including *EGFR*, *IGF2* and *FBLN1* (Fig. 2D; Table S3). Also, 249 genes showed enrichment in CTB3 and, of these, many associate with trophoblast stem cell and regenerative properties (i.e. *YAP1*, *TP63*), Wnt signaling (*LRP5*) and cell-matrix interactions (*MMP14*, *MMP15*, *ITGA6*, *LAMA5*, *LAMB1*) (Fig. 2D). The CTB2 state showed enrichment for 251 genes, including the Wnt ligand *WNT7A*, the laminin receptor *BCAM*, and the transcriptional modulators *TFAP2A* and *CITED2* (Fig. 2D). Expression levels of the top 24 CTB2-enriched genes (FDR ranked) plotted by cell state over a simple branched pseudotime [as performed previously using Monocle 2 in Vento-Tormo et al. (2018) and Treissman et al. (2020)] showed consistent enrichment of all genes to the CTB origin (Fig. 2E). Notably, some genes (e.g. *BCAM*, *IFI6*) show a consistent downregulation along the SCT and EVT trajectories, whereas others (e.g. *DUSP9*, *GSTA3*, *VAMP8*) show transient decreases followed by increases in gene levels approaching SCT or EVT trajectory endpoints (Fig. 2E). scVelo-alignment of *BCAM*, one of the most enriched genes within the CTB2 state (Table S2), shows that *BCAM* levels are highest within the predicted CTB origin (overlapping primarily with CTB2) and gradually decrease as CTB differentiate into SCTp or cCTB/EVT (Fig. 2F).

Correlating *BCAM* with lineage trajectories using Monocle2 and Slingshot also showed *BCAM* levels to be enriched in cells aligning

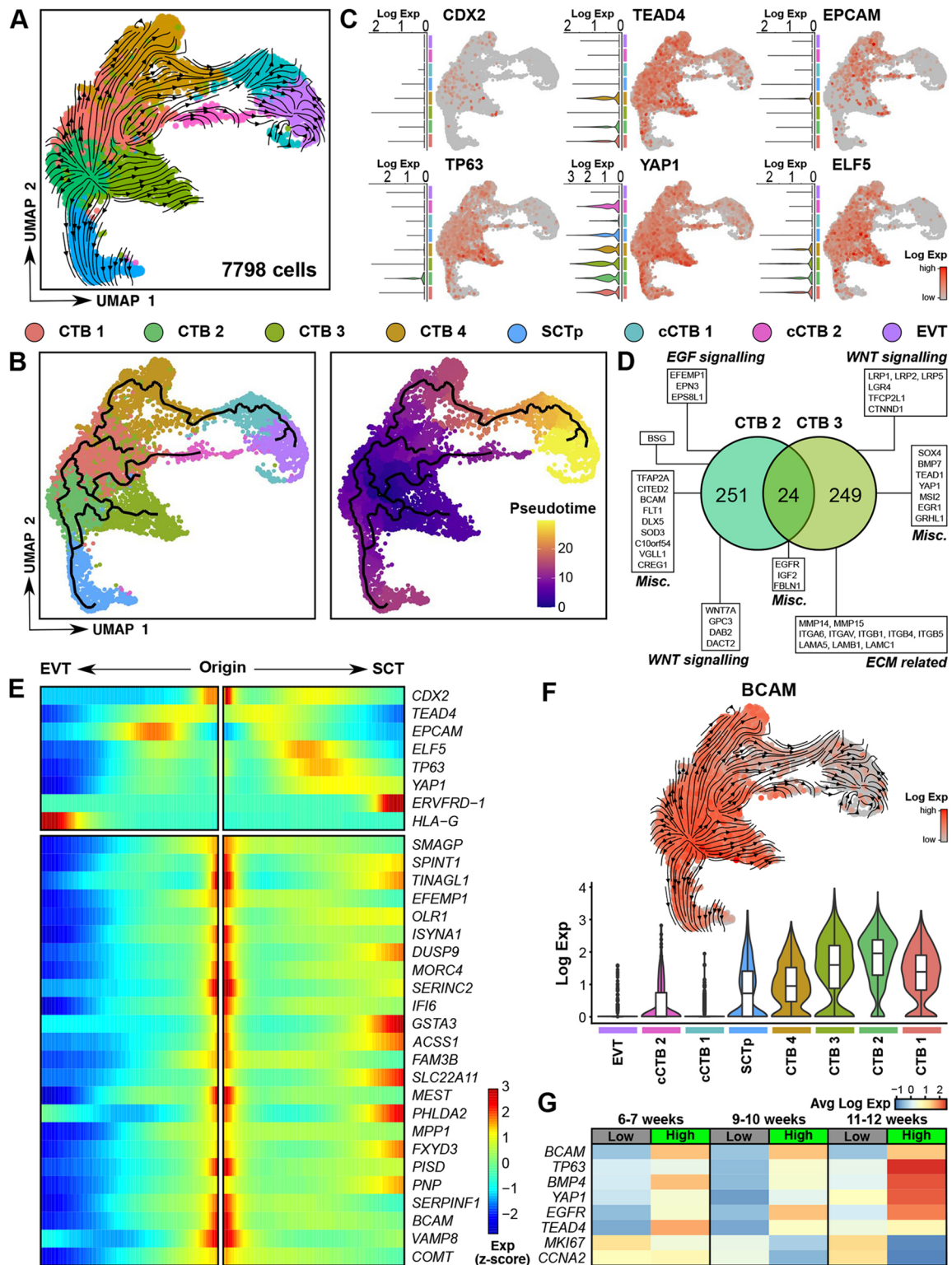
to the predicted origins (Fig. S3B-D). Trophoblasts expressing high levels of *BCAM* show gradual increases in known trophoblast progenitor genes from 6-7, 9-10 and 11-12 weeks GA, suggesting that the correlation of *BCAM* with progenitor-like properties strengthens as pregnancy progresses (Fig. 2G). In contrast, trophoblasts expressing low/no levels of *BCAM* show higher levels of proliferation-associated *MKI67* and *CCNA2* but markedly lower levels of progenitor-associated genes (Fig. 2G). Taken together, we identify a primitive CTB state within the first trimester placenta that shows molecular features consistent with that of putative CTB stem cells. We also identify multiple genes aligning with this origin, many of which have not been previously described in trophoblast progenitor biology, including *BCAM*.

### Trophoblast stem cells recapitulate scRNA-seq-informed CTB heterogeneity and hierarchy

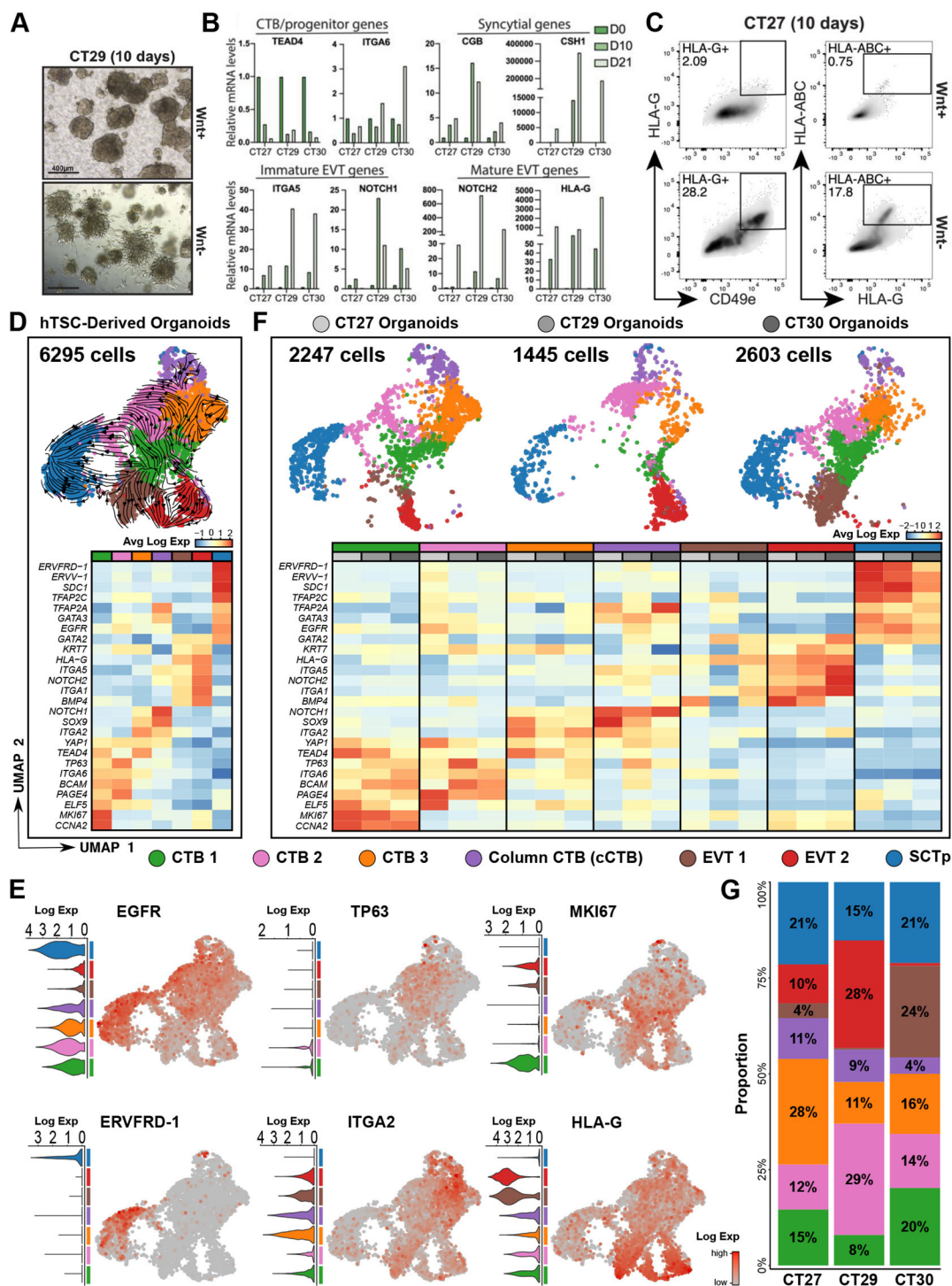
Recently described regenerative human trophoblast models provide an improved method for studying trophoblast differentiation *in vitro* (Okada et al., 2018). Three-dimensional organoid cultures established from human trophoblast stem cell lines (hTSCs) also reproduce molecular features underlying CTB maintenance and differentiation (Saha et al., 2020). To examine the reproducibility of trophoblast-intrinsic differentiation programs defined by scRNA-seq analysis of chorionic villous tissues, we subjected three lines of hTSC organoids to scRNA-seq analysis before and after EVT differentiation. hTSC organoids (CT27, CT29 and CT30 hTSC lines) were cultured for 7 days in regenerative Wnt<sup>+</sup> (containing the Wnt potentiators R-spondin and CHIR99021) and EVT-differentiating Wnt<sup>-</sup> (lacking R-spondin and CHIR99021) media before single cell workflow. To confirm the intactness of our system, a separate cohort of hTSC organoids were cultured over 21 days in these media and differentiation was assessed by light microscopy, flow cytometry and qPCR analysis. hTSC-derived organoids cultured in regenerative media established large spherical colonies characterized in part by elevated expression of the CTB progenitor genes *TEAD4* and *ITGA6* (Fig. 3A,B). Moreover, expression of surface MHC-I (HLA-A, -B, -C) by flow cytometry did not detect robust frequencies of MHC-I ligands in undifferentiated cultures (Fig. 3C). Comparatively, hTSC organoids cultured in EVT differentiating Wnt<sup>-</sup> media showed characteristic cellular extensions from the organoid body (Fig. 3A), as well as a marked decrease in levels of *TEAD4* mRNA (Fig. 3B). As expected, both immature (*ITGA5*, *NOTCH1*) and mature (*NOTCH2*, *HLA-G*) EVT marker genes, as well as HLA-C and HLA-G surface expression increased along the EVT differentiation timeline (Fig. 3B,C). Syncytial formation occurs in both regenerative Wnt<sup>+</sup> and EVT Wnt<sup>-</sup> conditions, and this was reflected at the gene level, where *CGB3* and *CSH1* levels markedly increased over 21 days (Fig. 3B). Together, these data indicate that the 3D organoid model to study CTB regeneration and differentiation is intact.

Regenerative and EVT-differentiated hTSC organoid cells were then used to generate single cell cDNA libraries following 10x Genomics Chromium workflow. Combining single cell transcriptomic data from CT27, CT29 and CT30 organoids resolved seven distinct cell states; three CTB (CTB1-3), a precursor to SCTp, a putative column/EVT progenitor (cCTB), and two EVT-like states (EVT1, EVT2) (Fig. 3D). Hallmark gene markers of trophoblast-lineage (*TFAP2A*, *TFAP2C*), progenitor CTB (*TEAD4*, *YAP1*, *TP63*, *ITGA6*, *ELF5*), general CTB (*EGFR*), SCT (*ERVFRD-1*, *ERVV-1*, *SDC1*), immature EVT (*ITGA5*, *ITGA2*, *NOTCH1*) and mature EVT (*ITGA1*, *NOTCH2*, *HLA-G*),





**Fig. 2. scVelo-informed cell differentiation trajectories identify a primitive CTB origin.** (A) UMAP plot overlay with projected RNA velocity (scVelo) vector field, indicating trophoblast directionality (represented by the velocity arrows) away from a putative origin towards endpoint SCTp and EVT states. Cell cluster names and colors correspond to those in Fig. 1C. (B) UMAP plot with constructed trajectory graph overlay. Cells are color coordinated by cluster identity as defined in Fig. 1C (left) and by relative pseudotime value (right). (C) UMAP plots and violin plots showing gene expression (red) of progenitor- and stem cell-associated genes in all scRNA-seq-informed trophoblast states. (D) Venn diagram comparing elevated genes within CTB states 2 and 3. (E) A heatmap was constructed using inferred pseudotime and the top 24 upregulated genes in CTB2. Included are known genes that align with the CTB (*CDX2*, *TEAD4*, *EPCAM*, *TP63*, *YAP1*, *ELF5*), the SCT (*ERVFRD-1*) and EVT (*HLA-G*) states. (F) scVelo-overlaid UMAP and violin plots showing *BCAM* levels in all first trimester trophoblast states. Boxplots are overlaid demonstrating the interquartile ranges and median log gene expression level. Violin plots show cluster averaged sample distribution of normalized Log2 expression levels. Box plot inside violin plot shows median and interquartile range. (G) Heatmap showing average progenitor gene expression in BCAM high (log expression > 2.0) and BCAM low (log expression = 0.0) trophoblasts in early (6-7 weeks GA), mid (9-10 weeks GA) and late (11-12 weeks GA) first trimester samples.



**Fig. 3. scRNA-seq analyses in hTSC organoids recapitulates a primitive CTB origin.** (A) Brightfield microscopy images of regenerative (Wnt<sup>+</sup>) and EVT-differentiated (Wnt<sup>-</sup>) hTSC-derived organoids following 10 days of culture. Scale bars: 400  $\mu$ m. (B) Relative measurement of mRNA levels by qPCR in EVT-differentiated hTSC-derived organoids (CT27, CT29, CT30) over 21 days (D0-D21) of culture. Shown are CTB/progenitor (*TEAD4*, *ITGA6*), syncytial (*CGB3*, *CSH1*), immature EVT (*ITGA5*, *NOTCH1*) and mature EVT (*NOTCH2*, *HLA-G*) genes. Each hTSC line was differentiated as one ( $n=1$ ) replicate. (C) Representative flow cytometry plots showing cell frequencies of HLA-G<sup>+</sup>, CD49e<sup>+</sup> and HLA-ABC<sup>+</sup> cells in regenerative (Wnt<sup>+</sup>) and EVT-differentiated (Wnt<sup>-</sup>) hTSC-derived (CT27) organoids. CT27 differentiation was performed three ( $n=3$ ) independent times. (D) UMAP and gene heatmap plot of combined CT27, CT29 and CT30 single cell hTSCs following culture in regenerative (Wnt<sup>+</sup>) and EVT-differentiated (Wnt<sup>-</sup>) conditions after 7 days in culture. UMAP plots are overlain with projected RNA velocity (scVelo) vector field, demonstrating seven distinct trophoblast clusters representing three states of CTB (CTB1, CTB2, CTB3), a cCTB state, two EVT-like states (EVT1, EVT2) and an SCT precursor state (SCTp). Heatmap shows the average expression of trophoblast- and proliferation-associated genes in each cell state. (E) UMAP plots and violin plots showing gene expression (red) of CTB, cCTB, SCT and EVT-associated genes in all scRNA-seq-informed hTSC-derived organoid states. Violin plots show cluster averaged sample distribution of normalized Log2 expression levels. (F) UMAP and gene heatmap plot derived from individual hTSC lines (CT27, 2247 cells; CT29, 1445 cells; CT30, 2603 cells) subjected to scRNA-seq analyses. (G) Proportion of clusters derived from each cell line (CT27, CT29, CT30) in the combined hTSC-derived organoid scRNA-seq-defined state.

as well as gene markers of proliferation (*MKI67*, *CCNA2*) are shown as a heatmap and in UMAP-generated cluster projections showing individual cell states (Fig. 3D,E Table S2). CTB1-3 shared moderate overlap in gene expression, though CTB1 distinctly co-expressed proliferative and progenitor genes (Fig. 3D). CTB2 and CTB3 also expressed progenitor genes; however, levels were comparatively lower in CTB3 than in CTB1 or CTB2 states (Fig. 3D). The putative column/EVT progenitor state, cCTB, preferentially expressed *ITGA5*, *ITGA2*, *SOX9* and the immature EVT marker *NOTCH1*; at the gene level, this unique state resembles the EVT progenitor population described by Lee et al. (2018) (Fig. 3D). States resembling terminally differentiated cell types showed strong enrichment for SCT- and EVT-associated genes, with EVT2 showing the highest levels of *HLA-G* and *NOTCH2* (Fig. 3D).

When the combined hTSC organoid object is split into individual hTSC cell lines (CT27, 2247 cells; CT29, 1445 cells; CT30, 2603 cells), comparison across cell lines showed notable variances in cell heterogeneity (Fig. 3F,G). Although defined SCTp and EVT endpoint states are evident in all three cell lines, variances defined by gene expression in CTB and cCTB states (Fig. 3F), as well as cluster proportions in EVT1 and EVT2 states (Fig. 3G), suggest that the hTSC lines have distinct cell differentiation kinetics or pathways unique to themselves. Subsetting and re-clustering cells from each individual hTSC cell line reduced the total number of cell states from seven to five (Fig. S4A). This is normal and likely because of transcriptomic variance between cell lines (Fig. 3F) and reduced cell numbers in each condition. Nonetheless, the overall makeup of each hTSC cell line is consistent and by combining single cell data from each hTSC organoid or by re-clustering and plotting data as individual stem cell lines, scVelo revealed a common CTB origin aligning to CTB1 (Fig. 3D; Fig. S4A). Similar to scVelo modeling in chorionic villi, 3D rendering of hTSC organoid UMAP clusters highlights multiple pathways extending from CTB1 to EVT, whereas a more direct single path is shown for SCTp establishment extending from CTB3 (Supplementary file 2). Examining gene signatures within these origins showed consistent enrichment of the trophoblast progenitor genes *TP63* and *YAP1*. However, we did see variances in *ELF5* levels between hTSC lines. Specifically, *ELF5* was enriched in CTB2 in the CT27 hTSC line but aligned more with the origin and the CTB1 state in all three cell lines (Fig. 3F). Consistent expression levels for the proliferative gene markers *MKI67* and *CCNA2* were observed within CTB1 in all hTSC lines, whereas CTB1 and CTB2 states showed more enrichment of progenitor/stem genes *TEAD4* and *YAP1* in CT27 and CT29 lines (Fig. 3F). When averaged across all hTSC lines, however, we observed consistent expression of *ELF5*, *BCAM*, *MKI67* and *CNNA2* within the CTB1 state that overlapped with the predicted origin (Fig. 3D).

#### Integration of chorionic villi and hTSC organoid single cell data identifies two progenitor origins

Having generated scRNA-seq datasets from both chorionic villi and hTSC-derived organoids, we next set out to examine how accurately *in vitro* hTSC differentiation aligns with *in vivo* differentiation trajectories and overall cell complexity. Comparison of transcriptional signatures between the scVelo-determined origins in CTB from chorionic villi and hTSC organoids highlights that there is some overlap between CTB2 (*in vivo*) and CTB1 (*in vitro*), as well as notable differences (Fig. 4A; Table S3). We found that 5084 genes are shared, whereas 5664 and 4946 genes are enriched within the origin states of chorionic villi and organoids, respectively (Fig. 4A). A differential gene expression analysis between CTB2

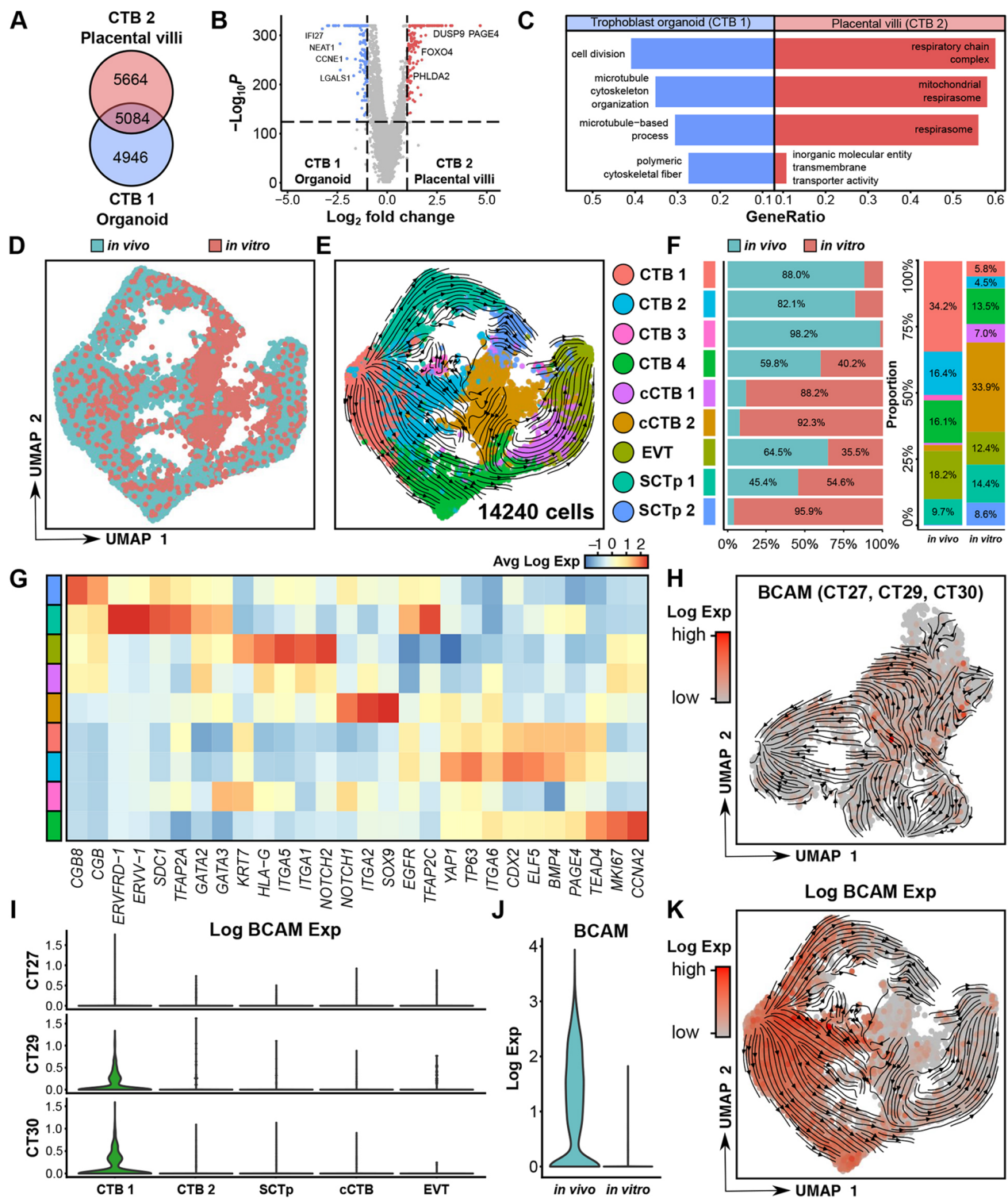
(villi) and CTB1 (organoid) states revealed 13,364 differentially expressed genes (DEGs) (Fig. 4B; Table S4). Genes enriched within the placental villi origin include *ACSS1* (encoding an acyl-CoA synthetase), genes related to ATP synthase regulation [e.g. *ATP5L* (also known as *ATP5MG*), *ATP5B* (*ATP5F1B*)], the transcription factor *FOXO4*, and the imprinted gene *PHLDA2* shown to function as a tumor suppressor and to regulate placental growth (Tunster et al., 2016; Wang et al., 2018). Top genes enriched in organoid CTB1 progenitors included many genes associated with cell cycle progression (e.g. *CDK1*, *CCNE1*, *MKI67*), the metalloprotease *ADAM15*, the immunomodulator *LGALS1*, and the xenobiotic efflux transporter *ABCG2*, shown to be selectively enriched within progenitors and stem cells in other tissue systems (Bhattacharya et al., 2007; Fatima et al., 2012). Pathway analysis of villous CTB2 and organoid CTB1 signatures shows that organoid progenitors have enrichment of pathways underlying cell division and microtubule assembly, indicative of a pro-proliferative state, whereas villous CTB2 progenitors showed enrichment of genes related to mitochondrial respiration, mitochondrial function, and inorganic transporter activity (Fig. 4C). These findings suggest that, although there are conserved processes shared between villous and hTSC organoid progenitors, there are also differences that should be acknowledged as possible limitations of current hTSC models.

To better explore how EVT and SCT differentiation kinetics are conserved in the hTSC organoids, we integrated chorionic villi/decidua (11 placentas, four decidua) datasets with organoid single cell data (Fig. 4D). Cell clustering identifies nine clusters representing four CTB (CTB 1-4), two SCTp (SCTp1/2), two cCTB (cCTB1/2) and one EVT state (Fig. 4E). Notably, cell state proportion is driven by cell source; chorionic villi contribute to ~90% of CTB1-3, whereas hTSC organoids contribute to less than 10% of these CTB states (Fig. 4F). Conversely, hTSC organoids contributed to ~90% of column trophoblast content (cCTB1, cCTB2) (Fig. 4F). Though differences in trophoblast state proportions were seen in chorionic villus- and organoid-derived trophoblasts, hTSC organoids did reproduce all *in vivo* states. Further, when average gene expression from each state is used to correlate similarities between chorionic villi and hTSC organoid states, an overall strong correlation is seen (Fig. S4B).

scVelo pseudotime analysis of the integrated dataset identified two potential origins (Fig. 4E). One origin overlapped with CTB1 and aligned well with the chorionic villi and organoid origins described previously (Fig. 4E). The new second origin aligned predominantly with cCTB2, where velocity arrows fed into CTB4, cCTB1 and EVT, suggesting these cells may represent EVT progenitors (Fig. 4E). Whereas chorionic villi contributed robustly to CTB1 cells residing within origin 1, hTSC organoids contributed more to the cCTB2 state that overlapped with origin 2 (Fig. 4E,F). Importantly, dataset integration overall conserves the CTB states previously characterized; CTB1 and CTB2 gene signatures align closely with the previous chorionic villi CTB2 and CTB3 states, and accordingly express the highest levels of trophoblast progenitor genes (e.g. *YAP1*, *TP63*, *ELF5*) (Fig. 4G).

Consistent with data from primitive CTB in chorionic villi, *BCAM* was shown to be one of the most highly expressed and FDR-ranked genes shared between organoid CTB1 and villous CTB2 states (Table S3). *BCAM* also highly overlapped with the scVelo-predicted origin in trophoblast organoids, as it did with the *in vivo* origin in villous trophoblasts (Fig. 4H,I; Fig. S4C), though overall *BCAM* levels in organoid CTB are lower than levels in CTB observed *in vivo* (Fig. 4J). Consistent with these findings, chorionic villi/organoid single cell integration clearly showed strong





**Fig. 4. Integration of *in vivo* and *in vitro* scRNA-seq datasets reveals two origin states.** (A) Venn diagram comparing genes shared and distinct between CTB2 in chorionic villi and CTB1 in hTSC organoids. (B) Volcano plot showing DEGs between CTB2 in chorionic villi and CTB1 in organoids. Select genes are highlighted. (C) Top pathways enriched in CTB2 (placental villi) and CTB1 (hTSC organoids) gene signatures identified in Fig. 4B. (D) UMAP plot of 14,240 integrated trophoblasts from our *in vivo* and *in vitro* datasets color coordinated by dataset source. (E) Integrated UMAP plot color coordinated by cell cluster identity demonstrating nine distinct trophoblast clusters: CTB1-4, SCTp1/2, cCTB1/2 and EVT. The UMAP plot is overlain with the projected RNA velocity (scVelo) vector field, indicating trophoblast directionality away from two putative origins. (F) Bar plots showing the proportion of cells in each defined state from *in vivo* and *in vitro* datasets (left) and the proportional relationship of each defined state within our *in vivo* and *in vitro* conditions (right). (G) Heatmap showing the average expression of known trophoblast and proliferation gene markers. (H) scVelo-overlaid UMAP plot showing *BCAM* levels in all (CT27, CT29, CT30) scRNA-seq-informed hTSC-derived organoid states. (I) Violin plots showing log *BCAM* expression for all scRNA-seq-informed organoid states identified in either the CT27, CT29 or CT30 datasets. (J) Violin plot comparing the pseudobulk log *BCAM* expression between our *in vivo* and *in vitro* datasets. Violin plots show cluster averaged sample distribution of normalized Log2 expression levels. (K) scVelo-overlaid UMAP plot showing *BCAM* levels in all integrated *in vivo* plus *in vitro* trophoblast states.

correlation of *BCAM* with the CTB1 progenitor origin; *BCAM* levels are markedly lower in the cCTB2 predicted origin and further decrease as cells differentiated into SCTp or EVT (Fig. 4K). These data show that molecular programs are largely conserved in hTSC organoids, that similarities in overall trophoblast heterogeneity are maintained, and that the cell adhesion molecule *BCAM* is a shared marker of regenerative CTB in chorionic villi and hTSC organoids.

### **BCAM confers increased CTB regenerative potential**

To our knowledge, *BCAM* localization within the human placenta has never been reported. To this end, we probed early (6.3 week) and late (11.2 week) first trimester placental villi specimens with *BCAM* antibody (Fig. 5A). A specific and strong *BCAM* signal was observed in CTB of both developmental time-points, and serial sections stained with keratin 7 (KRT7) and HLA-G showed that *BCAM* was present within proximal column cells, though the intensity of the signal was less than in villous CTB (Fig. 5A). Serial sections of chorionic villi stained with KRT7 and human chorionic gonadotrophin (hCG) clearly showed that *BCAM* is absent from the SCT outer layer (Fig. 5A). Consistent with this, in hTSC-derived organoids cultured in regenerative conditions, *BCAM* signal was specific to the outer CTB layer and absent within the inner hCG<sup>+</sup> syncytium (Fig. 5B). Dual labeling with *BCAM* and Ki67 antibody showed that *BCAM*<sup>+</sup> CTB also co-express the proliferation marker (Fig. 5B).

To test the regenerative potential of *BCAM*-expressing cells, *BCAM*<sup>hi</sup> and *BCAM*<sup>lo</sup> CTB were sorted from CT29 hTSC organoids, seeded at a density of 125 cells/μl within Matrigel domes, and cultured over 11 days (Fig. 5C,D). Although the colony take-rate did not differ between *BCAM*<sup>hi</sup> and *BCAM*<sup>lo</sup> CTB organoids (Fig. S5A), the colony size of *BCAM*<sup>hi</sup> organoids appeared to be larger by day 7, with *BCAM*<sup>hi</sup>-derived colonies being significantly larger than *BCAM*<sup>lo</sup>-derived colonies on days 9–11 in culture (Fig. 5D,E). Flow cytometry assessment of day 11 *BCAM*<sup>hi</sup>- and *BCAM*<sup>lo</sup>-derived organoids showed that >90% of *BCAM*<sup>hi</sup>-sorted cells retained high levels of *BCAM* surface expression (Fig. 5F), whereas the frequency of CTB expressing high levels of *BCAM* in *BCAM*<sup>lo</sup>-sorted cells was 70%, suggesting that the *BCAM* phenotype is dynamic. Though a minor EGFR<sup>lo</sup> population was detected in endpoint organoids from *BCAM*<sup>hi</sup>-sorted CTB, a 3-fold expansion in this population was evident in organoids derived from *BCAM*<sup>lo</sup>-sorted CTB, suggesting that *BCAM*<sup>lo</sup> CTB do not retain regenerative properties as robustly as *BCAM*<sup>hi</sup> CTB (Fig. 5F). However, our observation that *BCAM*<sup>lo</sup> CTB generate a population of cells expressing high levels of *BCAM* suggests that the opposite may be true, that *BCAM*<sup>lo</sup> cells develop a greater potential for regenerative ability. To test the possibility that *BCAM* levels affect differentiation potential, *BCAM*<sup>hi</sup> and *BCAM*<sup>lo</sup> organoids were established and subjected to EVT differentiation over 10 days. Consistent with our previous findings, *BCAM*<sup>hi</sup> organoids formed larger colonies than *BCAM*<sup>lo</sup> organoids; however, EVT outgrowth did not differ (Fig. S5B). Similarly, expression and localization of α5 integrin, an immature EVT marker, was also not different between *BCAM* groups (Fig. S5C). Further, levels of early (*NOTCH1*) and late (*HLA-G*) EVT genes did not differ, suggesting that *BCAM*<sup>lo</sup> CTB do not have an increased differentiation potential (Fig. S5D). Comparing DEGs between CTB expressing high or low levels of *BCAM* in organoid scRNA-seq data showed that elevated *BCAM* levels associate with higher levels of the paternally expressed transcription factor *PEG10*, and the stem cell renewal genes *EPCAM* and *HMG3* (Nemeth et al., 2006), indicating a positive relationship between *BCAM* and regenerative ability (Fig. 5G).

To specifically examine whether *BCAM* plays a functional role in promoting trophoblast organoid growth, two *BCAM*-targeting siRNAs were transfected into CT29 hTSCs before and during organoid development, and organoid growth was monitored over 7 days. Non-transfected and non-silencing siRNA-transfected (siCTRL) CT29 hTSCs served as controls. siRNA-directed knockdown of *BCAM* showed a 40–60% reduction in *BCAM* protein levels at days 3 (mid) and 7 (endpoint) of culture (Fig. 6A, B). Although non-transfected and siCTRL organoids formed large and robust colonies, *BCAM* silencing led to a consistent reduction in organoid colony size (Fig. 6C). Whereas control hTSCs formed greater frequencies of medium ( $2.0\text{--}4.9\times 10^4\ \mu\text{m}^2$ ) and large ( $>5.0\times 10^4\ \mu\text{m}^2$ ) organoids, *BCAM*-silenced hTSCs formed greater frequencies of small ( $<2.0\times 10^4\ \mu\text{m}^2$ ) colonies (Fig. 6D). Similar to our observation that *BCAM*<sup>lo</sup>-sorted CTB showed an enhanced expansion of an EGFR<sup>lo</sup> population, *BCAM* silencing in organoids led to a modest but significant increase in the frequency of EGFR<sup>lo</sup> trophoblasts, suggesting that a reduction in *BCAM* may result in greater plasticity or accelerated differentiation (Fig. 6E). However, *BCAM* silencing in differentiating hTSC cultured in 2D did not alter cell morphology or day 3 levels of *NOTCH1*, *NOTCH2* and *HLA-G* (Fig. S5E,F). We attempted to also examine gene levels at day 6 of differentiation; however, iterative rounds of siRNA transfection resulted in high cell death in both controls and *BCAM*-silenced cultures, precluding the assessment of this time-point. Although adequate amounts of RNA from one day 6 experimental replicate shows that *HLA-G* levels increase as expected, levels did not differ between controls and *BCAM*-silenced cultures (Fig. S5F). We also examined whether *BCAM* affects the rate of CTB proliferation, and showed in day 7 endpoint organoids that *BCAM*-silencing resulted in a reduction in the proportion of BrdU<sup>+</sup> cells following a 4 h pulse compared with BrdU<sup>+</sup> frequencies in non-silencing controls, though this difference was moderate and did not reach statistical significance (Fig. 6F). Taken together, we show that *BCAM* identifies a CTB population with enhanced regenerative growth ability, but also that *BCAM* functionally contributes to CTB progenitor expansion.

### **DISCUSSION**

Here, we applied scRNA-seq of human placentas at two GA ranges within the first trimester to further our understanding of human trophoblast development. In doing this, we provide novel transcriptomic data from 7798 human trophoblasts, resolve the first trimester trophoblast landscape at increased resolution, and clarify how trophoblast heterogeneity changes during the first trimester. Using state-of-the-art single cell lineage trajectory tools, we trace back the lineage commitments made by developing CTB to identify a trophoblast progenitor origin marked by increased expression of the cell-surface glycoprotein *BCAM*, where *BCAM* was shown to provide a functional growth advantage in regenerative organoids. The association of *BCAM* with the predicted CTB origin highlights an additional utility of using *BCAM* as a new cell-surface progenitor marker for future studies identifying and defining CTB progenitors.

Previous studies have explored the placental transcriptome at single cell resolution (Liu et al., 2018; Suryawanshi et al., 2018; Vento-Tormo et al., 2018). Although these studies have contributed important insight into placental cell heterogeneity, understanding CTB hierarchy and gene pathways underlying CTB regeneration and commitment to EVT or SCT was not a focus of these studies. Our study identifies four unique CTB states that are distinguishable by state of metabolism (protein translation, CTB1; mitochondrial regulation, CTB3), cell cycle (CTB4), and molecular readouts of



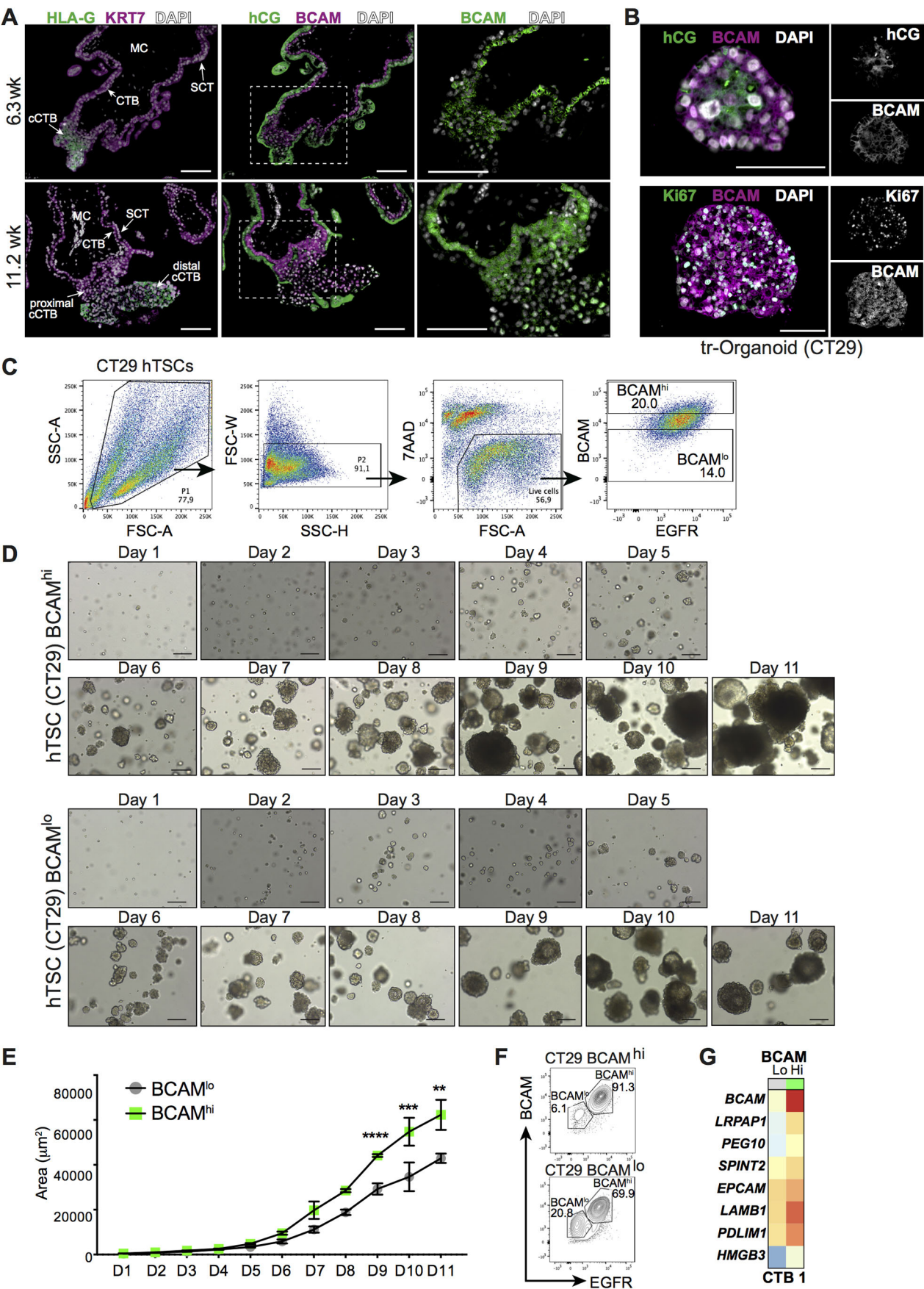


Fig. 5. See next page for legend.



**Fig. 5. BCAM enrichment establishes a growth advantage in trophoblast organoids.** (A) Immunofluorescence (IF) images of BCAM and KRT7, hCG and HLA-G in early (6.3 weeks) and late (11.2 weeks) first trimester placental villi. Nuclei are stained with DAPI. (B) IF images of BCAM, hCG and Ki67 in hTSC-derived organoid (CT29) cultured in regenerative (Wnt+) media for 7 days. (C) Cell sorting workflow for isolating BCAM<sup>hi</sup> and BCAM<sup>lo</sup> EGFR<sup>+</sup> CTB in hTSC (CT29)-derived organoids. (D) Brightfield images of hTSC (CT29) BCAM<sup>hi</sup>- and BCAM<sup>lo</sup>-derived organoids cultured over 11 days in regenerative Wnt<sup>+</sup> media. (E) Colony area ( $\mu\text{m}^2$ ) of BCAM<sup>hi</sup> and BCAM<sup>lo</sup> hTSC (CT29)-derived organoids cultured over 11 days (in D). Each condition was performed on three ( $n=3$ ) independent occasions. Data are median $\pm$ s.d. and statistical analyses were performed using a two-way ANOVA and multiple comparisons were controlled for using a Bonferroni post-test. \*\* $P<0.01$ ; \*\*\* $P<0.001$ ; \*\*\*\* $P<0.0001$ . (F) Flow cytometric panels showing expression of BCAM and EGFR in BCAM<sup>hi</sup> and BCAM<sup>lo</sup> CT29-derived organoids following 11 days of culture. (G) Heatmap showing levels of select DEGs from CTB expressing high (hi) or low (lo) levels of BCAM in the scRNA-seq-defined CTB1 state of CT29-derived organoids following 7 days of culture. Scale bars: 100  $\mu\text{m}$  (A); 200  $\mu\text{m}$  (B,D).

respiration and transporter activity (CTB2). Concurrent single cell analyses using hTSC-derived organoids identifies similar cell states to those found *in vivo*. However, differences in the number of cell clusters and in the expression patterns of specific genes in organoids derived from unique hTSC lines does suggest CTB clusters may represent snapshots of transient cell states captured during CTB maintenance and differentiation, and are not specific cell types per se.

Using scVelo, Monocle3 and Slingshot, we identified a CTB origin that likely represents a primitive stem cell-like niche. CTB2 was identified by scVelo, Monocle3 and Slingshot trajectory modeling tools as the CTB state overlapping with this predicted origin. We were initially surprised to see that the predicted origin did not directly align with CTB3, as this state showed robust levels of many classical CTB progenitor genes such as *YAP1*, *TEAD4* and *TP63*, whereas CTB2 expressed lower levels of these known trophoblast progenitor genes. However, CTB2 did show enrichment for some important CTB-associated genes such as *SPINT1* (Walentin et al., 2015), and further showed elevated levels for the Wnt ligand *WNT7A* and other Wnt-modifying factors such as *GPC3*, *DAB2* and *DACT2* (Jiang et al., 2012; Capurro et al., 2014; Wang et al., 2015), suggesting that this state is responsive to WNT signaling, an important feature underlying *in vitro* CTB progenitor maintenance (Okae et al., 2018).

Comparisons between our placental tissue sequencing data and our trophoblast organoid sequencing data demonstrate similarities in general CTB, SCT and EVT function, but also pinpoint subtle differences between the distinct states of CTB found in each stem cell-derived organoid condition and our tissue CTB. Comparison of the CTB progenitor niche identified in both the organoid and tissue datasets identifies the common transcription factors *YAP1*, *SPINT2* and *MSI1* as well as consistent upregulation of *BCAM*, supporting the designation of BCAM as a CTB progenitor marker. Interestingly, quantification of mRNA levels of known CTB-related genes (*TEAD4* and *ITGA6*) as well as SCT (*CGB3* and *CSH1*) and EVT related genes (*ITGA5*, *NOTCH1*, *NOTCH2* and *HLA-G*) on days 0, 11 and 21 of EVT differentiation indicate that hTSC-specific organoid colonies show different rates of differentiation or capacities to differentiate. Our EVT-differentiated organoid transcriptomes were sequenced on day 7 of differentiation, indicating that the differences we are observing may in part be due to groups of organoid colonies being at a different stages of differentiation, with CT29-derived organoids showing the most rapid capacity for EVT differentiation, followed

by CT27 and CT30. These differences in differentiation kinetics or capacities should be taken into consideration by researchers aiming to use these regenerative trophoblast lines for future research.

Integration of our chorionic villous and organoid datasets demonstrated a strong correlation between our *in vivo* and *in vitro* data, validating the use of hTSC-derived organoids as a model to study trophoblast differentiation. Interestingly, integration revealed that two CTB progenitor states exist in the first trimester placenta, one representing a primitive CTB origin state marked by elevated *BCAM*, and one representing a more specified column CTB progenitor cell state, marked by *ITGA2* and *NOTCH1*. In 8-9 week GA placental villi, *ITGA2* is specific to proximal cells of anchoring placental villi, near the basement membrane (Lee et al., 2018). It is thought that as chorionic villi develop, trophoblast progenitors within close proximity to uterine stroma adopt a unique gene expression profile that predisposes their differentiation along the extravillous pathway. We show that the transcriptome of *ITGA2*<sup>+</sup> column progenitors express EVT-associated genes such as *NOTCH1* and *ITGA5*, as well as the CTB/progenitor-associated genes *GATA3* and *ELF5*, suggesting that this population harbors progenitor characteristics with a preference to differentiate along the EVT pathway.

We identify *BCAM* as a gene expressed by CTB progenitors, where its expression is particularly elevated in the most upstream CTB states. BCAM interacts predominantly with its ligand LAMA5, a component of the laminin-511 heterotrimers found in the placental extracellular matrix (Zen et al., 1999; Kikkawa et al., 2013). BCAM is an antigen upregulated in ovarian carcinoma, and facilitates tumor migration by competing for laminin binding with integrins (Kannan et al., 2015; Guadall et al., 2019). Here, we show that flow cytometry enrichment of BCAM<sup>+</sup> CTB significantly potentiates organoid growth. Interestingly, no differences were observed in the capacity of BCAM<sup>hi</sup> or BCAM<sup>lo</sup> CTB to form organoid colonies from single cells, suggesting that BCAM<sup>hi</sup> CTB may not have intrinsically greater stem-like properties over BCAM<sup>lo</sup> CTB. Rather, BCAM may enhance organoid growth by potentiating survival signals induced by cell-ECM engagement, or by initiating/driving CTB proliferation pathways. However, we show that genes associated with stem cell regeneration are elevated in BCAM<sup>hi</sup> hTSC-derived organoids compared with organoids derived from BCAM<sup>lo</sup> cells. Specifically, we show increased expression of *EPCAM* and *HMGB3*, indicating that BCAM may confer increased organoid growth by controlling stem cell maintenance or regeneration. In our experimental organoid system, even though BCAM<sup>lo</sup> organoids form smaller colonies by day 11, assessment of BCAM expression in BCAM<sup>lo</sup> organoids at endpoint shows that BCAM levels and the frequency of BCAM-expressing cells is comparable with BCAM<sup>hi</sup> CTB. This suggests that BCAM expression in progenitor CTB is dynamic, and that requirements for growth, stemness or survival result in the induction of BCAM expression.

Interestingly, BCAM<sup>hi</sup> and BCAM<sup>lo</sup> CTB showed no difference in their capacity for extravillous differentiation; however, we did not explore the influence of BCAM expression on SCT differentiation. It is possible that BCAM regulates CTB differentiation along the villous path, with increased SCT production inside our 3D organoids serving as one explanation for the growth advantage observed in our BCAM<sup>hi</sup> cultures. Future experiments are needed to explore this possibility. Consistent with a role of BCAM in controlling organoid growth, BCAM silencing also resulted in the formation of smaller organoids, where cell cycle analysis suggests that BCAM regulates progenitor entry/exit into the cell cycle. Future research should focus on how

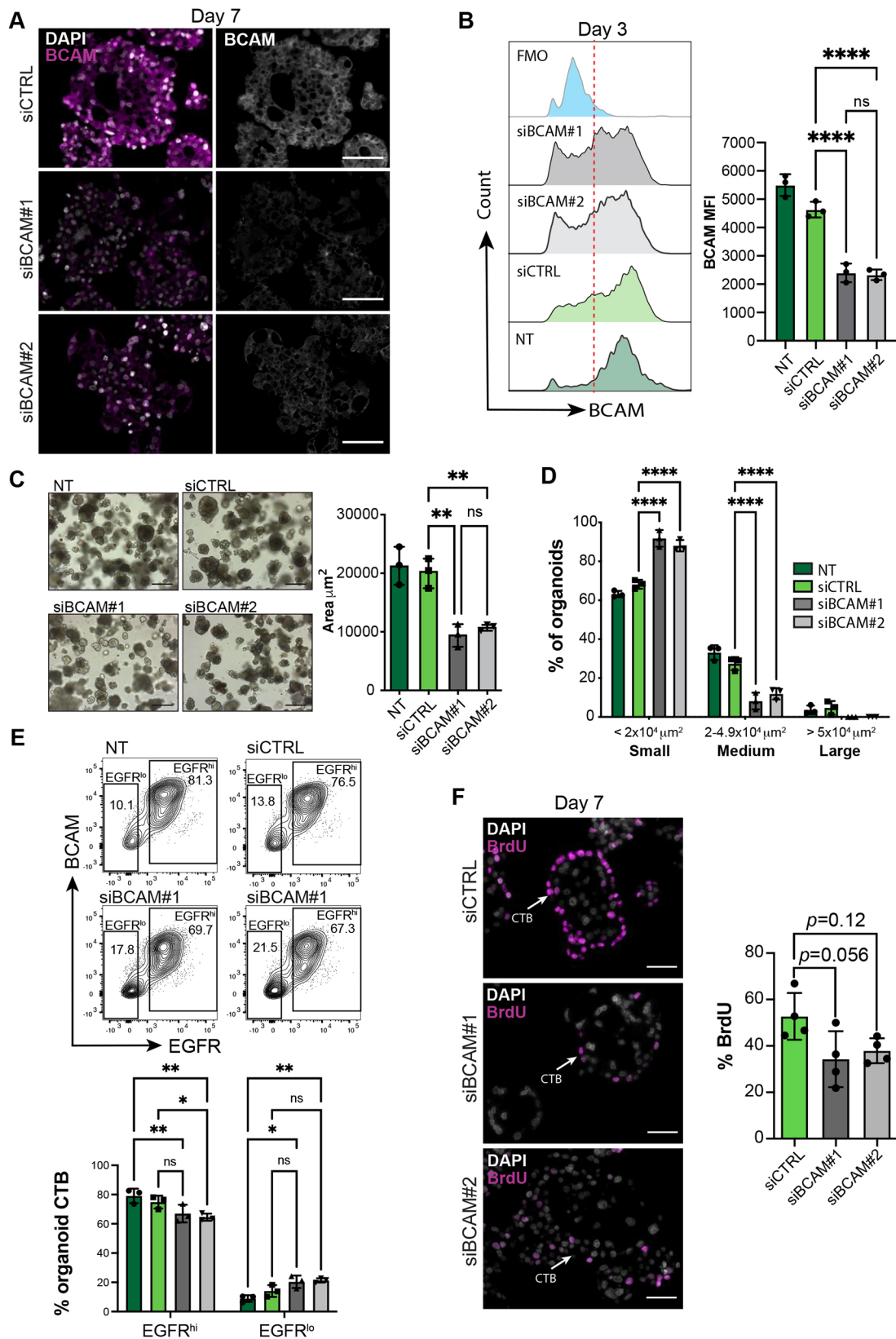


Fig. 6. See next page for legend.

BCAM potentially creates a supportive stem/regenerative cell niche, possibly by mediating specific cell-ECM engagement signals important in progenitor maintenance.

The identification of a novel cell-surface human CTB progenitor cell marker provides improved methods for isolating and studying the mechanisms regulating human CTB progenitor cells,

**Fig. 6. BCAM promotes CTB organoid growth.** (A) Immunofluorescence (IF) images of BCAM in CT29 organoids transfected with control siRNA (siCTRL) and BCAM-targeting siRNAs (siBCAM#1, siBCAM#2). Nuclei are stained with DAPI. (B) Representative flow cytometry plots showing BCAM levels in control (NT, siCTRL) and BCAM-silenced (siBCAM#1, siBCAM#2) CT29 organoids. FMO, fluorescence minus one. Bar graph to the right shows quantification of BCAM median fluorescence intensity (MFI). Statistical analyses between groups were performed using ANOVA and two-tailed Tukey post-test; significance  $P < 0.05$ . Each condition was performed on four ( $n=4$ ) independent occasions. (C) Brightfield images of CT29 organoids transfected with control (NT, siCTRL) or BCAM-silencing siRNA (siBCAM#1, siBCAM#2) following 7 days of culture in regenerative Wnt<sup>+</sup> media. Bar graph to the right shows area quantification of control and BCAM-silenced organoids. Median values are shown and statistical analyses between groups were performed using ANOVA and two-tailed Tukey post-test; significance  $P < 0.05$ . Each condition was performed on three ( $n=3$ ) independent occasions. (D) Bar graph shows proportions of small, medium and large colonies in control- and BCAM-silenced CT29 organoids. Median values are shown and statistical analyses were performed using a two-way ANOVA, multiple comparisons were controlled for using a Bonferroni post-test; significance  $P < 0.05$ . (E) Flow cytometry plots showing EGFR and BCAM frequencies in control and BCAM-silenced CT29 organoids at day 3 of organoid culture in regenerative media. Bar graphs show median values of EGFR<sup>hi</sup> and EGFR<sup>lo</sup> frequencies, and statistical analyses between groups were performed using ANOVA and two-tailed Tukey post-test; significance  $P < 0.05$ . Each condition was performed on three ( $n=3$ ) independent occasions. (F) IF images showing BrdU incorporation in control (siCTRL) or BCAM-silenced day 7 organoids following a 4 h BrdU pulse. Bar graphs show median values of BrdU frequencies in each group. Each condition was performed on three or four ( $n=3$ ,  $n=4$ ) independent occasions. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\*\* $P < 0.0001$ . ns, not significant. Data are median  $\pm$  s.d. Scale bars: 100  $\mu$ m (A); 200  $\mu$ m (C,F).

trophoblast differentiation and human placental development. We have used scRNA-seq to define trophoblast differentiation trajectories in early human placental development. Importantly, we identify BCAM as a new CTB progenitor marker that confers enhanced clonal growth. Future work is needed to clarify how increased BCAM expression facilitates improved CTB growth, and the exact molecular paths through which this occurs. We hope that future experiments will be able to capitalize on the identification of BCAM as a cell-surface CTB progenitor marker through which these cells can be isolated and further studied.

## MATERIALS AND METHODS

### Patient recruitment and tissue collection

Placental tissues were obtained with approval from the Research Ethics Board on the use of human subjects, University of British Columbia (H13-00640). All samples were collected from women (19 to 39 years of age) providing written informed consent undergoing elective terminations of pregnancy at British Columbia's Women's Hospital, Vancouver, Canada. First trimester placental tissues ( $n=9$ ) were collected from participating women (GA: 5-12 weeks) having confirmed viable pregnancies by ultrasound-measured fetal heartbeat. Patient clinical characteristics i.e., height and weight were additionally obtained to calculate body mass index (BMI: kg/m<sup>2</sup>) and all consenting women provided self-reported information via questionnaire to having first-hand exposure to cigarette smoke or taken anti-inflammation or anti-hypertensive medications during pregnancy. Patient and sample phenotype metadata are summarized in Table S1.

### Human trophoblast organoids

#### Establishment and culture

hTSC CT27 (RCB4936), CT29 (RCB4937) and CT30 (RCB4938) from Riken BRC Cell Bank, were established from first trimester CTB by Okae et al. (2018) and used here to generate organoid cultures following a modification of previously described protocols (Haider et al., 2018; Turco et al., 2018). In summary, these cells were re-suspended undiluted in ice-cold growth factor-reduced Matrigel (GFR-M, Corning). Cell suspensions (40  $\mu$ l) containing 5-20  $\times 10^3$  cells were carefully seeded in the center of each

well of a 24-well plate. Plates were placed at 37°C for 2 min. Following this, the plates were turned upside down to promote equal spreading of the cells throughout the Matrigel domes. After 15 min, the plates were flipped back to normal and the Matrigel domes were overlaid with 0.5 ml pre-warmed trophoblast organoid media containing advanced DMEM/F12 (Gibco), 10 mM HEPES, 1  $\times$  B27 (Gibco), 1  $\times$  N2 (Gibco), 2 mM L-glutamine (Gibco), 100  $\mu$ g/ml Primocin (Invivogen), 100 ng/ml R-spondin (PeproTech), 1  $\mu$ M A83-01 (Tocris), 100 ng/ml rhEGF (PeproTech), 3  $\mu$ M CHIR99021 (Tocris) and 2  $\mu$ M Y-27632 (EMD Millipore). Organoids were cultured for 7-21 days.

### EVT differentiation

EVT differentiation of trophoblast organoids was performed using an adaptation of a previously published protocol (Okoe et al., 2018). Briefly, organoids were cultured until >50% colonies reached at least 100  $\mu$ m in diameter (~3-5 days). Following this, organoids were maintained in EVT differentiation media containing advanced DMEM/F12, 0.1 mM 2-mercaptoethanol (Gibco), 0.5% penicillin/streptomycin (P/S) (Gibco), 0.3% bovine serum albumin (Sigma-Aldrich), 1% ITS-X supplement (Gibco), 100 ng/ml NRG1 (Cell Signaling Technology), 7.5  $\mu$ M A83-01 (Tocris) and 4% Knockout Serum Replacement (Gibco). Media was exchanged for EVT differentiation media lacking NRG1 at around day 7-10 of differentiation following the visible outgrowth of cells. Differentiation was stopped at day 7 for 10x scRNA-seq and flow cytometry analysis. Differentiation was stopped at day 21 for qPCR analysis.

### Placental tissue and trophoblast organoid scRNA-seq library construction

#### Placental tissue

Healthy placental villi were physically scraped from chorionic membrane by scalpel and enzymatically digested in 0.5 ml collagenase/hyaluronidase (Stemcell Technologies) and 1.5 ml DMEM/F12 with 500  $\mu$ l P/S and 500  $\mu$ l 100 $\times$  antibiotic/antimycotic at 37°C for 45 min. Placental villi cell suspensions were diluted in 1  $\times$  Hank's Balanced Salt Solution (HBSS) containing 2% fetal bovine serum (FBS), pelleted at 1200 rpm (280 g) for 5 min, and resuspended in 0.25% trypsin-EDTA. The solution was diluted again in 1  $\times$  HBSS containing 2% FBS, pelleted at 1200 rpm for 5 min, and resuspended in dispase and 1 mg/ml DNase I. Liberated cells were enriched by passing digest through a 40  $\mu$ m filter.

#### Trophoblast organoids

For all organoid cultures, media was aspirated and organoids were diluted in Cell Recovery Solution at 4°C for 40 min on either culture day 0 in complete media (non-differentiated organoids) or culture day 7 in differentiation media (EVT-differentiated organoids). Cells were then washed with cold PBS and centrifuged at 1000 rpm (195 g) for 5 min. Cells were disaggregated in TrypLE containing 5% DNase I at 37°C for 7 min before the addition of 100  $\mu$ l complete media. The resulting cell suspension was then sieved through a 40  $\mu$ m filter.

#### Library construction

Single cell suspensions (tissue- and organoid-derived) were stained with 7-Aminoactinomycin D (7-AAD) viability marker and live/dead sorted using a Becton Dickinson (BD) FACS Aria. Cell suspensions were loaded onto the 10x Genomics single cell controller for library preparation. Placental villi-derived single cell suspensions were prepared using the chromium single cell 3' reagent v2 chemistry kit, following standard manufacturer protocol for all steps. Non-differentiated and differentiated trophoblast organoid-derived single cell suspensions were prepared using the chromium single cell 3' reagent Next GEM v3.1 chemistry kit, following standard manufacturer protocol for all steps. Single cell libraries were sequenced on an Illumina Nextseq 500 instrument at a sequencing depth of 50,000 reads/cell for tissue-derived cells and 150,000 reads/cell for organoid-derived cells. For all samples, the 10x Genomics Cell Ranger version 3.0 software was used for STAR read alignment and counting against the hg19 reference transcriptome. scRNA-seq datasets are publicly available on the GEO repository (GSE174481).



### Data repository integration

Droplet-based first trimester scRNA-seq decidual and placental tissue data was obtained from the public repository ArrayExpress (E-MTAB-6701) (Vento-Tormo et al., 2018). Raw BAM files were downloaded and pre-processed using 10x Genomics Cell Ranger version 3.0 under identical parameters as previously described for our placental villi and trophoblast organoid samples, again using the hg19 reference. Sample specific metrics are summarized in Table S1.

### scRNA-seq data analysis

The code used for all subsequent data processing and analyses is available at <https://github.com/MatthewJShannon>.

### Data pre-processing and quality control

A total of 50,790 single cells from maternal-fetal interface samples and 6354 cells from hTSC organoids were sequenced and pre-processed using the Seurat R package (version 4.0.1) (Butler et al., 2018; Stuart et al., 2019). To ensure quality cells were used in downstream analyses, cells containing fewer than 1000 detected genes and greater than 20% mitochondrial DNA content were removed. Individual samples were scored based on expression of G2/M and S phase cell cycle gene markers, scaled, normalized and contaminating cell doublets were removed using the DoubletFinder package version 2.0.3 (McGinnis et al., 2019) with default parameters and a doublet formation rate estimated from the number of captured cells. To minimize bias due to read dropouts the 'FindVariableFeatures' function was used with a 'vst' selection method to prioritize highly variable genes in each sample for all downstream analyses. Following pre-processing, all samples were merged and integrated using cell pairwise correspondences between single cells across sequencing batches. During integration, the Pearson residuals obtained from the default regularized negative binomial model were re-computed and scaled to remove latent variation in percent mitochondrial DNA content as well as latent variation resulting from the difference between the G<sub>2</sub>M and S phase cell cycle scores.

### Cell clustering, identification and trophoblast sub-setting

Cell clustering was accomplished in Seurat at a resolution of 0.900 using 50 principal components. Cell identity was determined through observation of known gene markers found within the maternal fetal interface (vimentin, *PTPRC*, *NCAM1*, *CD14*, etc.) and identifying cluster marker genes via the 'FindAllMarkers' function using a model-based analysis of single cell transcriptomics (MAST) GLM-framework implemented in Seurat (Finak et al., 2015). Gene markers were selected as genes with a minimum log fold change value >0.20 and an adjusted *P*-value <0.05 within a specific cluster. Trophoblasts were identified as expressing *KRT7*, *EGFR*, *TEAD4*, *TP63*, *TFAP2A*, *TFAP2C*, *GATA2*, *GATA3*, *HLA-G*, *ERVFRD-1* transcripts at a level greater than zero, and *VIM*, *WARS1*, *PTPRC*, *DCN*, *CD34*, *CD14*, *CD86*, *CD163*, *NKG7*, *KLRD1*, *HLA-DPA1*, *HLA-DPB1*, *HLA-DRA*, *HLA-DRB1*, *HLA-DRB5*, *HLA-DQA1* transcripts at a level equal to zero. Trophoblasts were subset and re-clustered in Seurat at a resolution of 0.375 using 30 principal components. As mentioned previously, trophoblast cluster identity was determined through observation of known trophoblast gene marker expression, proliferative gene marker expression, and from unique cluster genes identified by differential expression analysis in Seurat. Identified cluster marker genes are summarized in Table S2. Trophoblast projections were visualized using Uniform Manifold Approximation and Projections (UMAPs). hTSC organoid trophoblasts were clustered at a resolution of 0.300 using 35 principal components. Again, trophoblast organoid cluster identity was determined through observation of known trophoblast gene marker expression, proliferative gene marker expression, and from unique cluster genes identified by differential expression analysis in Seurat, as previously described. Organoid projections were similarly visualized by UMAP. Clonal-specific hTSC organoids were subset and re-clustered at resolutions of 0.200, 0.325 and 0.250, using 20 principal components and visualized by UMAP for CT27 (2247 cells), CT29 (1445 cells) and CT30 (2603 cells), respectively. The integrated placental villous (*in vivo*) and hTSC-derived organoid (*in vitro*) trophoblast data were clustered at a resolution of 0.700 using 50 principal components.

Trophoblasts were subset and re-clustered in Seurat as previously described at a resolution of 0.300 using 30 principal components. Cluster identity was determined by observation of known trophoblast gene marker expression, proliferative gene marker expression, and from unique cluster genes identified by differential expression analysis in Seurat, as previously described. Identified cluster marker genes are summarized in Table S2. Integrated projections were similarly visualized by UMAP.

### Pseudobulk data comparison

Raw expression data were independently averaged and log transformed in each cell from GSE174481 and E-MTAB-6701 datasets to create 'pseudobulk' data for GSE174481 and E-MTAB-6701. Pseudobulk data were then plotted as a scatter plot using the ggplot2 package (version 3.3.5) with GSE174481 plotted on the x-axis and E-MTAB-6701 plotted on the y-axis. Pearson correlation coefficients of the relationship between these pseudobulk samples was calculated using the 'cor' function from the stats R package (version 4.0.5). For the correlation heatmap shown in Fig. S5B, the 'cor' function was used as described and the data was plotted as a heatmap using the ggplot2 package.

### Differential gene expression

Differential expression analyses were performed in Seurat using the 'FindMarkers' function using a MAST framework on the raw count data. Parameters were limited to genes with a log<sub>2</sub> fold-change difference between groups of 0.20 and genes detected in a minimum 10% cells in each group. Results were filtered for positive gene markers and significant DEGs were taken as having a fold-change difference >0.25 and an adjusted *P*-value <0.05. To enrich for the top genes, a modified minimum log<sub>2</sub> fold-change difference between groups of 0.30 was applied for the Venn diagram comparison of CTB2 and CTB3 shown in Fig. 2C. To include the whole-cell transcriptome in DEG analysis, parameters were modified to include a minimum log fold-change difference between groups of '-INF', a minimum gene detection of '-INF' in cells in each group and a minimum difference in the fraction of detected genes within each group of '-INF', allowing all genes to be compared. Results were visualized using the R package EnhancedVolcano (version 1.8.0), with a fold-change cut-off of 0.25 and a *P*-value cut off of 0.00005 taken as the thresholds for significance. For the volcano plot shown in Fig. S1F, a fold-change cut-off of 2.00 and a *P*-value cut off of 10e<sup>-100</sup> were taken as the thresholds for significance.

### Gene ontology

GO analysis of significant DEGs was performed using the clusterProfiler R package version 3.18.1 (Yu et al., 2012) and the top enriched terms (Benjamini-Hochberg corrected *P*-value <0.05) were visualized using the R package enrich plot (version 1.10.2).

### Pseudotime

The Monocle2 R package version 2.18.0 (Trapnell et al., 2014; Qiu et al., 2017a,b) was used to explore the differentiation of progenitor CTB into specialized SCT and EVT. The trophoblast cell raw counts were loaded into Monocle2 for semi-supervised single cell ordering in pseudotime, using *CDX2* and *SPINT1* expression to denote cells at the origin and *HLA-G* and *ERVFRD-1* gene expression to indicate progress towards EVT and SCT, respectively. Branched expression analysis modelling (BEAM) was applied to observe the expression of the top CTB2 progenitor gene markers (Table S2) over pseudotime, progressing towards either EVT or SCT. Results were visualized using the 'plot\_cell\_trajectory' function. To resolve CTB differentiation trajectories at higher resolution, the Monocle3 package version 1.1.0 (Cao et al., 2019) was used. Trophoblast raw counts were input, clustered, and the principal graph was created using the 'learn\_graph' function. The 'order\_cells' function was then used to order cells in pseudotime, designating the trophoblast root within the CTB2 state. In addition to Monocle, the Slingshot R package (version 1.8.0; Street et al., 2018) was used to further explore CTB differentiation trajectories. Trophoblast raw counts were input and the 'slingshot' wrapper function was used for trajectory inference using the clusters identified in Seurat. The

trophoblast root was designated in the CTB2 state and the end clusters were designated as either the EVT or SCTp state. The 'plotGenePseudotime' function was then used under default parameters to visualize raw gene expression levels in individual cells across pseudotime for each lineage identified from the 'slingshot' function output.

### Lineage trajectory

The Velocity package (La Manno et al., 2018) was run in terminal using the command line interface to generate sample-specific files containing count matrices of unspliced, spliced and ambiguous RNA transcript abundances. The velocity.R (version 0.6) implementation of the package was then used to merge each file with its corresponding 10x sample during data pre-processing in Seurat. The final, processed trophoblast object and UMAP embeddings were exported to a Jupyter Notebook and the scVelo Python package (Bergen et al., 2020) was used to read, prepare and recover holistic trophoblast gene splicing kinetics. Velocities were projected on top of the extracted UMAP embeddings generated in Seurat.

### Flow cytometry analysis and cell sorting

#### Organoid single cell workflow

Single cells from CT29 organoids were isolated by incubation with Cell Recovery Solution (Corning) for 40 min at 4°C. Then, cells were washed in ice-cold PBS and dissociated into single cells using TrypLE™ supplemented with 5% DNase I for 7 min at 37°C.

#### Phenotyping EVT differentiated CT27, CT29 and CT30 organoids

Following single cell dissociation, organoid cells were stained with EGFR BV605 (clone AY13; 1:50; Biolegend), HLA-G APC (clone MEMG-9; 1:50; Abcam), HLA-ABC Pacific Blue (clone W6/32; 1:100; Biolegend), CD49e/ITGA5 PeCy7 (clone eBioGoH3; 1:200; Thermo Fisher Scientific), CD49f/ITGA6 Alexa 488 (clone P1D6; 1:200; Thermo Fisher Scientific) and BCAM BV711 (clone B64; 1:50; BD Biosciences).

#### Phenotyping BCAM<sup>hi</sup> and BCAM<sup>lo</sup> organoids

Single cells were incubated with anti-human BCAM BV421 (clone B64; 1:50; BD Biosciences), EGFR BV605 and CD49f/ITGA6 PeCy7 for 30 min in the dark on ice, followed by viability staining with 7-AAD. Samples were analyzed using LSRII (Becton Dickinson) and a minimum of 20,000 events were acquired.

#### hTSC BCAM<sup>hi/lo</sup> sorting

CT29 hTSCs were dissociated using TrypLE™ (Thermo Fisher Scientific) with 5% DNase I (Sigma-Aldrich) for 7 min at 37°C. Following resuspension in PBS supplemented with antibiotic-antimycotic (Gibco) and 1% P/S (Sigma-Aldrich), cells were stained with BCAM BV421 and EGFR PeCy7 (clone AY13; 1:50; Biolegend) for 30 min at 4°C in the dark, followed by incubation with 7-AAD (Thermo Fisher Scientific) for viability assessment. CT29 BCAM<sup>hi</sup> and BCAM<sup>lo</sup> populations were sorted using a FACS Aria (BD Biosciences).

#### Trophoblast organoid growth assay

CT29 sorted BCAM<sup>hi</sup> and BCAM<sup>lo</sup> cells were seeded at a density of 5000 cells in 40 µl of Matrigel in the presence of trophoblast organoid media. Phase contrast images were taken with AxioObserver inverted microscope (Carl Zeiss) using a 20× objective, and organoids were further analyzed by qPCR, immunofluorescence and flow cytometry. Organoid formation efficiency was assessed by measuring organoid number and size (area in µm<sup>2</sup>). Images were taken daily and analyzed using ImageJ (National Institutes of Health). The scale was set by measuring the scale bar in pixels (285.5 pixels=200 µm).

#### Trophoblast organoid EVT differentiation assay

The CT29 hTSC line was used to FACS sort BCAM<sup>hi</sup> and BCAM<sup>lo</sup> cells which were seeded at a density of 4000 cells in 40 µl of Matrigel in the presence of trophoblast organoid media and cultured for 5 days. Following this, organoids were maintained in EVT differentiation media (described above) for 7 days. Brightfield images were taken with AxioObserver

inverted microscope (Carl Zeiss) using 20× objective, and differentiated organoids were further analyzed by qPCR and immunofluorescence.

### RNA isolation

Prior to RNA extraction, organoids were dissociated into single cells as described above. Total RNA was prepared from these cells using RiboZol reagent (VWR Life Science) followed by RNeasy Mini kit (Qiagen) according to the manufacturers' protocols. RNA purity was confirmed using a NanoDrop Spectrophotometer (Thermo Fisher Scientific).

### cDNA synthesis and qPCR

RNA was reverse transcribed using qScript cDNA SuperMix synthesis kit (Quanta Biosciences) and subjected to qPCR (ΔΔCT) analysis, using PowerUP SYBR Green Master Mix (Thermo Fisher Scientific) on a StepOnePlus Real-Time PCR System (Thermo Fisher Scientific) and forward and reverse primer sets for *TEAD4* (F: 5'-CGGCACCATTCACCTC-CAACG-3'; R: 5'-CTGCGTCATTGTGTCGATGGGC-3'), *ITGA6* (F: 5'-CA-CTCGGGAGGACAACGTGA-3'; R: 5'-ACAGCCCTCCCGTTCTGTG-3'), *CEB3* (F: 5'-GCCTCATCCTTGGCGCTAGA-3'; R: 5'-TATACCTC-GGGGTTGTGGGG-3'), *CSH1* (F: 5'-GACTGGGCAGATCCTCAAG-CA-3'; R: 5'-CCATGCGCAGGAATGTCTCG-3'), *ITGA5* (F: 5'-GGGT-CGGGGGCTTCACTTA-3'; R: 5'-ACACTGACCCCGTCTGTTCC-3'), *HLA-G* (F: 5'-CCCACGCACAGACTGACAGA-3'; R: 5'-AGGTGCGA-GCCAATCATCA-3'), *NOTCH1* (F: 5'-GCCTGAATGGCGGGAAGT-GT-3'; R: 5'-TAGTCTGCCACGCCTCTGC-3'), *NOTCH2* (F: 5'-ACTC-GGGGCTACTCTGTGA-3'; R: 5'-TGTCTCCCTACAACGCTCC-3'), *GAPDH* (F: 5'-AGGGCTGCTTTAACTCTGGT-3'; R: 5'-CCCCACTT-GATTTTGGAGGGA-3'). All raw data were analyzed using StepOne Software v2.3 (Thermo Fisher Scientific). The threshold cycle values were used to calculate relative gene expression levels. Values were normalized to endogenous *GAPDH* transcripts.

### Immunofluorescence microscopy

Placental villi (6-12 weeks GA; n=2) were fixed in 2% paraformaldehyde overnight at 4°C. Tissues were paraffin embedded and sectioned at 6 µm onto glass slides. Immunofluorescence was performed as previously described. Briefly, organoids or placental tissues underwent antigen retrieval by heating slides in a microwave for 5×2-min intervals in citrate buffer (pH 6.0). Sections were incubated with sodium borohydride for 5 min at room temperature (RT), followed by Triton X-100 permeabilization for 5 min at RT (Aghababaei et al., 2015). Slides were blocked in 5% normal goat serum/0.1% saponin for 1 h at RT and incubated with combinations of the indicated antibodies overnight at 4°C: mouse monoclonal anti-HLA-G (clone 4H84; 1:100; Exbio); rabbit monoclonal anti-cytokeratin 7 (clone SP52; 1:50; Ventana Medical Systems), rabbit polyclonal anti-BCAM (1:100; ab244280; Abcam); mouse monoclonal anti-hCG beta (clone 5H4-E2; 1:100; Abcam), anti-BrdU (1:1000; Bu20a; Cell Signaling Technology), anti-α5 integrin (1:100; 4267; Cell Signaling Technology), anti-EGFR (1:50; D38B1; Cell Signaling Technology). Following overnight incubation, sections and coverslips were washed with PBS and incubated with Alexa Fluor goat anti-rabbit 568 and goat anti-mouse 488 conjugated secondary antibodies (1:200; A11036, A32723; Life Technologies) for 1 h at RT, washed in PBS and mounted using ProLong Gold mounting media containing DAPI (Life Technologies).

Slides were imaged using an AxioObserver inverted microscope (Carl Zeiss) using 20× Plan-Apochromat/0.80NA or 40× Plan-Apochromat oil/1.4NA objectives (Carl Zeiss). An ApoTome .2 structured illumination device (Carl Zeiss) set at full Z-stack mode and five phase images was used for image acquisition.

### siRNA knockdown of BCAM

Two BCAM-targeting siRNAs (siBCAM #1, J-010608-05-0002; siBCAM #2, J-010608-06-0002; Dharmacon) were transfected into CT29 hTSCs using Lipofectamine RNAiMAX (Thermo Fisher Scientific) at a concentration of 100 nM. For controls, cells were transfected with ON-TARGETplus non-silencing siRNA (siCTRL; D-001810-01-20; Dharmacon) or cultured in the presence of transfection reagent alone

(non-treated; NT). CT29 hTSCs cultured in 2D were seeded then transfected using Lipofectamine RNAiMAX (Thermo Fisher Scientific) to achieve efficient *BCAM* silencing before organoid establishment. Following 48 h, hTSCs cells were transfected in suspension before seeding with *BCAM* and CTRL siRNAs for a second time and embedded in Matrigel to initiate 3D trophoblast organoids or cultured in 2D with EVT differentiation media following the protocol described in Okae et al. (2018) for mRNA assessment. After 3 or 7 days post-transfection, organoids were either cell dissociated from Matrigel using Cell Recovery Solution (BD Biosciences) for *BCAM* cell-surface quantification using flow cytometry, or fixed in 4% paraformaldehyde (PFA), embedded in paraffin and sectioned onto glass slides for immunofluorescence (IF) assessment of *BCAM*, respectively.

### BrdU incorporation assay

Pulse-chase labeling with BrdU (Sigma-Aldrich) was conducted on CT29 hTSC-derived organoids. Two days following 3D hTSC organoid establishment and reverse control (siCTRL) and *BCAM* siRNA (siBCAM#1, siBCAM#2) transfection, organoids were exposed to a 4 h pulse with culture media containing 10  $\mu$ M BrdU. Following 4 h of labeling, organoids were washed in PBS and fixed in 4% PFA overnight. Organoids were paraffin embedded and sectioned for IF microscopy: IF staining with anti-BrdU antibody (1:200, Bu20a, Sigma-Aldrich) with the addition of a 30-min incubation in 2 M hydrochloric acid between permeabilization and sodium borohydride steps.

### Statistical analysis

Data are reported as median values with standard deviations. All calculations were carried out using GraphPad Prism 9 software. For single comparisons, Mann–Whitney non-parametric unpaired *t*-tests were performed. For multiple comparisons, one-way Kruskal–Wallis ANOVA followed by Dunn's multiple comparison test was performed. One-way ANOVA followed by Tukey post-test was performed for all other multiple comparisons. The differences were accepted as significant at  $P < 0.05$ . For scRNA-seq statistical analyses, please refer to 'scRNA-seq data analysis' section in Materials and Methods.

### Acknowledgements

The authors extend their sincere gratitude to the hard work of staff at British Columbia's Women's Hospital's CARE Program for recruiting participants to our study, and the staff at the University of British Columbia Biomedical Research Centre, especially Yiwei Zhao and Ryan Vander Werff, for assistance with all drop-seq library preparation and sequencing. We also wish to acknowledge the support of Dr Francis Lynn (University of British Columbia) in providing access to and training in using his laboratory's 10x Genomics Chromium system.

### Competing interests

The authors declare no competing or financial interests.

### Author contributions

Conceptualization: M.J.S., A.G.B.; Methodology: M.J.S., J.B., B.C., J.W., G.L.M., J.S.Y., J.T., H.T.L.; Formal analysis: M.J.S., B.C., A.G.B.; Investigation: M.J.S., J.B., J.W., G.L.M.; Resources: A.G.B.; Writing - original draft: M.J.S., A.G.B.; Writing - review & editing: J.B., B.C., J.W., G.L.M., J.S.Y., J.T., H.T.L., P.M.L., A.G.B.; Supervision: A.G.B.; Project administration: A.G.B.; Funding acquisition: P.M.L., A.G.B.

### Funding

This work was supported by Natural Sciences and Engineering Research Council of Canada Discovery (RGPIN-2020-05378) and Accelerator (RGPAS-2020-00013) Grants (to A.G.B.), a Canadian Institutes of Health Research operating grant (201809PJT-407571-CIA-CAAA to A.G.B.), a BC Children's Hospital Foundation Catalyst Grant (to P.M.L.), and a Canadian Institutes of Health Research Master's graduate studentship (to M.J.S.).

### Data availability

The ArrayExpress accession number for the publicly available data reported in this paper is E-MTAB-6701. The GEO accession number for publicly available data reported in this paper is GSE174481.

### Peer review history

The peer review history is available online at <https://journals.biologists.com/dev/article-lookup/doi/10.1242/dev.199840>.

### References

- Aghababaei, M., Hogg, K., Perdu, S., Robinson, W. P. and Beristain, A. G. (2015). ADAM12-directed ectodomain shedding of E-cadherin potentiates trophoblast fusion. *Cell Death Differ.* **22**, 1970–1984. doi:10.1038/cdd.2015.44
- 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. doi:10.1038/s41587-020-0591-3
- Bhattacharya, S., Das, A., Mallya, K. and Ahmad, I. (2007). Maintenance of retinal stem cells by *Abcg2* is regulated by notch signaling. *J. Cell Sci.* **120**, 2652–2662. doi:10.1242/jcs.008417
- Burton, G. J., Cindrova-Davies, T. and Turco, M. Y. (2020). Review: Histotrophic nutrition and the placental-endometrial dialogue during human early pregnancy. *Placenta* **102**, 21–26. doi:10.1016/j.placenta.2020.02.008
- Butler, A., Hoffman, P., Smibert, P., Papalexi, E. and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420. doi:10.1038/nbt.4096
- 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. doi:10.1038/s41586-019-0969-x
- Capurro, M., Martin, T., Shi, W. and Filmus, J. (2014). Glypican-3 binds to Frizzled and plays a direct role in the stimulation of canonical Wnt signaling. *J. Cell Sci.* **127**, 1565–1575. doi:10.1242/jcs.140871
- Chang, C.-W., Wakeland, A. K. and Parast, M. M. (2018). Trophoblast lineage specification, differentiation and their regulation by oxygen tension. *J. Endocrinol.* **236**, R43–R56. doi:10.1530/JOE-17-0402
- Fatima, S., Zhou, S. and Sorrentino, B. P. (2012). *Abcg2* expression marks tissue-specific stem cells in multiple organs in a mouse progeny tracking model. *Stem Cells* **30**, 210–221. doi:10.1002/stem.1002
- Finak, G., McDavid, A., Yajima, M., Deng, J., Gersuk, V., Shalek, A. K., Slichter, C. K., Miller, H. W., McElrath, M. J., Plic, M. et al. (2015). MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* **16**, 278. doi:10.1186/s13059-015-0844-5
- Fowden, A. L., Forhead, A. J., Sferuzzi-Perri, A. N., Burton, G. J. and Vaughan, O. R. (2015). Review: Endocrine regulation of placental phenotype. *Placenta* **36**, S50–S59. doi:10.1016/j.placenta.2014.11.018
- Guadall, A., Cochet, S., Renaud, O., Colin, Y., Le Van Kim, C., De Brevern, A. G. and El Nemer, W. (2019). Dimerization and phosphorylation of Lutheran/basal cell adhesion molecule are critical for its function in cell migration on laminin. *J. Biol. Chem.* **294**, 14911–14921. doi:10.1074/jbc.RA119.007521
- Haider, S., Meinhardt, G., Velicky, P., Otti, G. R., Whitley, G., Fiala, C., Pollheimer, J. and Knöfler, M. (2014). Notch signaling plays a critical role in motility and differentiation of human first-trimester cytotrophoblasts. *Endocrinology* **155**, 263–274. doi:10.1210/en.2013-1455
- Haider, S., Meinhardt, G., Saleh, L., Fiala, C., Pollheimer, J. and Knöfler, M. (2016). Notch1 controls development of the extravillous trophoblast lineage in the human placenta. *Proc. Natl Acad. Sci. USA* **113**, E7710–E7719. doi:10.1073/pnas.1612335113
- Haider, S., Meinhardt, G., Saleh, L., Kunihs, V., Gamperl, M., Kaindl, U., Ellinger, A., Burkard, T. R., Fiala, C., Pollheimer, J. et al. (2018). Self-renewing trophoblast organoids recapitulate the developmental program of the early human placenta. *Stem Cell Rep.* **11**, 537–551. doi:10.1016/j.stemcr.2018.07.004
- Jauniaux, E., Moffett, A. and Burton, G. J. (2020). Placental implantation disorders. *Obstet. Gynecol. Clin. North Am.* **47**, 117–132. doi:10.1016/j.ogc.2019.10.002
- Jiang, Y., He, X. and Howe, P. H. (2012). Disabled-2 (Dab2) inhibits Wnt/ $\beta$ -catenin signalling by binding LRP6 and promoting its internalization through clathrin: Dab2 regulates clathrin-mediated endocytosis of LRP6. *EMBO J.* **31**, 2336–2349. doi:10.1038/emboj.2012.83
- Kannan, K., Coarfa, C., Chao, P.-W., Luo, L., Wang, Y., Brinegar, A. E., Hawkins, S. M., Milosavljevic, A., Matzuk, M. M. and Yen, L. (2015). Recurrent *BCAM*-AKT2 fusion gene leads to a constitutively activated AKT2 fusion kinase in high-grade serous ovarian carcinoma. *Proc. Natl Acad. Sci. USA* **112**, E1272–E1277. doi:10.1073/pnas.1501735112
- Kikkawa, Y., Ogawa, T., Sudo, R., Yamada, Y., Katagiri, F., Hozumi, K., Nomizu, M. and Miner, J. H. (2013). The Lutheran/Basal cell adhesion molecule promotes tumor cell migration by modulating integrin-mediated cell attachment to laminin-511 protein. *J. Biol. Chem.* **288**, 30990–31001. doi:10.1074/jbc.M113.486456
- Knöfler, M., Haider, S., Saleh, L., Pollheimer, J., Gamage, T. K. J. B. and James, J. (2019). Human placenta and trophoblast development: key molecular mechanisms and model systems. *Cell. Mol. Life Sci.* **76**, 3479–3496. doi:10.1007/s00018-019-03104-6



- La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastrioti, M. E., Lönnerberg, P., Furlan, A. et al. (2018). RNA velocity of single cells. *Nature* **560**, 494–498. doi:10.1038/s41586-018-0414-6
- Lee, C. Q. E., Gardner, L., Turco, M., Zhao, N., Murray, M. J., Coleman, N., Rossant, J., Hemberger, M. and Moffett, A. (2016). What is trophoblast? A combination of criteria define human first-trimester trophoblast. *Stem Cell Rep.* **6**, 257–272. doi:10.1016/j.stemcr.2016.01.006
- Lee, C. Q. E., Turco, M., Gardner, L., Simons, B., Hemberger, M. and Moffett, A. (2018). Integrin  $\alpha 2$  marks a niche of trophoblast progenitor cells in first trimester human placenta. *Development* **145**, dev162305. doi:10.1242/dev.162305
- Liu, Y., Fan, X., Wang, R., Lu, X., Dang, Y.-L., Wang, H., Lin, H.-Y., Zhu, C., Ge, H., Cross, J. C. et al. (2018). Single-cell RNA-seq reveals the diversity of trophoblast subtypes and patterns of differentiation in the human placenta. *Cell Res.* **28**, 819–832. doi:10.1038/s41422-018-0066-y
- McGinnis, C. S., Murrow, L. M. and Gartner, Z. J. (2019). DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. *Cell Systems* **8**, 329–337.e4. doi:10.1016/j.cels.2019.03.003
- Moffett, A., Chazara, O. and Colucci, F. (2017). Maternal allo-recognition of the fetus. *Fertil. Steril.* **107**, 1269–1272. doi:10.1016/j.fertnstert.2017.05.001
- Nemeth, M. J., Kirby, M. R. and Bodine, D. M. (2006). Hmgb3 regulates the balance between hematopoietic stem cell self-renewal and differentiation. *Proc. Natl Acad. Sci. USA* **103**, 13783–13788. doi:10.1073/pnas.0604006103
- Okae, H., Toh, H., Sato, T., Hiura, H., Takahashi, S., Shirane, K., Kabayama, Y., Suyama, M., Sasaki, H. and Arima, T. (2018). Derivation of human trophoblast stem cells. *Cell Stem Cell* **22**, 50–63.e6. doi:10.1016/j.stem.2017.11.004
- Papuchova, H., Kshirsagar, S., Xu, L., Bougleux Gomes, H. A., Li, Q., Iyer, V., Norwitz, E. R., Strominger, J. L. and Tilburgs, T. (2020). Three types of HLA-G+ extravillous trophoblasts that have distinct immune regulatory properties. *Proc. Natl Acad. Sci. USA* **117**, 15772–15777. doi:10.1073/pnas.2000484117
- Pollheimer, J., Vondra, S., Baltayeva, J., Beristain, A. G. and Knöfler, M. (2018). Regulation of placental extravillous trophoblasts by the maternal uterine environment. *Front. Immunol.* **9**, 2597. doi:10.3389/fimmu.2018.02597
- Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H. A. and Trapnell, C. (2017a). Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982. doi:10.1038/nmeth.4402
- Qiu, X., Hill, A., Packer, J., Lin, D., Ma, Y.-A. and Trapnell, C. (2017b). Single-cell mRNA quantification and differential analysis with Censur. *Nat. Methods* **14**, 309–315. doi:10.1038/nmeth.4150
- Saha, B., Ganguly, A., Home, P., Bhattacharya, B., Ray, S., Ghosh, A., Rumi, M. A. K., Marsh, C., French, V. A., Gunewardena, S. et al. (2020). TEAD4 ensures postimplantation development by promoting trophoblast self-renewal: An implication in early human pregnancy loss. *Proc. Natl Acad. Sci. USA* **117**, 17864–17875. doi:10.1073/pnas.2002449117
- Street, K., Risso, D., Fletcher, R. B., Das, D., Ngai, J., Yosef, N., Purdom, E. and Dudoit, S. (2018). Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477. doi:10.1186/s12864-018-4772-0
- 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. doi:10.1016/j.cell.2019.05.031
- Suryawanshi, H., Morozov, P., Straus, A., Sahasrabudhe, N., Max, K. E. A., Garzia, A., Kustagi, M., Tuschl, T. and Williams, Z. (2018). A single-cell survey of the human first-trimester placenta and decidua. *Sci. Adv.* **4**, eaau4788. doi:10.1126/sciadv.aau4788
- Tilburgs, T., Crespo, Á. C., Van Der Zwan, A., Rybalov, B., Raj, T., Stranger, B., Gardner, L., Moffett, A. and Strominger, J. L. (2015). Human HLA-G+ extravillous trophoblasts: immune-activating cells that interact with decidual leukocytes. *Proc. Natl Acad. Sci. USA* **112**, 7219–7224. doi:10.1073/pnas.1507977112
- 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. doi:10.1038/nbt.2859
- Treissman, J., Yuan, V., Baltayeva, J., Le, H. T., Castellana, B., Robinson, W. P. and Beristain, A. G. (2020). Low oxygen enhances trophoblast column growth by potentiating differentiation of the extravillous lineage and promoting LOX activity. *Development* **147**, dev181263. doi:10.1242/dev.181263
- Tunster, S. J., Creeth, H. D. J. and John, R. M. (2016). The imprinted Phlda2 gene modulates a major endocrine compartment of the placenta to regulate placental demands for maternal resources. *Dev. Biol.* **409**, 251–260. doi:10.1016/j.ydbio.2015.10.015
- Turco, M. Y. and Moffett, A. (2019). Development of the human placenta. *Development* **146**, dev163428. doi:10.1242/dev.163428
- Turco, M. Y., Gardner, L., Kay, R. G., Hamilton, R. S., Prater, M., Hollinshead, M. S., Mcwhinnie, A., Esposito, L., Fernando, R., Skelton, H. et al. (2018). Trophoblast organoids as a model for maternal–fetal interactions during human placentation. *Nature* **564**, 263–267. doi:10.1038/s41586-018-0753-3
- Vento-Tormo, R., Efremova, M., Botting, R. A., Turco, M. Y., Vento-Tormo, M., Meyer, K. B., Park, J.-E., Stephenson, E., Polański, K., Goncalves, A. et al. (2018). Single-cell reconstruction of the early maternal–fetal interface in humans. *Nature* **563**, 347–353. doi:10.1038/s41586-018-0698-6
- Walentin, K., Hinze, C., Werth, M., Haase, N., Varma, S., Morell, R., Aue, A., Pötschke, E., Warburton, D., Qiu, A. et al. (2015). A Grhl2-dependent gene network controls trophoblast branching morphogenesis. *Development* **142**, 1125–1136. doi:10.1242/dev.113829
- Wang, S., Dong, Y., Zhang, Y., Wang, X., Xu, L., Yang, S., Li, X., Dong, H., Xu, L., Su, L. et al. (2015). DACT2 is a functional tumor suppressor through inhibiting Wnt/ $\beta$ -catenin pathway and associated with poor survival in colon cancer. *Oncogene* **34**, 2575–2585. doi:10.1038/onc.2014.201
- Wang, X., Li, G., Koul, S., Ohki, R., Maurer, M., Borczuk, A. and Halmos, B. (2018). PHLDA2 is a key oncogene-induced negative feedback inhibitor of EGFR/Erbb2 signaling via interference with AKT signaling. *Oncotarget* **9**, 24914–24926. doi:10.18632/oncotarget.3674
- Yu, G., Wang, L.-G., Han, Y. and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS J. Integr. Biol.* **16**, 284–287. doi:10.1089/omi.2011.0118
- Zen, Q., Cottman, M., Truskey, G., Fraser, R. and Telen, M. J. (1999). Critical factors in basal cell adhesion molecule/lutheran-mediated adhesion to Laminin. *J. Biol. Chem.* **274**, 728–734. doi:10.1074/jbc.274.2.728
- Zhou, Y., Fisher, S. J., Janatpour, M., Genbacev, O., Dejana, E., Wheelock, M. and Damsky, C. H. (1997). Human cytotrophoblasts adopt a vascular phenotype as they differentiate. A strategy for successful endovascular invasion? *J. Clin. Invest.* **99**, 2139–2151. doi:10.1172/JCI119387