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

Matthew J Shannon, Jennet Baltayeva, Barbara Castellana, Jasmin Wächter, Gina L McNeill, Ji Soo Yoon, Jenna Treissman, Hoa T Le, Pascal M Lavoie and Alexander G Beristain
DOI: 10.1242/dev.199840

Editor: Samantha Morris

Review timeline

Original submission:	26 May 2021
Editorial decision:	12 July 2021
First revision received:	28 October 2021
Editorial decision:	29 November 2021
Second revision received:	3 December 2021
Accepted:	7 December 2021

Original submission

First decision letter

MS ID#: DEVELOP/2021/199840

MS TITLE: Cell trajectory modeling identifies a primitive trophoblast state defined by BCAM enrichment

AUTHORS: Matthew J Shannon, Jennet Baltayeva, Barbara Castellana, Jasmin WÃ¤chter, Ji Soo Yoon, Jenna Treissman, Hoa T Le, Pascal M Lavoie, and Alexander G Beristain

I have now received all the referees' reports on the above manuscript, and have reached a decision. The referees' comments are appended below, or you can access them online: please go to BenchPress and click on the 'Manuscripts with Decisions' queue in the Author Area.

As you will see, the referees express considerable interest in your work, but have some significant criticisms and recommend a substantial revision of your manuscript before we can consider publication. If you are able to revise the manuscript along the lines suggested, which may involve further experiments, I will be happy receive a revised version of the manuscript. Your revised paper will be re-reviewed by one or more of the original referees, and acceptance of your manuscript will depend on your addressing satisfactorily the reviewers' major concerns. Please also note that Development will normally permit only one round of major revision.

We are aware that you may be experiencing disruption to the normal running of your lab that make experimental revisions challenging. If it would be helpful, we encourage you to contact us to discuss your revision in greater detail. Please send us a point-by-point response indicating where you are able to address concerns raised (either experimentally or by changes to the text) and where you will not be able to do so within the normal timeframe of a revision. We will then provide further guidance. Please also note that we are happy to extend revision timeframes as necessary.

Please attend to all of the reviewers' comments and ensure that you clearly highlight all changes made in the revised manuscript. Please avoid using 'Tracked changes' in Word files as these are lost in PDF conversion. I should be grateful if you would also provide a point-by-point response detailing

how you have dealt with the points raised by the reviewers in the 'Response to Reviewers' box. If you do not agree with any of their criticisms or suggestions please explain clearly why this is so.

Reviewer 1

Advance summary and potential significance to field

This manuscript describes single cell RNA sequencing (scRNAseq) of early and late first trimester human placenta, and describing different cytotrophoblast (CTB) progenitor states, leading to identification of BCAM-high cells as potential progenitor CTB, whose enrichment leads to enhanced growth of trophoblast organoids.
While there have been a lot of publications in this area, this manuscript advances the field by describing different CTB progenitor states, leading to identification as well as further characterization of one such CTB state. The authors are careful not to over-interpret their data; nevertheless, the manuscript suffers from a lack of clarity in several areas, appearing at times ready to "agree" with previously-published studies, without sufficient data. Specific suggestions for revision are listed below.

Comments for the author

There are multiple areas where the data are either poorly described or poorly interpreted; please address these in a revision:
Lines 153-155: Figure S1 shows that CTB1 (not CTB4) is majority S/G2M (unless the figure is incorrectly labeled?).
Lines 318-321: I'm not sure how accurate this statement is, as the starting material for BCAM-low organoids is 100% BCAM-low; so it appears that BCAM-low CTB can become BCAM-high, at a much higher rate than BCAM-high cells become BCAM-low?
Lines 335-338: Not clear what this "EGFR-lo" trophoblast is? What does "greater plasticity" mean? The authors should be able to distinguish between "greater plasticity" and "differentiation" with relatively simple assays: are the BCAM-low organoids differentiating prematurely and if so, into what lineage?
Lines 392-405: This is very confusing and needs to be re-worded; it appears that the authors are trying to make their data appear to be in agreement with this previous publication. However, there is no evidence presented in this manuscript about the ITGA2+ population being bipotent, so the authors should more clearly compare and contrast their data with those presented in Lee et al. (2018).
Additional minor comments:
Line 147: this should read "cCTB2 expresses, at comparable levels to CTB1…."
Line 152: Not sure if the majority of cCTB2 are in G1 (per Figure S1): the proportion of cells in G1 is greater in cCTB2 than cCTB1, but an equal proportion of cCTB2 appear to be in G1 and S phases.
Lines 265-268: Not sure if I agree with the assessment of the variability in ELF5 expression; while in CT27 and CT29, it is more aligned with the CTB2 state, in CT30, it appears to be most enriched in the cCTB state.

Reviewer 2

Advance summary and potential significance to field

This is a single cell RNA sequencing study of human first trimester placental villi as well as human trophoblast stem cell-derived organoids. In addition to a detailed scRNA-seq analysis identifying various cell states/populations in both entities, the authors try to correlate these datasets and conclude that largely, hTSCs recapitulate the in vivo differentiation trajectories. They also zoom in on a particular factor that was identified in both datasets, BCAM, as a stem/progenitor cell-enriched cell adhesion molecule that endows cell with a higher proliferation rate, albeit not with higher regenerative potential in single cell cloning experiments.

Comments for the author

The scRNA-seq data lack substantial technical detail. By default, the placental villus digest with collagenase/hyaluronidase will not capture syncytiotrophoblast cells, and even the extent to which EVT cells can passage through the Chromium device, which has a microfluidic diameter of 30µm, is questionable. The authors should compare their scRNA-seq dataset to bulk data of first trimester placental villi and scrutinize their data for key cell types they may be missing out on.
I could not see a reference to vital scRNA-seq data, i.e. the table with cell number should also include how many genes per cell they find expressed, and the total number of reads per cell obtained. How many dropouts were observed? These details are critical to evaluate the quality of the scRNA-seq dataset.

The authors start with over 50,000 cells, but report on only ~7,800. Presumably all the other cell types are villous msenchyme, described in Supp. Fig. 1. These are very unwieldy and hard to read UMAPs. The authors do not make any further reference of all these other cells. This should be done in the main text, beyond the routine pipeline graphs shown in Supp. Fig. 1.
The authors reference previous studies of exactly this type, indeed this is at least the fourth single cell analysis of first trimester placentas. While a blanket statement is made that overall their data are very consistent with the previous reports, this should be analyzed in far more detail. These single cell data become un-interpretable if every study identifies different numbers and identities of cell types and states. While some corroboration of findings by repeat studies is desirable, it needs to be made obvious how precisely these data overlap. UMAPs with and without inlcusion of the other data (and not just Vento-Tormo et al) is required.

The BCAM focus is a bit tenuous, as the gene is widely expressed in CTB and at the base of CCCs. Moreover, albeit being a cell adhesion molecule, the staining pattern is distinctly nuclear (e.g., Fig. 4B, bottom right). The quality and/or specificty of the antibody should be controlled for (by staining of KD cells, for example) and the nuclear localization explained.
The KD studies are poorly controlled. There is no indication that cells were selected after transfection. Hence, the majority of the cell pool will consist of untransfected cells. These experiments should ideally be repeated by sorting transfected cells (e.g. by co-labeling with a GFP marker) and re-plating of the KD cells only.
It is unclear why PEG10 is repeatedly named as a stem cell gene. It should be referred to as paternally expressed - not paternally imprinted - gene. In mice, Peg10 is syncytiotrophoblast-expressed. In human placentas, it does seem to have some reported function in VCT cells. The authors should identify why they reference it as stem cell marker.
What are "forward" and "reverse" transfections? To the best of my understanding of the Materials & Methods description, these were simply two repeat transfections of siRNA. Please clarify.
Finally, what does the sentence mean "In summary, these cells were re-suspended in ice-cold growth factor-reduced Matrigel (GFR-M, Corning) to a final concentration of 100%"? Final concentration of what? The Matrigel? I.e. it was left undiluted? If that is the interpretation, please just state "cells were resuspended in undiluted Matrigel".

Reviewer 3

Advance summary and potential significance to field

In this article, Shannon et al. performed single cell RNA sequencing (scRNA-seq) on the first trimester placentas across 5-12 weeks gestational age with a lineage trajectory analysis identifying the trophoblast progenitor origin and BCAM as a marker of trophoblast cells in this progenitor state.
This was followed by a similar analysis on in vitro generated trophoblast organoids, leading to observations in regards to similarities and the extent on which these organoids can model some of the cellular aspects of the early placenta. Furthermore, they use trophoblast organoid assays to validate several of their findings in regards to BCAM+ cells. In general this is a very nice paper, well timed which uses a good mix of in vivo and in vitro data, which no doubt will be of great interest for the field.

Comments for the author

The following are some few points that I think will help corroborate some of the findings and conclusions of manuscript as well as making some of the messages clearer.
Key points that in my opinion required attention before publication:
1. Could the authors provide some few examples of the expression of genes mentioned in page 6 and 7 in their UMAPs presented in figure 1 (Similar to Fig 2B). I believe it will help the readers.
2. I think that it is important that other pseudotime analyses are performed to complement the RNA velocity provided in figure 2 as different reductions and algorithms may suggest different or similar trajectories.
3. It is interesting to see that the scRNA-seq analysis of the hTSC organoid differentiation cultures resolve into similar cell type clusters as the placental trophoblast populations. Although integration of different datasets can be problematic and brings their own problems, if done appropriately can bring another level of understanding. Thus, I think that it will be very useful if an integrative analysis were to be provided on the scRNA-seq dataset of the hTSC organoid differentiation culture and the placental trophoblast populations.
4. Using the trophoblast organoid cultures it is shown that BCAMlo-sorted cells have a reduction in regenerative properties. On the other hand, does this also mean that cells in this population are more likely to be differentiating to the CTB and EVT lineages? Since the BCAMhigh and BCAMlow-sorted cells can both form organoids, can the authors also examine the differentiation potential of the organoids generated from these two sorted cell populations?
Other suggestion that I think will improve the manuscript or make it clearer:
5. It was concluded that BCAM enriches trophoblast cells with greater regenerative potential based on the different assays performed using the in vitro trophoblast organoid models. Have the authors examined if BCAM can be used to directly isolate cells from the placental tissue for establishment of hTSC lines or trophoblast organoid?
Of course this may no be possible at this stage or depending on the samples that the authors have access to, however I believe that this will be a very important experiment to have if possible.
6. Individual samples were grouped into early and late gestational age to examine differences in cellular heterogeneity (Fig. S1G).
I think that it will be highly valuable if the maternal/placental contribution could be deconvoluted (maybe using SNPs), rather than solely relying on theoretical origin. Having that relationship resolved may provide insights into differences among the same groups.
7. Single cell data suffers from dropout, so if a “signature” (made of many genes) could be calculated for the CTB1/2/3 populations, then this could be plotted into the different reduction plots (UMAPs).
8. At the moment cluster assignment and their relationship among samples in figure 3D is not clear. Maybe the authors can try to integrate the datasets or show a correlation between clusters?
9. It will be interesting to examine and discuss if there is any difference between the transcriptomic profile of the BCAM+ trophoblast populations across different gestational ages.
10. If possible it will be important to compare the differentiation potential of the siBCAM organoids and siCTRL organoids? Although in line 337-338 is suggested that BCAM reduction may result in accelerated differentiation, more assays could be performed to validate this in the trophoblast organoid system.
11. For Fig. 5F, a co-immunostaining with a SCT marker (hCG as in Fig. 4B)
should be included to indicate the presence of ST, otherwise don’t you think that the SCT indicator should be removed.
12. It would be more helpful for the readers to follow along if the authors could cite each of the sub-panels of the supplementary figures in a more explicit way (Fig. S1A, Fig. S1B, Fig. S1C etc) rather than just citing the whole supplementary figure in each statement in the manuscript.
13. In line 220-223, it would be less confusing for the readers if the authors could rephrase ‘three hTSC lines’ to ‘three lines of hTSC organoids’
since that it was the organoids that were used for sequencing rather than the monolayer hTSC lines.
14. The in-text figure citation in line 237 does not correlate to a figure panel showing expression of CGB and CSH1.

First revision

Author response to reviewers' comments

Response to Reviewers

We thank all three reviewers for providing us with insightful comments and criticism of our manuscript. We have performed new analyses of our single cell datasets, performed new experiments, re-organized the presentation of our data, and made clarifications within the manuscript that directly address all the specific concerns raised.
Overall, we feel that the changes we have made have considerably strengthened the quality of our work.

We would also like to share with the reviewers that we did encounter difficulties in recruiting patients/tissues needed for some of the comments/suggestions provided by the reviewers. The hospital clinic that coordinates patient/sample collection for our laboratory was not participating in research activities from March 2020 to September 2021 due to COVID-19 curtailment and REB amendments were required for patient consenting to start-up following COVID return to research. We also encountered many COVID-related backorders on key reagents like Matrigel needed for organoid culture. For these reasons we necessitated requesting/being granted an additional 1.5 months to complete the revisions.

Below are detailed responses to the Reviewer's concerns:

Reviewer 1:
Reviewer 1 Advance Summary and Potential Significance to Field: This manuscript describes single cell RNA sequencing (scRNAseq) of early and late first trimester human placenta, and describing different cytotrophoblast (CTB) progenitor states, leading to identification of BCAM-high cells as potential progenitor CTB, whose enrichment leads to enhanced growth of trophoblast organoids.

While there have been a lot of publications in this area, this manuscript advances the field by describing different CTB progenitor states, leading to identification as well as further characterization of one such CTB state. The authors are careful not to over-interpret their data; nevertheless, the manuscript suffers from a lack of clarity in several areas, appearing at times ready to "agree" with previously-published studies, without sufficient data. Specific suggestions for revision are listed below.

We thank Reviewer 1 for the thorough comments on our work and the acknowledgment of how our study advances the field by presenting a detailed characterization of distinct CTB states and in particular the progenitor state aligning with the predicted origin. The changes made have strengthened the message and quality of our findings. Below is the point-by-point response to each comment/concern.

Comments for the author:

1. Lines 153-155: Figure S1 shows that CTB1 (not CTB4) is majority S/G2M (unless the figure is incorrectly labeled?).

We thank the reviewer for identifying this mistake as CTB4 is the state with the majority of cells in S/G2M. Data shown in the previous Fig. S1 or cycle phases were incorrect and have now been updated accordingly in Fig. S1 and within the results Lines 156-161.

2. Lines 318-321: I'm not sure how accurate this statement is, as the starting material for BCAM-low organoids is 100% BCAM-low; so it appears that BCAM-low CTB can become BCAM-high, at a much higher rate than BCAM-high cells become BCAM-low?

The reviewer is correct in their interpretation that BCAM-low cells acquire high levels of BCAM expression at a potentially faster rate than BCAM-high cells lose BCAM expression. This interpretation could mean that BCAM- low cells acquire a greater state of regenerative ability following sorting and culture. However, the expansion of the EGFR-low population in BCAM-low

CTBs can also be interpreted as BCAM-low cells spontaneously differentiate (at a greater rate than BCAM-high cells?), as EGFR is a CTB marker and is down-regulated in SCT and EVT in vivo. To capture both interpretations better, we have added new text in Lines 372-375 that describe this alternative viewpoint.

3. Lines 335-338: Not clear what this "EGFR-lo" trophoblast is? What does "greater plasticity" mean? The authors should be able to distinguish between "greater plasticity" and "differentiation" with relatively simple assays: are the BCAM-low organoids differentiating prematurely and if so, into what lineage?

We are unsure of what the EGFR-low population is, though we initially interpreted this population as a subset of cells losing progenitor/stemness and developing into more differentiated states. To address whether BCAM high/low populations have different differentiation potentials, we performed new experiments where CTBs were sorted on level of BCAM expression, established as regenerative organoids for 5 days, and then subjected to 10 day EVT differentiation where RNA was collected and organoids were imaged at days 0, 3, 7, and 10. Further, endpoint day 10 organoids were PFA fixed and used for IF imaging of CTB and EVT markers.

While brightfield imaging showed that both BCAM-high and -low derived organoids develop EVT-like growth/extensions by day 7 of EVT culture, qualitatively we did not see any notable differences between BCAM groups. IF imaging of the CTB marker EGFR and the immature EVT marker a5 integrin (ITGA5) likewise showed that CTB composition and early EVT development are not different. Though consistent and in line with our previous data, BCAM-high derived organoids generated larger organoids than BCAM-low organoids.

We also aimed to examine mRNA levels of key progenitor (TP63, TEAD4), SCT (CSH1, CGA), immature (ITGA5, NOTCH1), and mature EVT (HLA-G, NOTCH2) genes at the four timepoints. However, following BCAM-low/-high sorting and organoid culture, we did not collect sufficient RNA quantity to assay all these genes, and instead opted to only examine levels of NOTCH1 and HLA-G. We recognize that this does not address if BCAM-low/high CTBs are committed more to forming the SCT lineage, and as such is a limitation of this work. Revision time-line constraints imposed by unforeseen reagent backorders needed for organoid culture (NRG1, Matrigel) did play a major role in limiting the scope of these experiments.

To summarize, these new experiments suggest that the differentiation potential of BCAM-high and BCAM-low organoids into the EVT lineage is not different. New data is presented in Fig. S5 and described as new text within the results on Lines 375-384.

4. Lines 392-405: This is very confusing and needs to be re-worded; it appears that the authors are trying to make their data appear to be in agreement with this previous publication. However, there is no evidence presented in this manuscript about the ITGA2+ population being bipotent, so the authors should more clearly compare and contrast their data with those presented in Lee et al. (2018).

We thank the reviewer for this comment. We agree with them that there is no concrete evidence that the ITGA2 population has a level of bipotency. New computational data performed for this revision involved integrating our chorionic villi and Vento-Tormo single cell datasets with our hTSC-derived organoid single cell data.
Combining these data and performing scVelo pseudotime modeling identified an additional origin/progenitor state aligning with ITGA2 expressing cells predominately derived from organoids. While this data does not address whether this population has a level of bipotency to differentiate into SCT or EVT, pseudotime modelling does show that ITGA2 cells contribute to column and EVT subsets, suggesting that these cells are a column progenitor.

We have revised the paragraph in the discussion to more accurately reflect what these cells may represent, including additional insights generated from new single cell analyses; Lines 464-466.

Additional minor comments:

Line 147: this should read "cCTB2 expresses, at comparable levels to CTB1…."

We thank the reviewer for noticing this error and have corrected this mistake.

Line 152: Not sure if the majority of cCTB2 are in G1 (per Figure S1): the proportion of cells in G1 is greater in cCTB2 than cCTB1, but an equal proportion of cCTB2 appear to be in G1 and S phases.

We apologize to the reviewer that the data presented in this figure was not correct due to a mistake generated in Seurat in R. We have replaced the old figure with a revised figure that now shows the correct single cell informed cell cycle state data. Of note, cCTB1 actually has a greater proportion of cells in S and G2M than cCTB2 which is interesting and perhaps unexpected considering the cCTB1 gene signature reflects a more mature column/EVT state.

Lines 265-268: Not sure if I agree with the assessment of the variability in ELF5 expression; while in CT27 and CT29, it is more aligned with the CTB2 state, in CT30, it appears to be most enriched in the cCTB state.

We recognize the limited interpretability of the data presented in the heatmaps of Fig. 3. Previously, we had initially subset cells of each hTSC cell line (CT27, CT29, and CT30), re-clustered this data, and then averaged gene expression within each cluster of each cell line. As a result of this, intra-cell line comparisons of average gene expression are possible, however any comparison across cell lines was not possible and this data needed to be modified to improve its interpretability.

To fix this error and allow for direct comparison across each hTSC cell line in our data, we instead began with the combined hTSC data (including cells from CT27, CT29, and CT30-derived organoids) and introduced a new variable that included the combined cluster identities as shown in Fig. 3D, but with the addition of the cell line from which all cells were derived (for example, cells in CTB1 of the combined hTSC data were given a label of either: CTB1_CT27, CTB1_CT29, or CTB1_CT30 - this maintained cluster identity and cell line origin within the combined dataset, allowing now for direct comparison across the three cell lines used). With this, we then averaged gene expression to the identities included in this new variable, providing an average gene expression level for every gene in each cluster as well as cell line. Next, to improve interpretation of the data, we decided to combine the previous three heatmaps (one heatmap for each cell line) into one heatmap that allowed for easier visualization of average log gene expression differences across the various states derived from either CT27, CT29, or CT30.

From this, we now see that ELF5 expression is predominantly found in CTB1 cells in each cell line condition, with some additional expression in CTB2 cells derived from CT27, and reduced expression seen in the cCTB state. As well, we see greatest ELF5 expression in organoids derived from the CT27 cell line.

Due to these modifications, we have updated our description and subsequent interpretation of the data presented in the heatmaps found in Fig. 3. Revised text outlining these updates are found in Lines 289-293 in the results.

Reviewer 2:

Reviewer 2 Advance Summary and Potential Significance to Field:

This is a single cell RNA sequencing study of human first trimester placental villi as well as human trophoblast stem cell- derived organoids. In addition to a detailed scRNA-seq analysis identifying various cell states/populations in both entities, the authors try to correlate these datasets and conclude that largely, hTSCs recapitulate the in vivo differentiation trajectories.
They also zoom in on a particular factor that was identified in both datasets, BCAM, as a stem/progenitor cell-enriched cell adhesion molecule that endows cell with a higher proliferation rate, albeit not with higher regenerative potential in single cell cloning experiments.

We thank the reviewer for their constructive and thorough comments.

Comments for the author:

1. The scRNA-seq data lack substantial technical detail. By default, the placental villus digest

with collagenase/hyaluronidase will not capture syncytiotrophoblast cells, and even the extent to which EVT cells can passage through the Chromium device, which has a microfluidic diameter of 30µm, is questionable. The authors should compare their scRNA-seq dataset to bulk data of first trimester placental villi and scrutinize their data for key cell types they may be missing out on.

The reviewer raises some important points about the limitations of the 10X Chromium microfluidics platform for generating single cell cDNA libraries. It is true that placental villous (and possibly organoid) syncytiotrophoblast will not be captured on this platform due to the large and atypical cellular/nuclear composition of this multinucleated structure; enzymatic digestion will simply result in the generation of free syncytial nuclei. The procedure in which enzymes are used to generate a single cell suspension from chorionic villi was a very challenging process due in part to syncytial nuclear release and resulting acellular DNA that gummed up our cell suspensions leading to excessive cell loss. Through trials of different DNaseI treatments we were able to greatly reduce the impact of syncytial DNA.

Despite not being able to capture true syncytia, we are able to capture a CTB population that are likely in the process of fusing; we term these cells SCT precursors (SCTp). These cells express many genes associated with the syncytiotrophoblast, and while not true syncytiotrophoblast, we can still model CTB differentiation into a downstream SCT precursor. Future single cell omics approaches will need to either utilize spatial single cell platforms or single nuclei seq platforms. Our lab recognizes this and we are currently investing time and effort to optimize single nuclei RNA seq workflow for future work.
We respectfully point out to the reviewer that the 10X Chromium microfluidics bore is 50 µm, not 30 µm (though it is true 10X Genomics has only tested in house suspended cells up to 30 µm in diameter). While this size may still limit the capture of truly large cells like polyploid EVT, it is important to keep in mind that cells in suspension are much smaller than when they are adherent to a substratum. So, while an EVT imaged by IHC within the decidual bed may be 100 µm in diameter, in suspension this cell may be as small as 50 µm. Our data suggests that we are capturing most trophoblasts within the placenta, and the Vento-Tormo dataset indicates they have captured a significant portion of decidual EVT as well. We do however agree with the reviewer that even with a 50 µm bore, we are likely missing the capture of some cells, and this is a limitation of the system.

We are unsure of what the reviewer is asking of us regarding the comparing of our single cell data to bulk RNA- seq data. It is difficult to draw comparisons from averaged RNA profiles (bulk seq) to single cell RNA profiles. The known gene markers used to characterize/define each cell state show that cell capture is efficient and that the data generated agrees with what is known about genes expressed in general trophoblast subtypes.

2. I could not see a reference to vital scRNA-seq data, i.e. the table with cell number should also include how many genes per cell they find expressed, and the total number of reads per cell obtained. How many dropouts were observed? These details are critical to evaluate the quality of the scRNA-seq dataset.

As suggested, we have added to the technical detail of overall scRNA-seq quality control described in Table 1. We have updated the table to include the fraction of reads in cells, the mean number of reads per cell, the median UMI counts per cell, the median genes captured per cell, and the total number of genes detected in each sample that we generated in-house (GSE174481). This information for E-MTAB-6701 was not previously published by Vento-Tormo et al and is not publicly available for their dataset. Unfortunately, due to a server error this information was deleted for our re-processing of E-MTAB-6701. We do not believe this to be a concern as the publication of this data and subsequent re-analysis of this dataset in multiple publications provides informal indication of its quality. As well, we have included this data for our new datasets included within this paper and have applied a strict quality control pipeline, detailed in the Single-Cell RNA Seq data analysis subsection of the methods (Lines 625-631), to ensure only quality cells were used in this study.

Regarding read dropouts, this is not something that is quantifiable in each sample. Common bioinformatic packages used to assess dropouts in single-cell RNA seq experiments (M3Drop, RESCUE, scDoc, SDImpute, etc.) address the dropout problem by imputing the missing gene

expression data. This practice can result in false positives being introduced, and therefore avoided to maintain data reproducibility. We do recognize that we were not clear in how we handled read dropouts in our dataset and have updated the Single-Cell RNA Seq data analysis methods subsection to clarify this. In short, we used the feature selection function in Seurat to focus on the top 2000 genes that showed cell-cell variability (Lines 595-597). This ensures that biological signal in our data is prioritized and that we are minimizing the possibility of measuring an artefact resulting from read dropouts.

3. The authors start with over 50,000 cells, but report on only ~7,800. Presumably all the other cell types are villous msenchyme, described in Supp. Fig. 1. These are very unwieldy and hard to read UMAPs. The authors do not make any further reference of all these other cells. This should be done in the main text, beyond the routine pipeline graphs shown in Supp. Fig. 1.

The reviewer is correct that approximately 7800 trophoblasts were used for downstream analyses following quality control computational pipelines. As can be seen within the UMAP showing maternal/decidual cells (Fig. 1B; Fig. S1F). The majority of the 50,790 cells came from the Vento-Tormo decidual samples, and the reviewer is correct that the majority of the captured cells are mesenchymal/stromal and immune cells of the decidua.
Placental mesenchymal and immune cells also contributed to a significant portion of captured cells; for our analyses we were interested only in the trophoblast component and therefore did not describe/analyze non- trophoblasts beyond their general identification inferred from gene expression.

As suggested, we have moved the single cell clustering data of the entire fetal/maternal landscape into Fig. 1.

4. The authors reference previous studies of exactly this type, indeed this is at least the fourth single cell analysis of first trimester placentas. While a blanket statement is made that overall their data are very consistent with the previous reports, this should be analyzed in far more detail. These single cell data become un-interpretable if every study identifies different numbers and identities of cell types and states. While some corroboration of findings by repeat studies is desirable, it needs to be made obvious how precisely these data overlap. UMAPs with and without inlcusion of the other data (and not just Vento-Tormo et al) is required.

The reviewer raises concern about dataset consistency and how well our data fits into the first trimester single cell placental landscapes previously described by Vento-Tormo et al, Suryawanshi et al, and Liu et al.

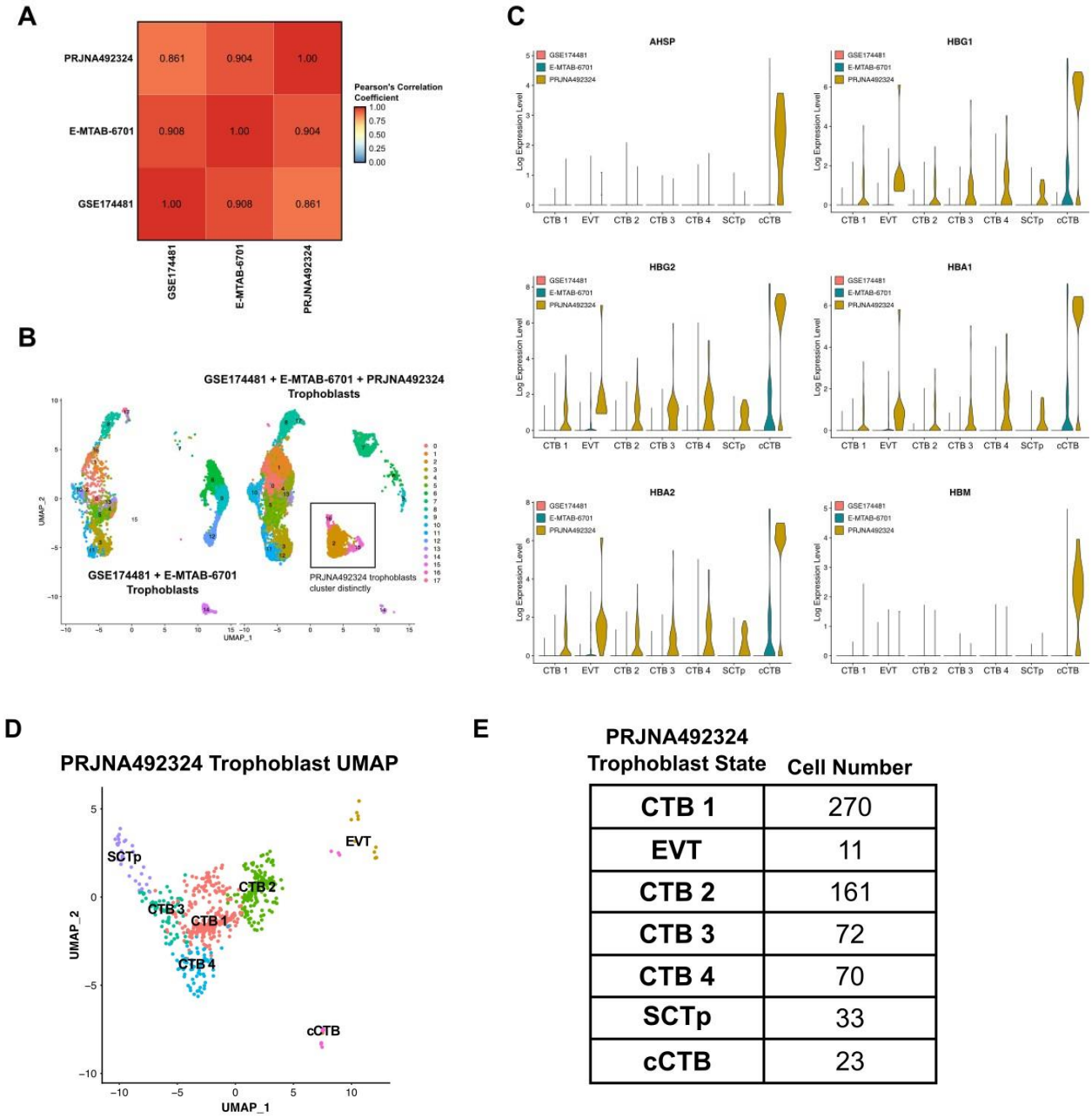
We recognize that high dimensional single cell gene expression data is inherently subject to bias, especially when factoring in varied sample preparation protocols, sequencing technology choice, data pre-processing, and data quality control filters that are not standardized across studies. Data pre-processing can greatly influence which cells are maintained in downstream analyses and, if not carefully controlled, can introduce bias. This is one likely source of the inconsistencies described by the reviewer when comparing multiple single cell atlases that are independently curated. In addition to this, cluster resolution is an important parameter used to modify the number of clusters generated in single cell data. Importantly, for higher resolution parameters the gene expression differences needed to call different states will be smaller. Lower resolution parameters therefore indicate large transcriptomic differences are needed define different clusters. Because this parameter is not controlled across single cell studies, the number of clusters found in different single cell atlases (or when these atlases are re-analyzed) is likely not consistent, and should be carefully interpreted. Within our study, we have used the clustree package to assess cluster stability at various resolutions and use this to select robust and ideal resolution parameters for our research questions.

This said, we did aim to integrate our data with all publicly available first trimester single cell placenta repositories to improve the power and robustness of our findings, and as well synergize the current first trimester single cell placenta transcriptomic landscape. Because of our concerns noted in the previous paragraph, we cautiously filtered through the available data to avoid introducing technical variability and bias in our study. As a result, we sought to only include data generated on the 10X genomics Chromium platform and only include data from patient tissues

that were digested to a single cell suspension in an unbiased manner (no EPCAM$^+$ or HLA-G$^+$ cell enrichment). In addition to this, we applied a strict data preprocessing pipeline, described in our methods, to ensure only high-quality cells were included in our analyses. Further, we ran a pseudo-bulk comparison of the whole gene expression profile of each potential dataset to ensure that only repositories with >90% gene expression similarity to our data (as defined using a Pearson's Correlation Coefficient) were integrated into our study. Based on these inclusion criteria, only a portion of the Vento-Tormo dataset was selected for integration with our data.

To ensure that our inclusion criteria was not too strict, we wanted to additionally observe how the 10X samples included in the Suryawanshi paper (n=2) integrated with the Vento-Tormo data and our data that was used in this paper. We did not do this with the Liu et al. data because their data was SmartSeq2, not 10X data, and only provided an n=1. Importantly, some concerns were noted upon integration with the Suryawanshi data:

- *First, **in panel A and B below**, we observed less than 90% correlation between our dataset and the Suryawanshi data. As well, clustering of cells by data source (GSE174481, E-MTAB-6701, and PRJNA492324) and not by gene expression profile was observed with our in-house data (GSE174481) clustering cohesively with the Vento-Tormo et al (E-MTAB-6701) data and the Suryawanshi et al data clustering distinctly.*
- *Second, **in panel C below**, looking at the PRJNA492324 data relative to our GSE174481 data and the E- MTAB-6701 data, we observed increased expression of hemoglobin-related genes such as AHSP, HBG1, HBG2, HBA1, HBA2, and HBM. We interpret this as possible contamination in the dataset that could introduce unwanted data artefacts if integrated.*
- *Third, **in panels D and E below**, further observation of the Suryawanshi (PRJNA492324) data on its own demonstrated reduced trophoblast capture (only 640 trophoblasts were identified) and poor representation of the first trimester trophoblast landscape with little SCTp (33 cells), column CTB (23 cells), and EVT (11 cells) captured in the data. While not a direct concern for data integrity, when taken together with the previous points described here we made the decision to focus our analyses on our data as well as the 10X genomics data available from Vento-Tormo et al. Together, this provided us with >50,000 high quality fetal/maternal interface cells, generated in two independent facilities using consistent experimental approaches/design and with >90% pseudo-bulk transcriptome similarity.*



E

PRJNA492324

Trophoblast State	Cell Number
CTB 1	270
EVT	11
CTB 2	161
CTB 3	72
CTB 4	70
SCTp	33
cCTB	23

***Figure legend**: Data Comparison between GSE174481, E-MTAB-6701, and PRJNA492324. **A)** Pseudo-bulk comparison of each data set. **B)** Uniform Manifold Approximation Plot (UMAP) of the integrated data used in the paper integrated with and without PRJNA492324. **C)** Violin plots of the expression of hemoglobin-related genes in the trophoblast states identified in each dataset. **D)** UMAP of the trophoblast data derived from PRJNA492324. E) The number of cells found in each cluster from PRJNA492324.*

5. The BCAM focus is a bit tenuous, as the gene is widely expressed in CTB and at the base of CCCs. Moreover, albeit being a cell adhesion molecule, the staining pattern is distinctly nuclear (e.g., Fig. 4B, bottom right). The quality and/or specificty of the antibody should be controlled for (by staining of KD cells, for example) and the nuclear localization explained.

The reviewer is correct in that BCAM is expressed in many states assessed by single cell analysis. Further, IF microscopy shows BCAM to be expressed strongly in most CTB, with levels gradually decreasing within cells of proximal and distal columns; SCT do not express detectable levels of BCAM as assessed by IF. However, when examining mRNA levels, BCAM is highest in the two CTB states (CTB2, CTB3) aligning with the predicted origin determined by scVelo pseudotime analysis.

This data shows that BCAM levels gradually decrease as CTB2/3 differentiate into EVT and SCT, and therefore suggests that BCAM may play a role in progenitor CTB biology. We have updated Fig. 2 to better illustrate BCAM mRNA levels in the 8 scRNA-seq informed states (Fig. 2F).

We agree with the reviewer that the image in the previous Fig.4B showed nuclear staining of BCAM in some cells. This staining pattern is not what we typically see by IF, and perhaps was the result of non-specific staining of dead cells in that specific organoid. We have replaced the previous BCAM IF image with an image that better reflects typical BCAM localization in trophoblast organoids (Fig. 5B). We also remind the reviewer that we isolate CTB by BCAM surface expression by FACS; this does not rule out that nuclear BCAM staining exists, but the ability to detect/sort on surface BCAM expression by flow does provide additional evidence that BCAM expression in our organoid system is present on the extracellular portion of the cell. Silencing of BCAM, shown in data presented in Fig. 6A, shows that levels of BCAM are reduced in siRNA targeting conditions, suggesting that the BCAM antibody does detect BCAM protein with a level of specificity.

6. The KD studies are poorly controlled. There is no indication that cells were selected after transfection. Hence, the majority of the cell pool will consist of untransfected cells. These experiments should ideally be repeated by sorting transfected cells (e.g. by co-labeling with a GFP marker) and re-plating of the KD cells only.

We respectfully point out to the reviewer that we used an siRNA approach to transiently knock-down BCAM in trophoblast organoids. This approach uses siRNA that are not tagged or ectopically expressed by an expression system. We agree with the reviewer that ideally a stable knock-down (tagged Lenti system) or knock-out (CRISPR) approach would be preferable, but we resorted to using the siRNA approach having encountered technical challenges optimizing Lenti viral transduction of vectors expressing BCAM shRNAs in hTSCs.

7. It is unclear why PEG10 is repeatedly named as a stem cell gene. It should be referred to as paternally expressed - not paternally imprinted - gene. In mice, Peg10 is syncytiotrophoblast-expressed. In human placentas, it does seem to have some reported function in VCT cells. The authors should identify why they reference it as stem cell marker.

The reviewer is correct, and we apologize for this mistake. We have now modified text in Line 386 and Lines 450-452 to indicate that PEG10 is a paternally expressed gene and not a transcription factor associated with stemness (reference to PEG10 was removed in the discussion).

8. What are "forward" and "reverse" transfections? To the best of my understanding of the Materials & Methods description, these were simply two repeat transfections of siRNA. Please clarify.

We apologize to the reviewer for not making the distinction between forward and reverse transfections clearer. Forward transfection refers to the classical lipid-based transfection protocol where cells are initially seeded (often overnight) prior to transfection. In forward transfection protocols, siRNA is combined with the lipid- delivery reagent and added to media the cells are cultured in.

This is in contrast to reverse transfections where the siRNA and lipid transfection reagent are combined directly with a single cell suspension immediately prior to seeding. We used these two techniques as they allowed us to perform sequential transfections in our organoid system to achieve BCAM knockdown over a 7-day time course.

9. Finally, what does the sentence mean "In summary, these cells were re-suspended in ice-cold growth factor-reduced Matrigel (GFR-M, Corning) to a final concentration of 100%"? Final concentration of what? The Matrigel? I.e. it was left undiluted? If that is the interpretation, please just state "cells were resuspended in undiluted Matrigel".

The reviewer is correct that our original wording was confusing. We have revised this to simply indicate that hTSCs were resuspended in undiluted Matrigel (Line 531).

Reviewer 3 Advance Summary and Potential Significance to Field...

In this article, Shannon et al. performed single cell RNA sequencing (scRNA-seq) on the first trimester placentas across 5-12 weeks gestational age, with a lineage trajectory analysis identifying the trophoblast progenitor origin and BCAM as a marker of trophoblast cells in this progenitor state. This was followed by a similar analysis on in vitro generated trophoblast organoids, leading to observations in regards to similarities and the extent on which these organoids can model some of the cellular aspects of the early placenta. Furthermore, they use trophoblast organoid assays to validate several of their findings in regards to BCAM+ cells. In general this is a very nice paper, well timed which uses a good mix of in vivo and in vitro data, which no doubt will be of great interest for the field.

We thank the reviewer for the encouraging remarks and the thorough and thoughtful comments. We have addressed every concern to the best of our abilities and feel the revised manuscript is stronger because of these changes.

Reviewer 3 Comments for the Author...

The following are some few points that I think will help corroborate some of the findings and conclusions of manuscript as well as making some of the messages clearer.

Key points that in my opinion required attention before publication:

1. Could the authors provide some few examples of the expression of genes mentioned in page 6 and 7 in their UMAPs presented in figure 1 (Similar to Fig 2B). I believe it will help the readers.

We agree that the inclusion of some UMAP plots depicting qualitative gene expression levels across our identified trophoblast states will improve the interpretation and readability of Fig. 1 and thank the reviewer for pointing this out. We have remodeled Fig. 1 to include 6 feature plots demonstrating gene expression of EGFR, TP63, MKI67, ERVRD-1, ITGA2, and HLA-G in our UMAP projection. These new data are presented in Fig. 1E and described in Lines 139-143 of the results section.

2. I think that it is important that other pseudotime analyses are performed to complement the RNA velocity provided in figure 2 as different reductions and algorithms may suggest different or similar trajectories.

We agree that the scVelo trajectory models described in our paper benefits from comparison against other pseudotime and lineage models derived from additional trajectory analyses and algorithms. We previously compared our scVelo results in Fig. 2 with pseudotime trajectories that were derived from both the Monocle2 and Monocle3 platform algorithms. We found here consistent ordering of trophoblasts along both the villous and extravillous differentiation paths, with consistent upregulation of basal cell adhesion molecule (BCAM) within an identified origin population residing in the CTB2 state.

To improve clarity of this, we have moved our Monocle3 pseudotime ordering from Fig. S3 to Fig. 2. To further demonstrate consistent observation of BCAM upregulation within our origin, we added a fourth algorithm for trajectory inference, SlingShot. The results of our SlingShot trajectory modelling are consistent with what is previously observed, with some deviation due to limitations in its underlying algorithm being unable to regress out the cell proliferation variable (which can bias trajectory inference). We present our Slingshot ordering in a new Fig. S3 with the Monocle2 pseudotime ordering inserted here as well. Together, we consistently observe BCAM upregulation within a putative trophoblast origin population within the CTB2 state. Together, description of this new data is written in Lines 188-197 of Results.

3. It is interesting to see that the scRNA-seq analysis of the hTSC organoid differentiation cultures resolve into similar cell type clusters as the placental trophoblast populations. Although integration of different datasets can be problematic and brings their own problems, if done appropriately can bring another level of understanding. Thus, I think that it will be very useful if an integrative analysis were to be provided on the scRNA-seq dataset of the hTSC organoid differentiation culture

and the placental trophoblast populations.

We thank the reviewer for this insightful suggestion and have included an integrative analysis on our scRNA-seq data derived from both in vivo placental trophoblasts and in vitro hTSC-derived organoids. This data is presented in a new Fig. 4 and clarifies the identification of an upstream CTB progenitor population with upregulated BCAM expression. Description of this new data is written in Lines 320-340 of the Results.

In vivo and in vitro data integration further resolved our trophoblast clustering into 9 states. Our in vivo placental trophoblast and in vitro hTSC organoid datasets proved to be similar with some interesting differences noted within the paper. Ultimately, consistent upregulation of BCAM within an origin population found in both the in vivo and in vitro data supports our use of hTSC-derived organoids as a model to explore CTB BCAM expression further.

4. Using the trophoblast organoid cultures it is shown that BCAMlo-sorted cells have a reduction in regenerative properties. On the other hand, does this also mean that cells in this population are more likely to be differentiating to the CTB and EVT lineages? Since the BCAMhigh and BCAMlow-sorted cells can both form organoids, can the authors also examine the differentiation potential of the organoids generated from these two sorted cell populations?

We thank the reviewer for this suggestion. This comment is similar to the comment provided by Reviewer 1 above (comment 3), and we provide a detailed response to how we performed new experiments to test if BCAM- lo or BCAM-hi organoids have different differentiation potentials.

Briefly, CTBs from CT29 cells were sorted into BCAM-lo and BCAM-hi and established in regenerative organoid media for 5 days. Following this, organoid media was switched to EVT differentiation media over a 10-day time-course. We focused exclusively on the EVT lineage potential and did not examine SCT potential as we encountered many obstacles in culturing organoids this past summer/fall due to significant reagent backorders. Matrigel, for example, was backordered for 8 weeks, and we were reluctant to use a substitute matrix (i.e hydrogel) so as to maintain culture/experimental consistency.

To summarize our findings, we did not find evidence that BCAM-lo CTBs differentiate more readily into EVT than BCAM-hi cells. However we did see, as before, that BCAM-hi CTB grow larger organoids (in both regenerative and EVT conditions). This new data is presented as new Fig. S5 and described in Lines 375-384.

Other suggestion that I think will improve the manuscript or make it clearer:

5. It was concluded that BCAM enriches trophoblast cells with greater regenerative potential based on the different assays performed using the in vitro trophoblast organoid models. Have the authors examined if BCAM can be used to directly isolate cells from the placental tissue for establishment of hTSC lines or trophoblast organoid? Of course this may no be possible at this stage or depending on the samples that the authors have access to, however I believe that this will be a very important experiment to have if possible.

This suggestion by the reviewer is very insightful and would greatly strengthen the overall understanding of BCAM in CTB regeneration capacity. We had initially planned on enriching for BCAM+ cells from chorionic villus preparations to examine if such preparations could enrich the isolation of regenerative CTB. Unfortunately, all first trimester tissue collections were halted in March 2020 due to COVID curtailments and were just resumed this past September. We agree that this will be important to do, and we are currently planning CTB isolation experiments to address this in future experiments. Given the length of time and placental sample numbers needed for this work, we feel this is not possible to perform for this revision given the time constraints and logistics we faced with re-instating our REB first trimester placenta study.

6. Individual samples were grouped into early and late gestational age to examine differences in cellular heterogeneity (Fig. S1G). I think that it will be highly valuable if the maternal/placental contribution could be deconvoluted (maybe using SNPs), rather than solely relying on theoretical origin. Having that relationship resolved may provide insights into differences among the same

groups.

We thank the reviewer for noting a need for applying SNP-based deconvolution methods to determine with greater precision the maternal/placental contributions to our data as presented in Fig. 1 and Fig. S1.

We agree that this would provide increased resolution for cell state origins, however, we were able to infer this information theoretically and with some technical insights from the decidual and placental preparations which were sequenced separately, providing a general understanding of the origin of a majority of the cells in each identified state. Further, our focus was specific to the origin and differentiation paths of progenitor placental trophoblasts, not on the origin of all cells within the fetal/maternal interface. As a result, we have not included a thorough SNP-based deconvolution approach within this paper, but note the importance of this for future work that may explore additional cell populations in the fetal/maternal interface and their possible fetal or maternal origin.

7. Single cell data suffers from dropout, so if a "signature" (made of many genes) could be calculated for the CTB1/2/3 populations, then this could be plotted into the different reduction plots (UMAPs).

We recognize the need to address the limitation of read dropouts in scRNAseq data and have clarified our approach to this in the Single-Cell RNA Seq data analysis subsection of the methods (Lines 592-594). We were not entirely certain if the signature mentioned here referred to the inference of a general CTB gene signature from our data or referred to a need for data imputation. We made attempts to generate and present a non- proliferative CTB gene signature derived from the top 10 genes upregulated within CTB1/2/3. However, due to word count limitations and concerns that this data was outside the scope of this paper, we decided to omit this data.

On the other hand, if this is suggesting the imputation of a gene signature for CTB 1/2/3 based on gene expression profiles in comparable cells, to account for read dropouts, we intentionally did not do this. This was because we wanted to avoid the potential introduction of false positives and technical bias within our data.
Further, we applied a feature selection algorithm in Seurat during data quality control (described in the methods) to ensure that only detected and variable genes were used in analysis (prioritizes biological signal over technical noise).

8. At the moment cluster assignment and their relationship among samples in figure 3D is not clear. Maybe the authors can try to integrate the datasets or show a correlation between clusters?

We agree that the relationship here is difficult to interpret and have modified the presentation of data in Fig. 3D to clarify the cluster assignment and relationship between the scRNAseq data derived from CT27, CT29, and CT30. All three cell lines used to generate scRNAseq libraries are shown integrated in Fig. 3D, and the individual distinctions between each cell line compartment of the integrated data are then displayed in Fig. 3F.

9. It will be interesting to examine and discuss if there is any difference between the transcriptomic profile of the BCAM+ trophoblast populations across different gestational ages.

We agree with the reviewer that this would be a helpful and interesting addition to the paper and have explored the transcriptomic profile of BCAM high and BCAM low CTBs (CTB states 1-3, omitting the CTB4 state which is predominated by a proliferative gene signature) within three gestational age groups across the first trimester (group 1: early, 6-7 weeks GA, group 2: mid, 9-10 weeks GA, and group 3: late, 11-12 weeks GA). This data is included as a new panel G in Fig. 2 and described in Lines 223-228.

10. If possible it will be important to compare the differentiation potential of the siBCAM organoids and siCTRL organoids? Although in line 337-338 is suggested that BCAM reduction may result in accelerated differentiation, more assays could be performed to validate this in the trophoblast organoid system.

We thank the reviewer for this important suggestion. We have performed BCAM knockdown experiments in hTSC to examine EVT differentiation potential (similar to what was performed using BCAM-low and -high sorted CTBs, though in 2D as opposed to 3D). Because we had initially used an siRNA approach to transiently knock-down BCAM in organoids, we opted to use the same approach in 2D hTSCs to maintain a level of consistency. We opted for 2D hTSC culture as opposed to 3D organoids because of the challenge of sequential siRNA treatments needed to maintain BCAM knockdown during an EVT differentiation timecourse. However, we found that iterative rounds of BCAM (and control) siRNA had a negative effect on overall cell viability. For this reason, we were only able to perform 3 independent rounds of BCAM knockdown in differentiating hTSCs up to day 3 of EVT differentiation. We did collect sufficient RNA from day 6 of differentiation from one experimental round. Unfortunately, we were unable to image or collect timepoints beyond day 6 in the other two rounds, as there was excessive cell death in the cultures.

Moving forward, we feel it will be essential to optimize more stable knock-down or knock-out approaches for interrogating gene function in hTSCs and organoids. We are working on optimizing a Lenti shRNA BCAM approach to sort out cells expressing the shRNA vector, but we are still characterizing these cells.

New data examining the affect of BCAM silencing in hTSCs is presented in Fig. S5 and described in Lines 401- 407.

11. For Fig. 5F, a co-immunostaining with a SCT marker (hCG as in Fig. 4B) should be included to indicate the presence of ST, otherwise don't you think that the SCT indicator should be removed.

We thank the reviewer for catching this oversight. We have removed the descriptive text within the IF images indicating the SCT core.

12. It would be more helpful for the readers to follow along if the authors could cite each of the sub-panels of the supplementary figures in a more explicit way (Fig. S1A, Fig. S1B, Fig. S1C etc) rather than just citing the whole supplementary figure in each statement in the manuscript.

We thank the reviewer for pointing this out and apologize for the lack of clarity. We now explicitly indicate each sub-panel for each supplementary figure in the text of the manuscript.

13. In line 220-223, it would be less confusing for the readers if the authors could rephrase 'three hTSC lines' to 'three lines of hTSC organoids' since that it was the organoids that were used for sequencing rather than the monolayer hTSC lines.

We have modified the sentence as suggested by the reviewer; Line 239.

14. The in-text figure citation in line 237 does not correlate to a figure panel showing expression of CGB and CSH1.

We thank the reviewer for catching this mistake. We now reference the correct figure panel Fig. 3B.

Second decision letter

MS ID#: DEVELOP/2021/199840

MS TITLE: Cell trajectory modeling identifies a primitive trophoblast state defined by BCAM enrichment

AUTHORS: Matthew J Shannon, Jennet Baltayeva, Barbara Castellana, Jasmin WÃ¤chter, Gina L McNeill, Ji Soo Yoon, Jenna Treissman, Hoa T Le, Pascal M Lavoie, and Alexander G Beristain

I have now received all the referees reports on the above manuscript, and have reached a decision. The referees' comments are appended below, or you can access them online: please go to BenchPress and click on the 'Manuscripts with Decisions' queue in the Author Area.

The overall evaluation is positive and we would like to publish a revised manuscript in Development, provided that the referees' comments can be satisfactorily addressed. Please attend to all of the reviewers' comments in your revised manuscript and detail them in your point-by-point response. If you do not agree with any of their criticisms or suggestions explain clearly why this is so.

We are aware that you may currently be unable to access the lab to undertake experimental revisions. If it would be helpful, we encourage you to contact us to discuss your revision in greater detail. Please send us a point-by-point response indicating where you are able to address concerns raised (either experimentally or by changes to the text) and where you will not be able to do so within the normal timeframe of a revision. We will then provide further guidance. Please also note that we are happy to extend revision timeframes as necessary.

Reviewer 1

Advance summary and potential significance to field

This manuscript describes single cell RNA sequencing (scRNAseq) of early and late first trimester human placenta, and describing different cytotrophoblast (CTB) progenitor states, leading to identification of BCAM-high cells as potential progenitor CTB, whose enrichment leads to enhanced growth of trophoblast organoids.
While there have been a lot of publications in this area, this manuscript advances the field by describing different CTB progenitor states, leading to identification as well as further characterization of one such CTB state. The authors have now added significant additional analysis of single cell RNAseq data from both TSC organoids and first trimester placenta, as well as performed additional experiments with their BCAM-hi and BCAM-lo organoids, strengthening their conclusions.

Comments for the author

The authors have thoroughly and sufficiently addressed all reviewers' concerns. I have no further suggestions.

Reviewer 2

Advance summary and potential significance to field

Overall, the authors have made a significant and welcome effort to thoroughly address all reviewers' comments. It is also appreciated that they requested the additional time to conduct the experiments requested. They added some more data and also included additional analyses that have improved the manuscript. The lack of preferential differentiation trajectory bias of BCAM-positive cells casts some doubt on how informative this marker really is. But overall, the study is still important and a great addition to the field.

After reading the authors' responses, this reviewer is now confused whether the 50,000 starting cells are the authors' 'own' sequenced cells, or whether this number is reached after inclusion of the Vento-Tormo data. I.e., are the samples listed under "Files 8-21" in Supp. Table 1 from the Vento-Tormo paper? This is now how I understand it. This reduces the overall power of the current analysis quite dramatically from my previous assumption. The authors should explicitly state in the text (Results, p. 6) how many cells originate from their own study, and how many of these are trophoblast (this number is then in the range of some 1,500-1,800 cells only as "original" data as opposed to re-analysis of previously published data, correct?). This essential information appears to be intentionally blurred or at least very hard to tease out.

A few responses to the authors' comments are not quite valid but are of minor concern and can be easily remedied. An siRNA approach does not preclude co-transfection with a GFP marker that would have allowed to enrich for transfected (and hence knocked down) cells. This is a minor point and does not justify a repeat experiment.

After explanation, the terminology of forward and reverse transfection is unnecessary and certainly not widely used. The simple difference is whether cells are pre-seeded before transfection or transfected in suspension on plating. Both methods are commonly used. The authors should just state it like this and not confuse the reader by the unconventional terminology.

The newly added BCAM IF photo in Fig. 5B (top) is blurry. Please could the authors replace this with a higher resolution photo.

Comments for the author

See above.

Reviewer 3

Advance summary and potential significance to field

As per my original assessment:
In this article, Shannon et al. performed single cell RNA sequencing (scRNA-seq)
on the first trimester placentas across 5-12 weeks gestational age, with a lineage trajectory analysis identifying the trophoblast progenitor origin and BCAM as a marker of trophoblast cells in this progenitor state. This was followed by a similar analysis on in vitro generated trophoblast organoids leading to observations in regards to similarities and the extent on which these organoids can model some of the cellular aspects of the early placenta.
Furthermore, they use trophoblast organoid assays to validate several of their findings in regards to BCAM+ cells. In general this is a very nice paper, well timed which uses a good mix of in vivo and in vitro data, which no doubt will be of great interest for the field.

Comments for the author

The authors have addressed, when possible, my main concerns with new experiments or analysis. As mentioned before, this is a very interesting manuscript and very timely. In my opinion, the manuscript is ready for publication.

Second revision

Author response to reviewers' comments

Response to Reviewers

We thank all three reviewers for reviewing our revised manuscript. Your comments and insights have been helpful for crafting a better manuscript, and we feel the previous and new changes we made based on the reviewer's suggestions have strengthened the manuscript's overall quality.

Reviewer 1:
Reviewer 1 Advance Summary and Potential Significance to Field:

This manuscript describes single cell RNA sequencing (scRNAseq) of early and late first trimester human placenta, and describing different cytotrophoblast (CTB) progenitor states, leading to identification of BCAM- high cells as potential progenitor CTB, whose enrichment leads to enhanced growth of trophoblast organoids. While there have been a lot of

publications in this area, this manuscript advances the field by describing different CTB progenitor states, leading to identification as well as further characterization of one such CTB state. The authors have now added significant additional analysis of single cell RNAseq data from both TSC organoids and first trimester placenta, as well as performed additional experiments with their BCAM-hi and BCAM-lo organoids, strengthening their conclusions.

We thank Reviewer 1 for the thorough comments on our work. We are happy that the reviewer feels our revisions have made our manuscript stronger.

Comments for the author:

The authors have thoroughly and sufficiently addressed all reviewers' concerns. I have no further suggestions.

We thank Reviewer 1 for revisiting our work, acknowledging our inclusion of new experiments/data, and providing their final opinion of our paper.

Reviewer 2:
Reviewer 2 Advance Summary and Potential Significance to Field:

Overall, the authors have made a significant and welcome effort to thoroughly address all reviewers' comments. It is also appreciated that they requested the additional time to conduct the experiments requested. They added some more data and also included additional analyses that have improved the manuscript. The lack of preferential differentiation trajectory bias of BCAM-positive cells casts some doubt on how informative this marker really is. But overall, the study is still important and a great addition to the field.

We thank the reviewer for acknowledging that the revisions we made strengthened our initial submission.

After reading the authors' responses, this reviewer is now confused whether the 50,000 starting cells are the authors' 'own' sequenced cells, or whether this number is reached after inclusion of the Vento-Tormo data. I.e., are the samples listed under "Files 8-21" in Supp. Table 1 from the Vento-Tormo paper? This is now how I understand it. This reduces the overall power of the current analysis quite dramatically from my previous assumption. The authors should explicitly state in the text (Results, p. 6) how many cells originate from their own study, and how many of these are trophoblast (this number is then in the range of some 1,500-1,800 cells only as "original" data as opposed to re-analysis of previously published data, correct?). This essential information appears to be intentionally blurred or at least very hard to tease out.

We apologise that we were not able to make clear the source and breakdown of dataset and cell numbers linked to each dataset. The 50,790 cells shown in Figure 1B as well as the 7,798 cells shown in Figure 1C are cells originating from both our data (GSE174481) and the Vento-Tormo data (E-MTAB-6701) following data integration.

In the previous revision Supplemental Figure 2A and 2C showed how each dataset contributes to cells that were sequenced and fed into downstream analyses; 12,794 cells from our in-house dataset (GSE174481), 37,996 cells from the Vento-Tormo dataset (E-MTAB-6701)]. The discrepancy in cell number is largely due to the inclusion of decidual cells in the Vento-Tormo dataset. We included decidual cells in our analyses as cells from the decidua contain interstitial EVT, a trophoblast population we likely did not capture in our own chorionic villous-derived dataset. Our in-house data is derived from 7 placental samples, that following stringent quality control measures, left us with 3843 high quality trophoblasts. This is to be directly compared with the placental villi (4 samples) and decidual EVT (4 samples) trophoblast contribution of the Vento-Tormo dataset leading to 3955 high quality trophoblasts following our quality control pipeline. In summary, our dataset contributed to about 50% of trophoblast content for all downstream bioinformatic analyses.

While we did not state the breakdown of these cell numbers in the main text of the results in previous revision, we now include these details and make reference to Supplemental Figure 2A and C. Please refer to changes on Page 6, lines 124-126, and lines 142-143. We hope these changes have made source of dataset and associated cell number clearer.

A few responses to the authors' comments are not quite valid but are of minor concern and can be easily remedied. An siRNA approach does not preclude co-transfection with a GFP marker that would have allowed to enrich for transfected (and hence knocked down) cells. This is a minor point and does not justify a repeat experiment.

The reviewer is correct that co-transfection strategies with marker constructs along with siRNA can provide a rough tool for transfection enrichment. We thank the reviewer for pointing this out to us.

After explanation, the terminology of forward and reverse transfection is unnecessary and certainly not widely used. The simple difference is whether cells are pre-seeded before transfection or transfected in suspension on plating. Both methods are commonly used. The authors should just state it like this and not confuse the reader by the unconventional terminology.

As requested, we have changed the wording to: "CT29 hTSCs cultured in 2D were seeded then transfected using Lipofectamine RNAiMAX (Thermo Fisher) to achieve efficient BCAM silencing prior to organoid establishment. Following 48 hr, hTSCs cells were transfected in suspension prior to 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" These changes can be found on page 28, lines 773-778. We agree with the reviewer that this new wording is more understandable than our previous description that used the terms "forward" and "reverse" transfections.

The newly added BCAM IF photo in Fig. 5B (top) is blurry. Please could the authors replace this with a higher resolution photo.

We respectfully point out to the reviewer that the top Fig. 5B image (hCG/BCAM) was not replaced or altered from the original submission. The image below this (Ki67/BCAM) was replaced in the previous revision as the reviewer was initially concerned about nuclear localization of the BCAM signal.

Nonetheless, the reviewer is right that the hCG/BCAM image is a little blurry. The blurriness is likely accentuated because the new representative Ki67/BCAM image was a maximum intensity projection of a deconvolution rendering from 120 Z-stacks (at 0.24 µm optical sections). This was my oversight, and I apologise as I was not aware the deconvolution image was used in the revision; I thought a wide-field image was being used as in the original submission.

To make the images consistent with each other and consistent with descriptions in the materials and methods, we have now removed the deconvolved Ki67/BCAM image and replaced it with an in-focus Z-stack wide-field image (imaged using a 40X oil 1.4NA objective).

Further, by examining the separate wide-field channels of the hCG/BCAM image, we find that the nuclei (DAPI) are not in focus, while the BCAM signal is in focus. In our experience, this is a common outcome in wide field microscopy applications. Therefore, to better show the BCAM and Ki67 signal more clearly, we now have restructured the image to show the merged image alongside separate Ki67 and BCAM channels. We feel, however, that the goal of showing BCAM localization within trophoblast organoids is retained regardless of which imaging technique we use. We hope the changes to Fig. 5 improve the clarity of BCAM overlap with proliferating CTB.

Reviewer 3:
Reviewer 3 Advance Summary and Potential Significance to Field.

As per my original assessment: In this article, Shannon et al. performed single cell RNA sequencing (scRNA-seq) on the first trimester placentas across 5-12 weeks gestational age, with a lineage trajectory analysis identifying the trophoblast progenitor origin and BCAM as a marker of trophoblast cells in this progenitor state. This was followed by a similar analysis on in vitro generated trophoblast organoids, leading to observations in regards to similarities and the extent on which these organoids can model some of the cellular aspects of the early placenta. Furthermore, they use trophoblast organoid assays to validate several of their findings in regards to BCAM+ cells. In general this is a very nice paper, well timed which uses a good mix of in vivo and in vitro data, which no doubt will be of great interest for the field.

We thank the reviewer for their input into our manuscript, and we are pleased that the reviewer feels our revisions have strengthened our findings/paper.

Reviewer 3 Comments for the Author...

The authors have addressed, when possible, my main concerns with new experiments or analysis. As mentioned before, this is a very interesting manuscript and very timely. In my opinion, the manuscript is ready for publication.

Third decision letter

MS ID#: DEVELOP/2021/199840

MS TITLE: Cell trajectory modeling identifies a primitive trophoblast state defined by BCAM enrichment

AUTHORS: Matthew J Shannon, Jennet Baltayeva, Barbara Castellana, Jasmin WÃ¤chter, Gina L McNeill, Ji Soo Yoon, Jenna Treissman, Hoa T Le, Pascal M Lavoie, and Alexander G Beristain
ARTICLE TYPE: Research Article

I am happy to tell you that your manuscript has been accepted for publication in Development, pending our standard ethics checks.