



# Single Nucleus RNA Sequence (snRNAseq) Analysis of the Spectrum of Trophoblast Lineages Generated From Human Pluripotent Stem Cells *in vitro*

## OPEN ACCESS

### Edited by:

Ruby Yun-Ju Huang,  
National Taiwan University, Taiwan

### Reviewed by:

Shurong Liu,  
Sun Yat-sen University, China  
Yu Zhao,  
Capital Medical University, China

### \*Correspondence:

Geetu Tuteja  
geetu@iastate.edu  
R. Michael Roberts  
robertsrm@missouri.edu

† These authors have contributed  
equally to this work and share first  
authorship

### Specialty section:

This article was submitted to  
Molecular Medicine,  
a section of the journal  
Frontiers in Cell and Developmental  
Biology

Received: 14 April 2021

Accepted: 21 June 2021

Published: 21 July 2021

### Citation:

Khan T, Seetharam AS, Zhou J,  
Bivens NJ, Schust DJ, Ezashi T,  
Tuteja G and Roberts RM (2021)  
Single Nucleus RNA Sequence  
(snRNAseq) Analysis of the Spectrum  
of Trophoblast Lineages Generated  
From Human Pluripotent Stem Cells  
*in vitro*.  
*Front. Cell Dev. Biol.* 9:695248.  
doi: 10.3389/fcell.2021.695248

Teka Khan<sup>1,2†</sup>, Arun S. Seetharam<sup>3,4†</sup>, Jie Zhou<sup>1,5</sup>, Nathan J. Bivens<sup>6</sup>, Danny J. Schust<sup>5</sup>,  
Toshihiko Ezashi<sup>1,2</sup>, Geetu Tuteja<sup>4\*</sup> and R. Michael Roberts<sup>1,2,7\*</sup>

<sup>1</sup> Christopher S Bond Life Sciences Center, University of Missouri, Columbia, MO, United States, <sup>2</sup> Division of Animal Sciences, Bond Life Sciences Center, University of Missouri, Columbia, MO, United States, <sup>3</sup> Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, IA, United States, <sup>4</sup> Genetics, Development and Cell Biology, Iowa State University, Ames, IA, United States, <sup>5</sup> Department of Obstetrics and Gynecology, University of Missouri School of Medicine, Columbia, MO, United States, <sup>6</sup> DNA Core Facility, University of Missouri, Columbia, MO, United States, <sup>7</sup> Department of Biochemistry, University of Missouri, Columbia, MO, United States

One model to study the emergence of the human trophoblast (TB) has been the exposure of pluripotent stem cells to bone morphogenetic protein 4 (BMP4) in presence of inhibitors of ACTIVIN/TGFB; A83–01 and FGF2; PD173074 (BAP), which generates a mixture of cytotrophoblast, syncytiotrophoblast, and cells with similarities to extravillous trophoblast. Here, H1 human embryonic stem cells were BAP-exposed under two O<sub>2</sub> conditions (20% and 5%, respectively). At day 8, single nuclei RNA sequencing was used for transcriptomics analysis, thereby allowing profiling of fragile syncytial structures as well as the more resilient mononucleated cells. Following cluster analysis, two major groupings, one comprised of five (2,4,6,7,8) and the second of three (1,3,5) clusters were evident, all of which displayed recognized TB markers. Of these, two (2 and 3) weakly resembled extravillous trophoblast, two (5 and 6) strongly carried the hallmark transcripts of syncytiotrophoblast, while the remaining five were likely different kinds of mononucleated cytotrophoblast. We suggest that the two populations of nuclei within syncytiotrophoblast may have arisen from fusion events involving two distinct species of precursor cells. The number of differentially expressed genes between O<sub>2</sub> conditions varied among the clusters, and the number of genes upregulated in cells cultured under 5% O<sub>2</sub> was highest in syncytiotrophoblast cluster 6. In summary, the BAP model reveals an unexpectedly complex picture of trophoblast lineage emergence that will need to be resolved further in time-course studies.

**Keywords:** human embryonic stem cells, BMP4, syncytiotrophoblast, placenta, trophoblast, extravillous trophoblast, snRNASeq

## INTRODUCTION

Early human gestation, especially the approximately 14 days between conception and the anticipated onset of the next menstruation period, is associated with high rates of pregnancy loss (Wilcox et al., 1988). It is the time in pregnancy when the trophoblast (TB) lineage emerges, implantation begins, and a trophoblast interface is established with the mother that allows for physical and nutritional support of the conceptus. Critically, it is also the period when signals from the conceptus trigger the phenomenon of maternal recognition of pregnancy, which prevents the mother from a return to ovarian cyclicity and provides the conceptus some measure of control over maternal physiology and its own future. TB, as the dominant lineage of the emerging placenta, has a major responsibility for all these functions. However, for a range of ethical and practical reasons, these first critical weeks of pregnancy have been largely inaccessible to experimental study. As a consequence, there has been interest in developing models that mimic some of these early events of placental TB emergence *in vitro*.

One approach that has been used to address our poor understanding of these enigmatic early days of pregnancy is extended embryo culture beyond the blastocyst stage (reviewed by Zhou et al. (2021), In Press), made possible because of the availability of “spare” human embryos from *in vitro* fertilization clinics and improved culture media that permitted development until at least day (d) 13 (Deglincerti et al., 2016; Shahbazi et al., 2016). These embryos, although separated from maternal influence, appear to follow a differentiation pathway quite similar to that inferred from the limited number of histological studies performed on *in vivo* material many years before (Hertig et al., 1956; Boyd and Hamilton, 1970). Single-cell RNA sequence (scRNAseq) analysis performed on such embryos at d 8, 10, and 12 of culture revealed the presence of emergent TB populations for syncytioTB and extravillous-like cytoTB that appeared to be similar to, but nonetheless distinct from, TBs that populate the villous placental structures that arise a little later in pregnancy (West et al., 2019). Extended human embryo culture has also provided insights into the mechanisms of TB lineage divergence around the time of implantation (Lv et al., 2019; West et al., 2019; Zhou et al., 2019).

A second model to study the TBs of early pregnancy has been to generate them from pluripotent stem cells, specifically human embryonic stem cells (hESCs) and induced pluripotent stem cells after exposure to BMP4. This pathway of directed differentiation was first reported by Xu et al. (2002) and continues to be used extensively, albeit with a range of modifications [Reviewed in Roberts et al. (2018); Horii et al. (2019, 2020)] and especially

the exclusion of FGF2 (Das et al., 2007). Among the refinements has been to include two pharmaceutical reagents, the ACTIVIN signaling inhibitor A83-01, and the FGF2 signaling inhibitor PD1730, so-called BAP medium (Amita et al., 2013; Yang et al., 2015). Under this regimen, virtually all cells become positive for the pan-trophoblast marker KRT7 by 48 h. HLA-G positive cells, indicative of extravillousTBs, appear soon after and reach a maximum by d 5–6, while the production of human chorionic gonadotropins (hCG) from syncytioTB begins at about d 5 and increases markedly thereafter. The colonies can be maintained in culture for approximately 8–9 days before regions of syncytium begin to loosen from the Matrigel substratum. The BAP model has allowed different-sized cell types, the largest being syncytium (syncytioTB) and the smallest mononucleated cytoTBs, to be partially characterized and compared (Yabe et al., 2016). Microarray analysis has also been performed on HLA-G positive cells isolated on magnetic beads, which demonstrated a resemblance in terms of its markers to extravillousTB derived from placenta (Telugu et al., 2013).

While microarray and RNAseq data from hESC colonies differentiated in response to BMP4 have shown the differentiated cells to be comprised only of TB, with no evidence for contributions from the main germ line lineages, particularly mesoderm (Schulz et al., 2008; Marchand et al., 2011; Telugu et al., 2013; Yabe et al., 2016; Jain et al., 2017), it has become clear that they displayed a different transcriptome profile from primary TBs isolated from the villous placenta. This realization led us to hypothesize that BAP best modeled the primitive placental structure that forms when the blastocyst first implants and establishes residency in the uterine wall. Although the original microarray and RNAseq studies were informative about the general nature of the TB that formed in response to directed differentiation, they provided little detail about the identity of the different sub-lineages that formed, their relationships to each other, and how they originated. Here we have employed a single nuclei RNAseq approach, which allowed an analysis of fragile syncytioTB structures as well as the more resilient cytoTBs, to display an unexpected heterogeneity of TB sub-lineages in d 8 cultures.

## MATERIALS AND METHODS

### Human Embryonic Stem Cell Culture and Trophoblast Differentiation

Human ESCs (H1; WA01) originated from WiCell Research Institute. They were cultured and maintained in Matrigel (BD Bioscience)-coated 6-well tissue culture plates (Thermo Scientific) on mTeSR1 medium (Stemcell Technologies) under two different O<sub>2</sub> concentration conditions, i.e., 20% and 5% at 37°C, as described previously (Yabe et al., 2016). The medium was replaced daily and cells (50,000/well) passaged approximately every 5–6 days. The protocol first described by Amita et al. (2013) was used to drive differentiation of the ESC to TB. Briefly, the mTeSR1 culture medium, which contains a high concentration of FGF2 (100 ng/ml), was replaced with DMEM/F12 medium (Thermo Scientific)

**Abbreviations:** ANOVA, Analysis of variance; AP, A83-01, PD1730; BAP, Bone morphogenic protein-4, A83-01, PD1730; bp, Base pair; CytoTB, Cytotrophoblast; GEMs, Gel Bead-in-Emulsion; hESCs, Human embryonic stem cells; UMAP, Uniform manifold approximation and projection; UMIs, Unique Molecular Identifiers; KOSR, Knock-out serum replacement; MEFs, Mouse embryonic fibroblast; NWRB, Nuclear wash and resuspension buffer; PCA, Principle component analyses; scRNAseq, Single cell RNA sequence; SNN, Shared nearest neighbor; snRNA, Single nucleus RNA sequence; SyncytioTB, Syncytiotrophoblast; TB, Trophoblast; t-SNE, t-distributed stochastic neighbor embedding; VST, Variance stabilizing transformation.

containing knock-out serum replacement (KOSR, Invitrogen) and low FGF2 (4 ng/ml) that had been conditioned by mouse embryonic fibroblast (MEFs) (**Supplementary Figure 1**). After 24 hours, the conditioned medium was replaced with daily changes of DMEM/F12/KOSR medium minus FGF2 but containing BMP4 (10 ng/ml), A83-01 (1  $\mu$ M), and PD1730740 (0.1  $\mu$ M), so-called BAP treatment, for 3 days and then the same medium without BMP4 for the following 4 days (AP treatment) (**Supplementary Figure 1A**). The content of hCG in the culture medium was measured on d 7 and d 8 by ELISA as described by Amita et al. (2013). The hCG concentration was normalized to  $10^5$  nuclei for each replicate cultured under the 20% O<sub>2</sub> and 5% O<sub>2</sub> conditions. One-way ANOVA was applied by using GraphPad Prism.

## Nuclei Isolation

On d 8 of differentiation, each culture well was rinsed twice with DMEM/F12 and the colonies were partially dispersed by using gentle cell disassociation reagent (GDR) (Stemcell Technologies) (0.6 ml/well; 7 min at 37°C). The colonies were fully dislodged from the substratum by gentle scraping, and nuclei were isolated from the disassociated cells by the protocol provided by 10X Genomics with minor modifications (10x Genomics, 2019). Briefly, 1 ml of chilled lysis buffer (10 mM Tris-HCl, pH 7.4; 10 mM NaCl; 3 mM MgCl<sub>2</sub>; and 0.1% Non-idet<sup>TM</sup> P40 Substitute in Nuclease-Free Water) was added to each well through a wide-bore pipet tip. After gently pipetting several times, the suspension was kept on ice for 5 min to achieve maximum lysis of cells and centrifuged (500  $\times$ g; 5 min at 4°C). The pellet was resuspended in 1 ml of nuclear wash and resuspension buffer (NWRB) (1X PBS with 1.0% w/v bovine serum albumin and 0.2 U/ $\mu$ l RNase inhibitor), cell debris were removed by filtration through a 40  $\mu$ m Nylon cell strainer (Falcon), and the solution was recentrifuged as above. The supernatant solution was removed, and the final nuclear pellet was suspended in 1 ml of NWRB. The complete workflow is documented in **Supplementary Figure 1**.

## Single Nuclear 3' RNA-Seq Library Preparation

Libraries were constructed by following the manufacturer's (10x Genomics, 2019) protocol with reagents supplied in the 10x Genomics Chromium Next Gel Bead-in-Emulsion (GEMs) Single Cell 3' Kit v3.1. Briefly, the concentrations of nuclei and intact cells in the nuclear preparation were measured with an Invitrogen Countess II automated cell counter. An aliquot of the nuclear suspension (1,000 nuclei per microliter), reverse transcription master mix, and partitioning oil were loaded on a Chromium Next GEM G chip with a nuclear capture target of 5,000 nuclei per library. Post-Chromium controller GEMs were transferred to a PCR strip tube, and reverse transcription was performed on an Applied Biosystems Veriti thermal cycler at 53°C for 45 min. cDNA was amplified for 14 cycles and purified by using Axygen AxyPrep MagPCR Clean-up beads. cDNA fragmentation, end-repair, A-tailing, and ligation of sequencing adaptors was performed according to manufacturer specifications. The final library was quantified with the Qubit HS DNA kit, and the fragment size was analyzed by means

of an Agilent Fragment Analyzer system. Libraries were pooled and sequenced on an Illumina NovaSeq S1 flow cell with a goal to generate 50,000 reads per nucleus with a sequencing configuration of 28 base pair (bp) on read1 and 91 bp on read 2 using unique dual indexes.

## Single-Nuclei Sequencing Analysis

The University of Missouri Informatics Research Core Facility pre-processed the data with the 10XGenomics/Cell Ranger software (v4.0.0) (Zheng et al., 2017), including demultiplexing, fastq file generation, read alignment, and UMI quantification. Cell Ranger was run with default options against the ENSEMBL GRCh38 reference genome (Schneider et al., 2017) with both pre-mRNA and mRNA feature files. The data were then processed by means of Seurat (v3.2.2) (Stuart et al., 2019), following the recommended practices for scRNAseq data with replicates. All steps used default options unless noted otherwise. Briefly, an expression matrix (count table) containing Unique Molecular Identifiers (UMIs) per nucleus per gene was imported for each replicate as the 10 $\times$  data object. Only nuclei with greater than 200 but less than 7,500 genes and less than 15% of genes originating from mitochondrial sources were retained. Once the nuclei were quality filtered, the data were imported as a Seurat object, and all mitochondrial and ribosomal protein-coding genes were removed from the expression matrix. The replicates for each condition were integrated by using the FindIntegrationAnchors and IntegrateData commands. The count matrix was scaled and normalized by variance stabilizing transformation (VST) with Seurat's ScaleData and NormalizeData commands, respectively. The 2,000 most variable features were then selected with the FindVariableFeatures command for the Principle Component Analyses (PCA), which was performed by the RunPCA command. The PCs generated by the PCA were assessed with ElbowPlot and JackStraw analyses by using up to 20 different components. The resulting PCs were used for Jaccard-weighted, shared nearest neighbor (SNN) distance calculations and graph generation. The graph was then subjected to Louvain clustering and Uniform Manifold Approximation and Projection (UMAP) for dimension reduction in order to visualize nuclear transcriptomic profiles in two-dimensional space. After changing the default assay to RNA in the Seurat object, FindMarkers was run to determine the marker genes (genes with a fold change of greater than or equal to 1.5 with an adjusted p-value of less than 0.05) in each cluster by comparing each gene's expression level against other clusters.

Cluster identities were approximated by using single-cell RNA-Seq data from first trimester human placenta (Vento-Tormo et al., 2018). We used cell-type specific genes as described in PlacentaCellEnrich (Jain and Tuteja, 2021) within the TissueEnrich Bioconductor package (Jain and Tuteja, 2019).

To perform differential expression analysis between O<sub>2</sub> conditions, metadata in the Seurat object was updated with a new column containing both information about the cluster and the condition. This column was set as Seurat objects identity, and differential expression analyses were performed by using the FindMarkers command. No down-sampling of any cluster was performed since the corresponding clusters had roughly the same number of cells. Genes were considered upregulated in a

condition if the fold change was greater than or equal to 1.5, and the adjusted p-value was less than 0.05.

The GitHub repository documenting all the analyses steps is available at [https://github.com/Tuteja-Lab/BAP.hESC.d8\\_snRNAseq](https://github.com/Tuteja-Lab/BAP.hESC.d8_snRNAseq). Sequencing data have been deposited in the Gene Expression Omnibus under accession ID GSE171768.

## RESULTS

### Heterogeneity of Cell Types

Exposure of either human embryonic stem cells or human induced pluripotent cells to BAP conditions for five days or more results in the initiation of syncytiotrophoblast formation evident as ruffled areas within the colonies and a steep increase in *CGA* and *CGB* expression and hCG production over subsequent days (Amita et al., 2013; Yabe et al., 2016). As previously described, these cells show downregulation of pluripotency genes, and upregulation of trophoblast marker genes compared to undifferentiated cells (Supplementary Table 1). *CGA*-expressing TBs are known to be confined to what has previously been demonstrated to be the syncytial patches, with little or no staining outside these areas (Das et al., 2007; Schulz et al., 2008; Amita et al., 2013; Yang et al., 2015), while other trophoblast markers, such as HLA-G, are confined to areas outside the syncytium. The heterogeneity of cell types is further demonstrable by the heterogeneity of staining within colonies and differences in intensities of staining for particular trophoblast markers. Together, these data indicate that snRNAseq analysis might be a fruitful approach to examining lineage divergence and cell type diversity.

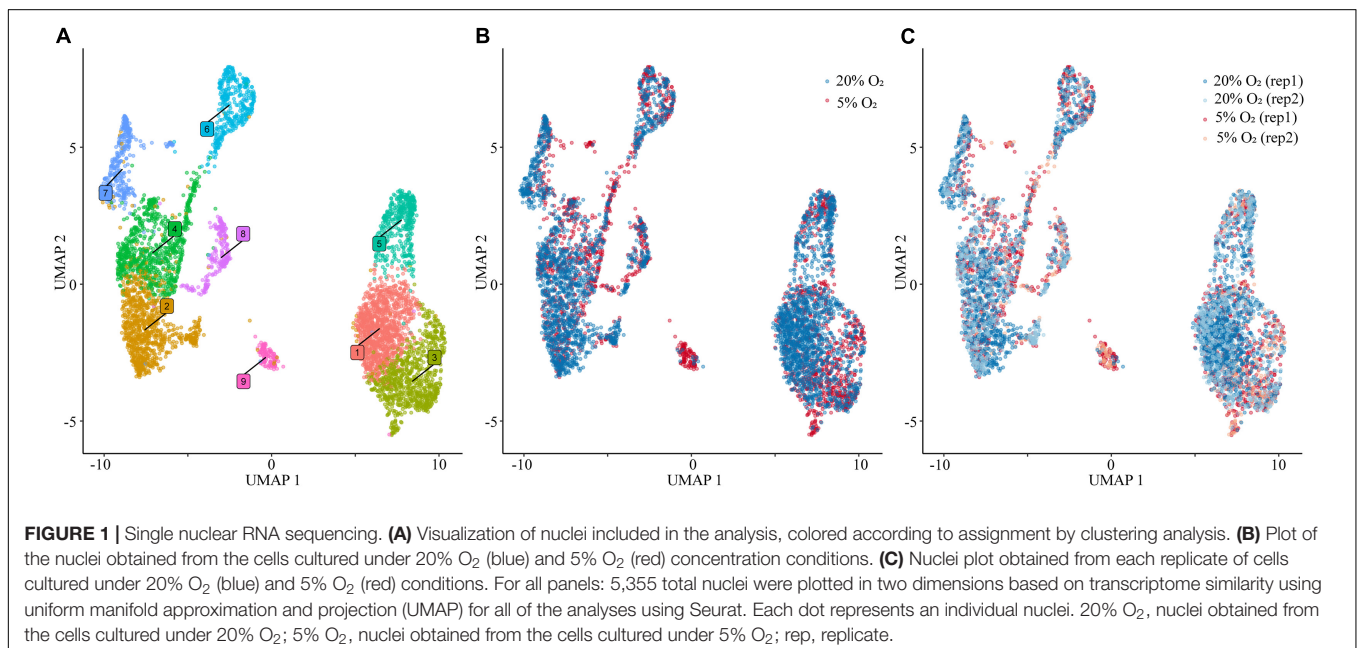
### Isolation of Nuclei

H1 ESCs were acclimated to 20% and 5% O<sub>2</sub> conditions prior to BAP-directed differentiation to TB (Supplementary Figure 1A).

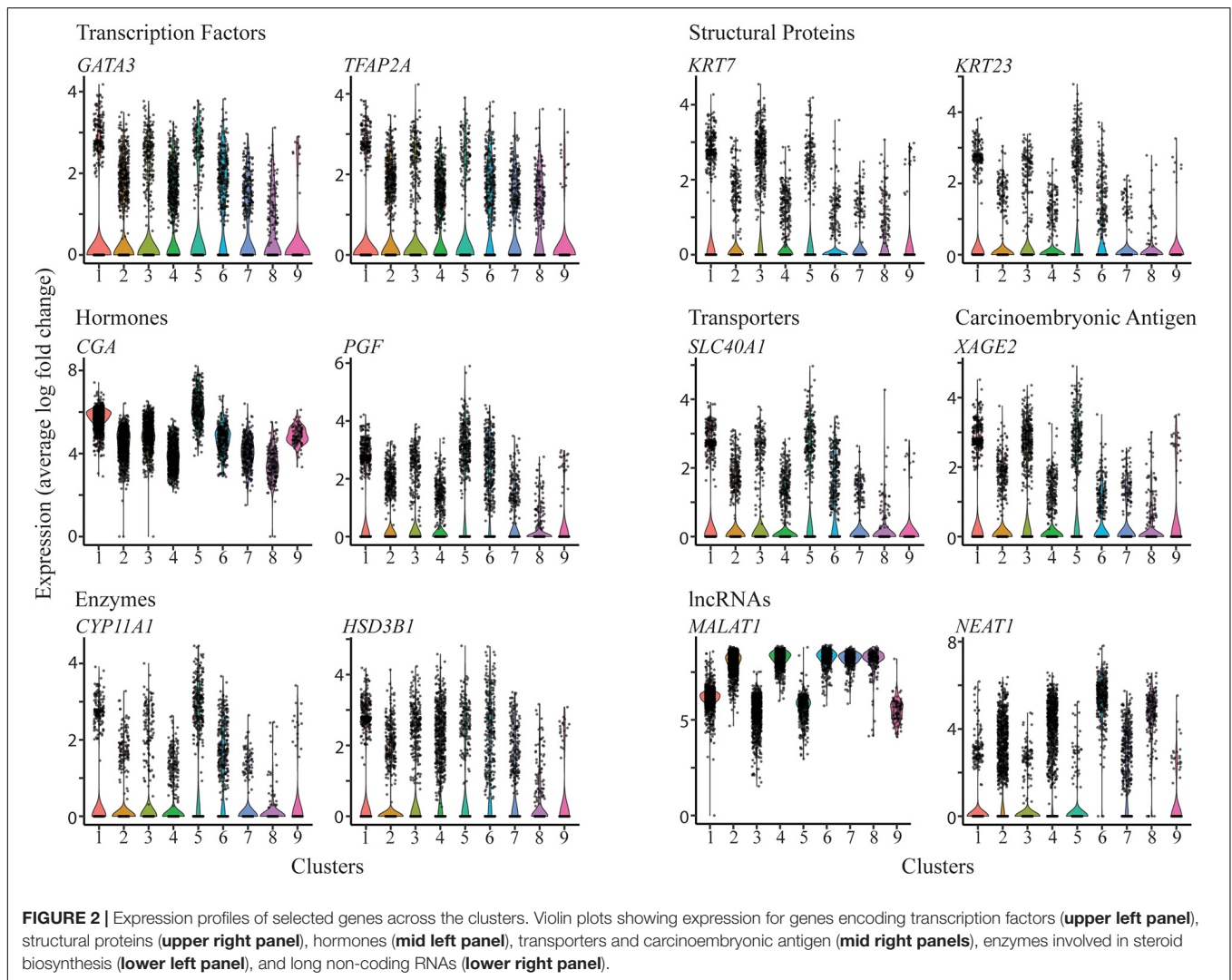
Sufficient hCG was released into the medium to be detected after d 5. By d7 and d8 copious amounts of hCG were being produced by cultures grown under both O<sub>2</sub> conditions (Supplementary Figure 1B), although previous studies had detected a significant lag between d6 and d8 under low O<sub>2</sub>, possibly reflective of a slower rate of differentiation than at 20% (Schulz et al., 2008; Amita et al., 2013). At d 8, the colonies were dissociated as completely as possible without destroying syncytial clumps, and nuclei were isolated (Supplementary Figure 1D). This experiment was repeated a second time, also under the same two O<sub>2</sub> conditions, in order to provide two replicates for each treatment group (Supplementary Figure 1A). The concentrations of nuclei used to prepare libraries for subsequent snRNAseq analysis were comparable, and the content of intact cells in all four preparations was 5% or less (Supplementary Figure 1F).

### snRNA-seq Analysis of Trophoblast Cells

After sequencing and data processing, nuclei were retained for analysis if they had between 200 and 7,500 unique genes detected, and less than 15% mitochondrial reads (Supplementary Figure 2 and Supplementary Table 2). With the possible exception of the first replicate from cultures under 20% O<sub>2</sub> conditions, the majority of mitochondrial contamination was associated with nuclei that contained a relatively low number of nuclear transcripts (Supplementary Figure 3). From the four preparations, this resulted in a total of 5,355 nuclei passing quality control filters (Supplementary Table 2). In addition to transcripts from mitochondrial genes, sequencing also revealed, as expected, the presence of ribosomal subunit RNAs in all nuclei. Accordingly, all data were filtered to remove mitochondrial and ribosomal RNA transcripts before further analysis with Seurat software.



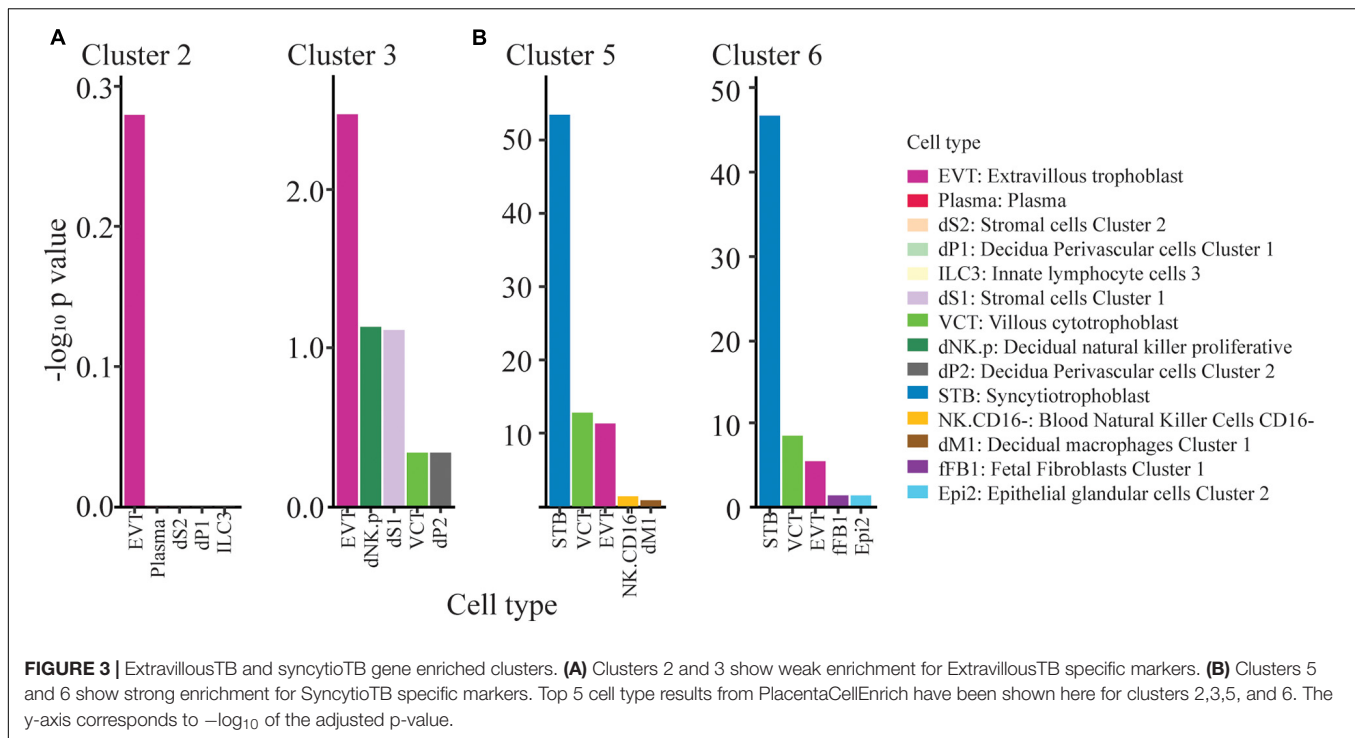




Following cluster analysis and projection of data in uniform manifold approximation and projection (UMAP) expression space, two major groupings, one comprised of five (2,4,6,7,8) and the second of three (1,3,5) clusters, were evident (**Figure 1**). Cluster 9 appeared to be separated from these two groupings. The gene expression signatures that distinguished the clusters are shown in **Supplementary Table 3** and illustrated graphically for each cluster in **Supplementary Figure 4**. The four different nuclear preparations replicated each other well (**Figure 1B**). Cluster 9 was again the exception and was comprised almost entirely by contributions from cells differentiated under 5% O<sub>2</sub>. All nine clusters contained nuclei that expressed genes that have traditionally been used in various combinations as markers for TB. These genes included, among others, ones encoding transcription factors, e.g., *GATA3*, *TFAP2A*, structural proteins, e.g., *KRT7*, *KRT23*, hormones, e.g., *CGA*, *PGF*, transporters, e.g., *SLC40A1*, the carcinoembryonic antigen *XAGE2*, and enzymes involved in steroid biosynthesis, e.g., *CYP11A1*, *HSD3B1*. However, it is also clear that the expression of any particular gene was variable between individual nuclei

within clusters and that mean expression values differed across clusters (**Figure 2**). In addition, all clusters had high levels of the placenta-enriched, long non-coding RNAs, *MALAT1* and *NEAT1* (**Figure 2**). Both of the latter were expressed most robustly in the grouping on the left (a) comprised of clusters 2, 4, 6, 7, and 8 compared to the right grouping (b) (clusters 1, 3, and 5) (**Figure 2**).

There was no significant expression of most marker genes examined for mesoderm, including *DLL3*, *FOXC1*, *RIPPLY*, *T/BRA*, *FOXA2*, *MIXL1*, endoderm, including *AFP*, *GATA4*, *GDF1*, *GDF3*, *MIXL2*, and ectoderm, including *FGF5*, *OTX2*, *SOX1*, *PAX6*. A few such marker genes, e.g., *TWIST2*, *GATA6*, *NES*, were detected at low levels (**Supplementary Figure 5**), but at least one of them (*TWIST2*) has a previously described functional association with TB (Ng et al., 2012). These data are generally consistent with the view expressed previously that the BAP-driven conversion of ESCs to TBs is largely, if not entirely, complete and leads to little or no differentiation of the ESCs along the main germline lineages (Amita et al., 2013; Yabe et al., 2016; Roberts et al., 2018).



## Trophoblast Nature of the Clusters

The question then arose as to the kind of TB represented in clusters 1–9. We used the PlacentaCellEnrich program (Jain and Tuteja, 2021) to determine if marker genes from clusters 1–9 were enriched for genes with cell-type specific expression in first trimester placenta, according to data from Vento-Tormo et al. (2018). This analysis provided strong evidence that clusters 5 and 6, were enriched for nuclei with a profile similar to that of placental syncytiotrophoblast, while cluster 3 embodied some features of extravillousTB from first trimester placenta (Figure 3). Cluster 2 showed most similarity to extravillousTB, although the adjusted P value when compared to the scRNAseq analysis of first trimester placenta conducted by Vento-Tormo et al. (2018) was not significant indicating that the similarities were low and therefore the cluster identity is unclear. Data for the remaining clusters 1, 4, 7, 8, and 9 were more equivocal, as they were not dominated by a particular set of TB markers, but rather expressed a majority of them (Figure 2). They may represent less well-differentiated cytoTB populations.

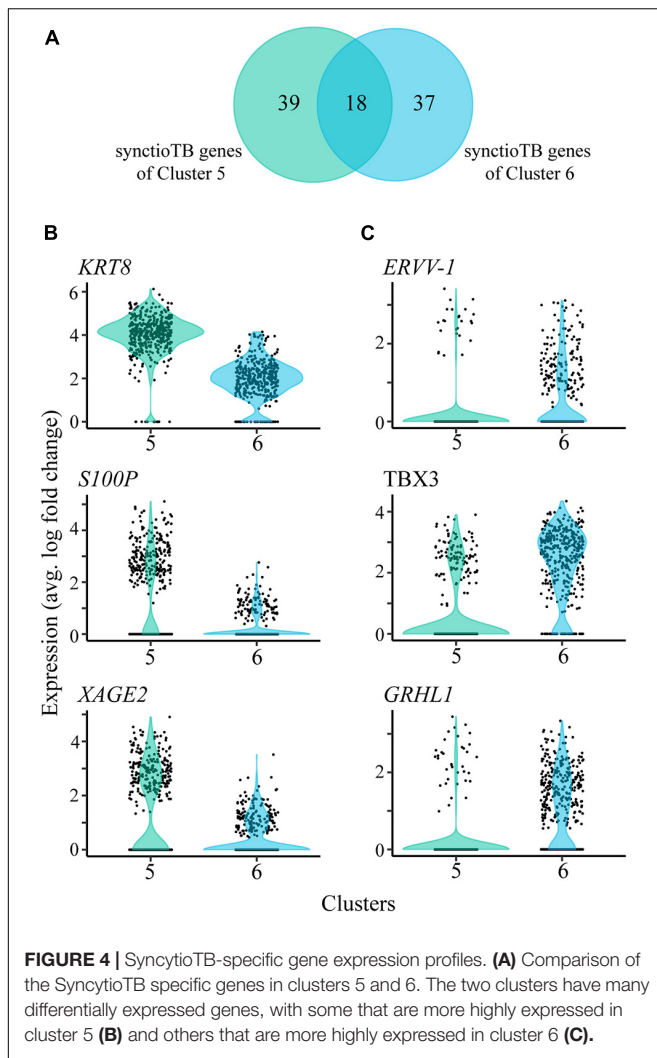
Clusters 5 and 6, which carry the more definitive hallmark features of placental syncytiotrophoblast, occupy the b and a cluster grouping (Figure 1A), respectively. However, they also differ markedly in the expression of several genes considered to be highly expressed in placental syncytiotrophoblast (Figure 4A, Supplementary Table 4). Of the 57 genes labeled as syncytiotrophoblast-specific in cluster 5 and the 55 genes labeled as syncytiotrophoblast-specific in cluster 6, only 18 were in common. Examples of genes up-regulated in cluster 5 compared to cluster 6 include *KRT8*, *S100P*, and *XAGE2* (Figure 4B). The reverse was observed for *ERVV-1*, and *TBX3* (Figure 4C), for example. A second endogenous retroviral gene, *ERVW-1*, had much

higher expression in cluster 6 than in cluster 5 (expression values 12.76 versus 4.27, respectively). In general, more transcription factors were identified with higher expression in cluster 6 (44 transcription factors) compared to cluster 5 (15 transcription factors).

Cluster 2 (in the major grouping a; Figure 1A) and Cluster 3 (group b; Figure 1A) each showed most similarity to first trimester human placental extravillousTB. Although the similarity was weak, the clusters also lacked indications of a major contribution of other TB cell types (Figure 3). The two are also clearly distinct from each other (Figure 1A). Among the transcripts that distinguish Cluster 2 from 3 are ones that encode TLE4, a transcriptional co-repressor that regulates WNT-mediated beta-catenin signaling, the procadherin PCDH9, and MAML2, a co-activator that binds to the intracellular domain of NOTCH receptors (Supplementary Figure 6). Cluster 3 noticeably possesses a group of upregulated genes that are overexpressed relative to Cluster 2 and whose functions are linked to the structure and organization of the cytoskeleton. Among these are *ACTG1*, *TMSB10*, and *TAGLN*, as well as three calcium binding proteins (*S100A11*, *S100A6*, and *S100A10*) (Supplementary Figure 6).

## Effects of O<sub>2</sub> Atmosphere

We identified differentially expressed genes between oxygen conditions in clusters 1–8. Cluster 9 is dominated by nuclei from 5% O<sub>2</sub>, which did not allow evaluation of differentially expressed genes. The number of differentially expressed genes for clusters 1–8 ranged from 37 to 188 (Supplementary Table 5). Of the TB clusters with most similarity to extravillousTB (cluster 2 and cluster 3), and the TB clusters predicted to be syncytiotrophoblast (cluster



5 and cluster 6), cluster 5 has the least number of differentially expressed genes (**Supplementary Table 5**). However, all of these clusters include genes associated with metabolism that were up-regulated in the 5% O<sub>2</sub> cultures (SLC2A3). Additionally, there are several other up-regulated transcripts in common in clusters 2,3,5, and 6, such as *CLIC3* and *FNI* (up-regulated in the 5% O<sub>2</sub> cultures) and *APOE*, *COL3A1*, and *LUM* (up-regulated in the 20% O<sub>2</sub> cultures; **Figure 5**, **Supplementary Figure 7**, **Supplementary Table 5**).

Because hypoxia inducible factor (HIF) is important in regulating gene expression in low O<sub>2</sub> conditions, we determined if previously published HIF targets (Ortiz-Barahona et al., 2010; Dengler et al., 2014) were upregulated in cells cultured under 5% O<sub>2</sub>. Indeed, we found that in all clusters, HIF targets that were differentially expressed in the data were almost always upregulated in nuclei from 5% O<sub>2</sub> rather than in nuclei from 20% O<sub>2</sub> (**Supplementary Table 6**). Certain HIF targets, such as *EGLN3* and *MMP2*, were upregulated in multiple clusters, whereas others, such as *ADM* and *VEGFA* were only upregulated in one cluster. The significance of these observations remains to be determined.

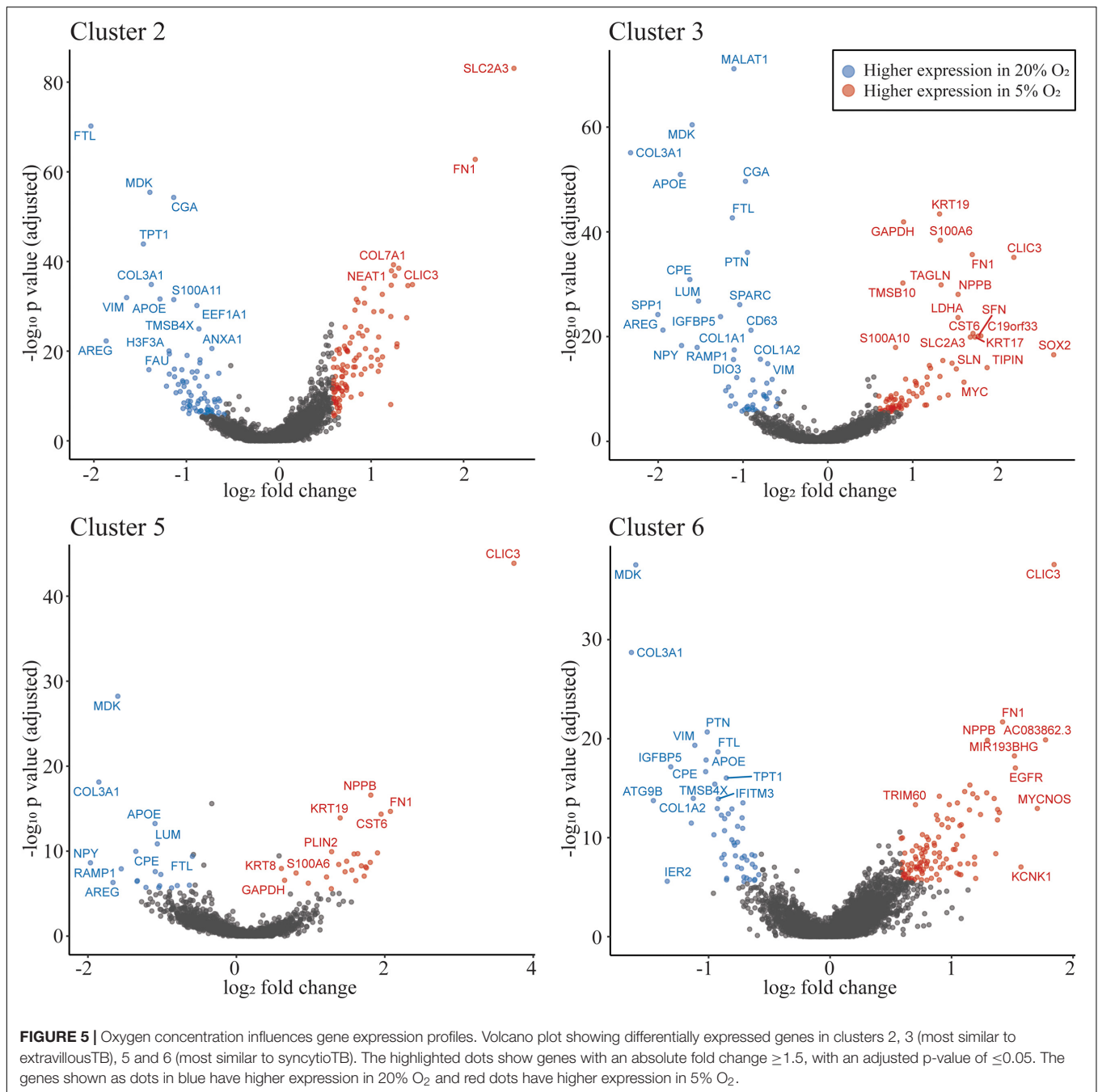
## DISCUSSION

The goals of this research were two. The primary objective was to study the heterogeneity in cells differentiated by the BAP protocol and determine whether there was evidence for the presence of different TB sub-lineages within the population. A second was to determine whether the oxygen atmosphere under which the colonies had been cultured influenced the state of differentiation following eight days of exposure to BAP conditions. Since syncytium, identified by its expression of CGA and CGB, comprises a significant and increasing fraction of the differentiating colonies after about d 5 (Yabe et al., 2016), and because such cellular structures are large and fragile, they were not amenable to scRNAseq analysis. Accordingly, we chose to analyze transcripts present in isolated nuclei, where the issues of cell breakage and size are less likely to confound outcomes. snRNAseq has recently been used to examine transcripts during mouse placental labyrinth development (Marsh and Belloch, 2020), where the evaluation of syncytial structures by single-cell procedures had proved to be difficult. It is also clear that the single-cell analysis performed on first trimester human placenta by Vento-Tormo et al. (2018) likely lacks a mature syncytioTB component (Roberts et al., 2021).

We note that the number of differentially expressed genes between O<sub>2</sub> conditions was variable across clusters. While many genes were identified in syncytioTB cluster 6 (169 differentially expressed genes), only 44 were identified in syncytioTB cluster 5. These results could be due to the proportion of nuclei associated with each cluster in each condition. Investigating additional stages of differentiation would allow assessing changes in overall cell populations between oxygen conditions. Much more significant differences are expected to occur in early differentiation where high O<sub>2</sub> appears to speed events (Amita et al., 2013), while low O<sub>2</sub> permits later stage catch-up as observed here with hCG production (**Supplementary Figure 1B**).

The data have confirmed what has previously been contested, namely that the BMP4-driven differentiation of pluripotent stem cells results in complete or near-complete conversion to TB. However, a surprise from the analysis was the unexpectedly large number of well-defined clusters and, particularly, the presence of two major groupings (a and b) (**Figure 1A**), each containing distinct syncytioTB, and what appear to be diverse cytoTB components. One particular distinction between grouping a and b was the high expression of the long non-coding RNAs *NEAT1* and *MALAT1* in a, with the latter particularly abundant and, like *CGA*, expressed in almost every nucleus analyzed. Although cluster 9 is distinct from the two groupings, the expression of *NEAT1* and *MALAT1* in this cluster was more similar to grouping b. Other long non-coding RNAs, *H19*, *SPRY4*, and *HOTAIR* that, like *NEAT1* (Gremlich et al., 2014) and *MALAT1* (Tseng et al., 2009), have been linked to placental TB development and placental pathologies (Basak and Ain, 2019), were expressed relatively weakly and lacked meaningful discriminatory power to distinguish clusters.

It was clear that among the resulting clusters recognized in the Seurat analysis were ones that bore strong similarity to syncytioTB (5 and 6) and weaker similarity to extravillousTB



(2 and 3) of first trimester placental TB. The enrichment for extravillousTB was non-significant in cluster 2, and there were only four genes (*PAPPA2*, *GCSH*, *ADAM12*, and *ASPSCR1*) contributing to the enrichment. Nevertheless, cluster 2, although clearly expressing TB markers, did not have similarities to any other placenta cell populations (Figure 3B). Other clusters, for example 4 and 1, that clearly did not conform strongly to either of these two sub-lineages but were sandwiched between them in the Seurat plots (Figure 1A), also expressed classical human TB markers and likely represented forms of cytoTB. We suspect that the upper regions of these two putative cytoTB

clusters represent cells providing a source of syncytiotB and that their lower regions are precursors of extravillousTB. Even the more enigmatic clusters, 7, 8, and 9, display multiple TB markers, but what cell types they represent in the tissue culture colonies from which they were derived remains as yet unclear. Even within the clusters, there is additional heterogeneity. For example, the lower tip of cluster 3 and cluster 9 uniquely express *SOX2* (Supplementary Figure 8), suggesting a shared origin. In scRNAseq studies on human placentas weeks 8 and 24 of gestation (Liu et al., 2018), in first trimester (weeks 6–11) material (Suryawanshi et al., 2018), and even in blastocyst-like structures



engineered from reprogrammed fibroblasts (Liu et al., 2021), it also has become clear that there are multiple distinguishable populations of cytoTBs. In the study of Liu et al. there were also three extravillousTB types (Liu et al., 2018). Moreover, cellular phenotypes as defined in terms of transcript content changed markedly between week 8 and week 24 of gestation (Liu et al., 2018). Perhaps it should be no surprise that multiple kinds of TBs comprise these cell populations derived directly from ESCs, which we have hypothesized likely represent the very early stages of *in vivo* placentation. A planned time-course experiment beginning when differentiation is initiated should enable us to infer cluster origins and interrelationships in a more complete manner.

Our ability to define two distinct syncytioTB clusters (5 and 6) is of particular interest. Do these nuclei represent different kinds of syncytioTB or nuclei with different cytoTB origins in a single type of syncytioTB? The fact that the transcripts for the presumed fusogens, ERVW-1 and ERVV-1, were primarily marker genes for nuclei in cluster 6 and less well expressed in cluster 5 could be evidence that syncytioTB formation requires the interaction of two select populations of presumptive syncytioTBs, one of which expresses the necessary syncytins (ERVs) on their surfaces, the other possibly bearing the appropriate “receptor” factors. Similarly, the absence of *MFSD2A*, which encodes the proposed receptor for ERVFRD-1, another proposed fusogen (Roberts et al., 2021), is puzzling. Perhaps another fusion partner for ERVFRD-1 exists or syncytialization of these TBs occurs without the involvement of ERVFRD-1, which has quite low expression in these preparations of ESC-derived TB.

Another observation of note is that *CGA*, whose translation product partners with one of the CGB isoforms to form the active placental hormone hCG, is expressed in most, if not all, nuclei (Figure 2), yet the protein itself can normally be detected only in syncytioTB and what appear to be its immediate precursor syncytioTB when H1 ESCs are differentiated to syncytioTB (Amita et al., 2013; Yabe et al., 2016). There seems to be three possible explanations: (1) *CGA* mRNA is not translated in cells that are not progressing to syncytioTB; (2) the protein product *CGA* is highly unstable in the context of cells outside the syncytial area and in absence of expression of its partner CGB with which it forms the hCG heterodimer; (3) *CGA* transcripts are unable to exit the nucleus for translation except in syncytioTB. Gene expression changes that are primarily regulated at the protein level have been noted previously in the mouse placenta and other cell types (Abdulghani et al., 2019). In fact, a general rule is that mRNA concentrations are relatively poor guides to protein levels (Ghaemmaghami et al., 2003; Ghazalpour et al., 2011; Schwanhauser et al., 2011). A number of other highly expressed transcripts, for example those for *GABRP* and *VTCN1*, are also abundant across clusters yet their proteins are restricted to emerging syncytioTB in BAP colonies. They are also expressed in placenta where they are localized to villous syncytioTB, primarily to the first trimester of pregnancy (Karvas et al., 2020). Interestingly, expression of the CD274 molecule, perhaps better known as B7-H1 or programmed cell death 1 ligand 1, also appears to be controlled at the translational level in placental TB cells (Holets et al., 2009). Both *VTCN1* and *CD274* bind to

receptors on lymphocytes. In cancer cells and possibly in TB they are considered regulators of immune tolerance (Holets et al., 2006; Schust et al., 2021).

In summary, the BAP model, which we have proposed represents TB associated with the implanting conceptus, reveals a relatively complex picture of TB emergence, including the appearance of at least two kinds of syncytioTB nuclei plus multiple cytoTB populations. Conducting similar analyses at earlier time points should elucidate how these lineages arose and diverged and perhaps provide insights into what occurs during the very earliest stages of human embryonic development. Such studies may also be revealing about how syncytioTB arises, mechanisms of cell fusion, and the possible roles of the non-coding RNAs *MALAT1* and *NEAT1*, and various transcription factors in directing events.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/geo/>, GSE171768. The GitHub repository documenting all the analyses is available at [https://github.com/Tuteja-Lab/BAP.hESC.d8\\_](https://github.com/Tuteja-Lab/BAP.hESC.d8_).

## AUTHOR CONTRIBUTIONS

TK, AS, JZ, NB, and TE performed the experiments. RR and TE conceived the project. RR, TE, TK, AS, and GT designed the study and were responsible for data interpretation. RR, TK, AS, and GT wrote the manuscript with contributions from all authors. RR, TE, GT, and DS sponsored the project through grants. All authors contributed to the article and approved the submitted version.

## FUNDING

The authors' research is supported by grants 1R01HD094937 and 1R21A1145071 from the National Institutes of Health.

## ACKNOWLEDGMENTS

We acknowledge the assistance of Mingyi Zhou and Ellen Kesler of the University of Missouri DNA Core for preparing libraries and generation of sequence. Christopher Bottoms of the University of Missouri Informatics Research Core assisted with initial data processing.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2021.695248/full#supplementary-material>

## REFERENCES

- Abdulghani, M., Song, G., Kaur, H., Walley, J. W., and Tuteja, G. (2019). Comparative Analysis of the Transcriptome and Proteome during Mouse Placental Development. *J. Proteome. Res.* 18, 2088–2099. doi: 10.1021/acs.jproteome.8b00970
- Amita, M., Adachi, K., Alexenko, A. P., Sinha, S., Schust, D. J., Schulz, L. C., et al. (2013). Complete and unidirectional conversion of human embryonic stem cells to trophoblast by BMP4. *Proc. Natl. Acad. Sci. USA* 110, E1212–E1221.
- Basak, T., and Ain, R. (2019). Long non-coding RNAs in placental development and disease. *Non-Coding RNA Invest.* 3, 1–26.
- Boyd, J. D., and Hamilton, W. J. (1970). *The human placenta*. Cambridge, MA: Heffer and Sons.
- Das, P., Ezashi, T., Schulz, L. C., Westfall, S. D., Livingston, K. A., and Roberts, R. M. (2007). Effects of fgf2 and oxygen in the bmp4-driven differentiation of trophoblast from human embryonic stem cells. *Stem Cell Res.* 1, 61–74. doi: 10.1016/j.scr.2007.09.004
- Deglinerti, A., Croft, G. F., Pietila, L. N., Zernicka-Goetz, M., Siggia, E. D., and Brivanlou, A. H. (2016). Self-organization of the in vitro attached human embryo. *Nature* 533, 251–254. doi: 10.1038/nature17948
- Dengler, V. L., Galbraith, M., and Espinosa, J. M. (2014). Transcriptional regulation by hypoxia inducible factors. *Crit. Rev. Biochem. Mol. Biol.* 49, 1–15. doi: 10.3109/10409238.2013.838205
- 10x Genomics (2019). *Chromium Single Cell V(D)J Reagent Kits With Feature Barcoding Technology for Cell Surface Protein, Document Number CG000186 Rev A, (2019, July 25)*. Pleasanton, CA: 10x Genomics.
- Ghaemmaghami, S., Huh, W. K., Bower, K., Howson, R. W., Belle, A., Dephoure, N., et al. (2003). Global analysis of protein expression in yeast. *Nature* 425, 737–741.
- Ghazalpour, A., Bennett, B., Petyuk, V. A., Orozco, L., Hagopian, R., Mungrue, N., et al. (2011). Comparative analysis of proteome and transcriptome variation in mouse. *PLoS Genet.* 7:e1001393. doi: 10.1371/journal.pgen.1001393
- Gremlich, S., Damnon, F., Reymondin, D., Braissant, O., Schittny, J. C., Baud, D., et al. (2014). The long non-coding RNA NEAT1 is increased in IUGR placentas, leading to potential new hypotheses of IUGR origin/development. *Placenta* 35, 44–49. doi: 10.1016/j.placenta.2013.11.003
- Hertig, A. T., Rock, J., and Adams, E. C. (1956). A description of 34 human ova within the first 17 days of development. *Am. J. Anat.* 98, 435–493. doi: 10.1002/aja.1000980306
- Holets, L. M., Carletti, M. Z., Kshirsagar, S. K., Christenson, L. K., and Petroff, M. G. (2009). Differentiation-induced post-transcriptional control of B7-H1 in human trophoblast cells. *Placenta* 30, 48–55. doi: 10.1016/j.placenta.2008.10.001
- Holets, L. M., Hunt, J. S., and Petroff, M. G. (2006). Trophoblast CD274 (B7-H1) is differentially expressed across gestation: influence of oxygen concentration. *Biol. Reprod.* 74, 352–358. doi: 10.1095/biolreprod.105.046581
- Horii, M., Bui, T., Touma, O., Cho, H. Y., and Parast, M. M. (2019). An Improved Two-Step Protocol for Trophoblast Differentiation of Human Pluripotent Stem Cells. *Curr. Protoc. Stem. Cell. Biol.* 50:e96.
- Horii, M., Touma, O., Bui, T., and Parast, M. M. (2020). Modeling human trophoblast, the placental epithelium at the maternal fetal interface. *Reproduction* 160, R1–R11.
- Jain, A., and Tuteja, G. (2019). TissueEnrich: Tissue-specific gene enrichment analysis. *Bioinformatics* 35, 1966–1967. doi: 10.1093/bioinformatics/bty890
- Jain, A., and Tuteja, G. (2021). PlacentaCellEnrich: a tool to characterize gene sets using placenta cell-specific gene enrichment analysis. *Placenta* 103, 164–171. doi: 10.1016/j.placenta.2020.10.029
- Jain, A., Ezashi, T., Roberts, R. M., and Tuteja, G. (2017). Deciphering transcriptional regulation in human embryonic stem cells specified towards a trophoblast fate. *Sci. Rep.* 7:17257.
- Karvas, R. M., McInturf, S., Zhou, J., Ezashi, T., Schust, D. J., Roberts, R. M., et al. (2020). Use of a human embryonic stem cell model to discover GABRP, WFDC2, VTCN1 and ACTC1 as markers of early first trimester human trophoblast. *Mol. Hum. Reprod.* 26, 425–440. doi: 10.1093/molehr/gaaa029
- Liu, X., Tan, J. P., Schröder, J., Aberkane, A., Ouyang, J. F., Mohenska, M., et al. (2021). Modelling human blastocysts by reprogramming fibroblasts into iBlastoids. *Nature* 2021, 1–6.
- Liu, Y., Fan, X., Wang, R., Lu, X., Dang, Y. L., Wang, H., 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
- Lv, B., An, Q., Zeng, Q., Zhang, X., Lu, P., Wang, Y., et al. (2019). Single-cell RNA sequencing reveals regulatory mechanism for trophoblast cell-fate divergence in human peri-implantation conceptuses. *PLoS Biol.* 17:e3000187. doi: 10.1371/journal.pbio.3000187
- Marchand, M., Horcajadas, J. A., Esteban, F. J., McElroy, S. L., Fisher, S. J., and Giudice, L. C. (2011). Transcriptomic signature of trophoblast differentiation in a human embryonic stem cell model. *Biol. Reprod.* 84, 1258–1271. doi: 10.1095/biolreprod.110.086413
- Marsh, B., and Belloch, R. (2020). Single nuclei RNA-seq of mouse placental labyrinth development. *eLife* 9:e60266.
- Ng, Y. H., Zhu, H., and Leung, P. C. (2012). Twist modulates human trophoblastic cell invasion via regulation of N-cadherin. *Endocrinology* 153, 925–936. doi: 10.1210/en.2011-1488
- Ortiz-Barahona, A., Villar, D., Pescador, N., Amigo, J., and del Peso, L. (2010). Genome-wide identification of hypoxia-inducible factor binding sites and target genes by a probabilistic model integrating transcription-profiling data and in silico binding site prediction. *Nucleic Acids Res.* 38, 2332–2345. doi: 10.1093/nar/gkp1205
- Roberts, R. M., Ezashi, T., Schulz, L. C., Sugimoto, J., Schust, D. J., Khan, T., et al. (2021). Syncytins expressed in human placental trophoblast. *Placenta* 2021:006
- Roberts, R. M., Ezashi, T., Sheridan, M. A., and Yang, Y. (2018). Specification of trophoblast from embryonic stem cells exposed to BMP4. *Biol. Reprod.* 99, 212–224. doi: 10.1093/biolre/iy070
- Schneider, V. A., Graves-Lindsay, T., Howe, K., Bouk, N., Chen, H.-C., Kitts, P. A., et al. (2017). Evaluation. *Genome. Res.* 27, 849–864.
- Schulz, L. C., Ezashi, T., Das, P., Westfall, S. D., Livingston, K. A., and Roberts, R. M. (2008). Human embryonic stem cells as models for trophoblast differentiation. *Placenta* 29, S10–S16.
- Schust, D. J., Bonney, E. A., Sugimoto, J., Ezashi, T., Roberts, R. M., Choi, S., et al. (2021). The Immunology of Syncytialized Trophoblast. *Int. J. Mol. Sci.* 22, 1767. doi: 10.3390/ijms22041767
- Schwanhauser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., et al. (2011). Global quantification of mammalian gene expression control. *Nature* 473, 337–342. doi: 10.1038/nature10098
- Shahbazi, M. N., Jedrusik, A., Vuoristo, S., Recher, G., Hupalowska, A., Bolton, V., et al. (2016). Self-organization of the human embryo in the absence of maternal tissues. *Nat. Cell. Biol.* 18, 700–708. doi: 10.1038/ncb3347
- Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M. III, et al. (2019). Comprehensive integration of single-cell data. *Cell* 177, 1888–1902.
- Suryawanshi, H., Morozov, P., Straus, A., Sahasrabudhe, N., Max, K. E. A., Garzia, A., et al. (2018). A single-cell survey of the human first-trimester placenta and decidua. *Sci. Adv.* 4:eaau4788. doi: 10.1126/sciadv.aau4788
- Telugu, B. P., Adachi, K., Schlitt, J. M., Ezashi, T., Schust, D. J., Roberts, R. M., et al. (2013). Comparison of extravillous trophoblast cells derived from human embryonic stem cells and from first trimester human placentas. *Placenta* 34, 536–543. doi: 10.1016/j.placenta.2013.03.016
- Tseng, J. J., Hsieh, Y. T., Hsu, S. L., and Chou, M. M. (2009). Metastasis associated lung adenocarcinoma transcript 1 is up-regulated in placenta previa increta/percreta and strongly associated with trophoblast-like cell invasion in vitro. *Mol. Hum. Reprod.* 15, 725–731. doi: 10.1093/molehr/gap071
- Vento-Tormo, R., Efreanova, M., Botting, R. A., Turco, M. Y., Vento-Tormo, M., Meyer, K. B., et al. (2018). Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature* 563, 347–353.
- West, R. C., Ming, H., Logsdon, D. M., Sun, J., Rajput, S. K., Kile, R. A., et al. (2019). Dynamics of trophoblast differentiation in peri-implantation-stage human embryos. *Proc. Natl. Acad. Sci. USA* 116:22635. doi: 10.1073/pnas.1911362116
- Wilcox, A. J., Weinberg, C. R., O'Connor, J. F., Baird, D. D., Schlatterer, J. P., Canfield, R. E., et al. (1988). Incidence of early loss of pregnancy. *N. Engl. J. Med.* 319, 189–194.
- Xu, R. H., Chen, X., Li, D. S., Li, R., Addicks, G. C., Glennon, C., et al. (2002). BMP4 initiates human embryonic stem cell differentiation to trophoblast. *Nat. Biotechnol.* 20, 1261–1264. doi: 10.1038/nbt761

- Yabe, S., Alexenko, A. P., Amita, M., Yang, Y., Schust, D. J., Sadovsky, Y., et al. (2016). Comparison of syncytiotrophoblast generated from human embryonic stem cells and from term placentas. *Proc. Natl. Acad. Sci. USA* 113, E2598–E2607.
- Yang, Y., Adachi, K., Sheridan, M. A., Alexenko, A. P., Schust, D. J., Schulz, L. C., et al. (2015). Heightened potency of human pluripotent stem cell lines created by transient BMP4 exposure. *Proc. Natl. Acad. Sci. USA* 112, E2337–E2346.
- Zheng, G. X., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., et al. (2017). Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* 8:14049. doi: 10.1038/ncomms14049
- Zhou, F., Wang, R., Yuan, P., Ren, Y., Mao, Y., Li, R., et al. (2019). Reconstituting the transcriptome and DNA methylome landscapes of human implantation. *Nature* 572, 660–664. doi: 10.1038/s41586-019-1500-0
- Zhou, J., West, R. C., Ehlers, E. L., Ezashi, T., Schulz, L. C., Roberts, R. M., et al. (2021). Modeling human peri-implantation placental development and function. *Biol. Reprod.* 105, 40–51. doi: 10.1093/biolre/ioab080

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Khan, Seetharam, Zhou, Bivens, Schust, Ezashi, Tuteja and Roberts. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.