Here and there a trophoblast, a transcriptional evaluation of trophoblast cell models.

Brian Joseph Cox (b.cox@utoronto.ca)
University of Toronto https://orcid.org/0000-0001-7146-6041
Kendra Naismith
University of Western Ontario: Western University

Research Article

Keywords:

Posted Date: March 8th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1405255/v1

Abstract

A recent explosion of methods to produce human trophoblasts and stem cells (hTSC) is fuelling a renewed interest in this tissue. The trophoblast is critical to reproduction by facilitating implantation, maternal physiological adaptations to pregnancy and the growth of the fetus through transport of nutrients between the mother and fetus. More broadly, the trophoblast has phenotypic properties that make it of interest to other fields. Its angiogenic and invasive properties are similar to tumours and could identify novel drug targets, and its ability to regulate immunological tolerance of the allogenic fetus could lead to improvements in transplantations. Within this review, we integrate and assess transcriptomic data of cell-based models of hTSC alongside in vivo samples to identify the utility and applicability of these models. We also integrate single-cell RNA sequencing data sets of human blastoids, stem cells and embryos to identify how these models may recapitulate early trophoblast development.

Introduction

The study of human development is challenging, but the investigation of the extraembryonic tissue of the trophoblast is "…a riddle, wrapped in a mystery, inside an enigma." (Winston Churchill, 1939). Fetal development has benefited from human embryonic stem cells (hESC) as a model [1]. These stem cell models greatly expanded our understanding of cell fate decisions and developmental lineages and led to innovations such as regenerative medicine and cell-based therapies[2]. The other half of development is the placenta, specifically the trophoblast lineage's extraembryonic cells. Our knowledge of this organ and its specific cell types lag behind fetal or somatic tissues[3], and potentially stem cell models may help accelerate research.

While mouse trophoblast stem cells (TSC) were generated in the previous millennium[4], the derivation of human trophoblast stem cells (hTSC) has been sought for several decades. Initial attempts were unsuccessful using conditions similar to the derivation of mouse TSC[5]. Recently, human trophoblast stem cells were derived from human blastocysts and first-trimester placenta cells[6]. This bona fide stem cell of the trophoblast, analogous to the trophoblast stem cells derived from mouse embryos, opens opportunities to study human reproduction. These cell lines are grown in 2D cultures and can be differentiated into syncytium and extravillous trophoblast. Two other groups recently published methods for deriving and maintaining trophoblast organoids derived from first-trimester trophoblast [7, 8]. These organoids can be grown in 3D culture and differentiated into syncytium and EVT-like cells.

Before deriving bona fide hTSC and organoids, methods utilizing BMP treatment to derive trophoblast were actively investigated by many research groupss[9–12]. This method is met with controversy by some groups proposing that the cells were not trophoblast but mesoderm or amnion-like cells[13–15]. The primary evidence for trophoblast presented by most publications was the presence of various trophoblast marker genes and phenotypes. mRNA and protein expression such as GATA3, TFAP2C and KRT7, and the production of estradiol and bHCG (CGA/B) hormones are considered definitive markers of trophoblast[16, 17]. Another hallmark was the ability of the cells to generate a syncytium[12, 17]. One line

of reasoning against BMP4 mediated methods of trophoblast derivation is that human ESC were cultured as primed cells that are more representative of a post-implantation stage, long after trophoblast cell fate decisions were determined. hESC can be transited to an earlier developmental state called naïve that is closer to the preimplantation stage of the blastocyst inner cell mass[18]. Interestingly, methods to create human trophoblast from naïve stage embryonic stem cells were established using the media conditions similar to hTSC derivation and maintenance[19–21].

Shockingly, human naïve embryonic stem cells are reported to generate blastocyst-like structures called blastoids[22–25]. By seeding a large number of cells (>100) into suspension under specific conditions, the aggregates differentiate into primitive endoderm/hypoblast, trophoblast and epiblast while maintaining spatial organization and a cavitated blastocoel. Cell isolates from blastoids were also used to derive hESC, hTSC and hypoblast cell lines.

Many methods appear to enable hESC to access early developmental states, including the trophoblast. This review aims to use available transcriptomic data to establish the similarity of hESC and trophoblast-derived models to in vivo sources. We also investigate the blastoids as a model of embryo cell fate specification and determination, specifically if blastoids achieve trophoblast cell fates through embryo-like processes.

Results

hESC to trophoblast conversion

We searched the Gene Expression Omnibus for datasets on hTSC and trophoblast derivation from hESC made available between July 1, 2018, and June 8, 2021, that were Homo sapiens. We identified 94 studies and selected the 20 data sets on trophoblast derivation or containing established trophoblast cell lines and differentiated cell types. We removed studies using only choriocarcinoma lines (BeWo, JEG3) or transfection-derived cell lines (HTR8/SVneo). Single-cell sequencing data sets on blastoids and cell models were assessed in a separate analysis. ArrayExpress was searched with the same criteria, yielding four more results with two meeting criteria for this study. In total, there were two microarray and 12 bulk RNA sequencing data sets. All data were processed from raw files using a consistent bioinformatics pipeline. Only one study was removed after failing to pass quality control. In the end, we proceeded with two microarray and 11 bulk RNA sequencing data sets(**Supplemental Table 1**) comprising 308 samples (**Supplemental Table 2**). For convenience, we identified embryonic and pluripotent stem cells as ESC.

Standard approaches of data integration and batch correction for bulk RNA sequencing did not work as many studies continued to show batch separation from each other (data not shown). Given the large number of samples (>300) we applied data integration methods from the R package Seurat, typically meant for single cell sequencing. After data integration, we applied a Uniform Manifold Approximation and Projection (UMAP) of clustered data. We observed a distribution and mixing of samples by study ID, indicating that experimental batch effects were likely removed or minimized (**Figure 1A**). We observed four data groupings composed of seven statistical clusters (**Figure 1B**). By convention, clusters are

numbered starting at 0 in order of population size. We observed that clusters 0, 1, 4 and 6 belong to the largest block of samples and clusters 2, 3 and 5 show a high degree of separation from the large block. Clusters 2 and 3 are close together, while 5 is remote (**Figure 1 B**). The organization of the clusters suggests the large group is closely related (clusters 0, 1, 4 and 6), clusters 2 and 3 are similar, and cluster 5 is unique (**Figure 1B**).

Assessing the clusters for co-grouping in vivo samples found that the placental samples (isolated trophoblasts and whole villous) are in cluster 3 along with the trophoblast organoids (**Figure 1C**). Cluster 2 contains hTSC lines derived from the placental trophoblast (**Figure 1C**). Associated with clusters 2 and 3 are samples derived from naïve human embryonic stem cells, suggesting that these are closest to in vivo and cell lines directly derived from trophoblast (**Figure 1C; Table 1 and 2**). This also indicates that hTSC and organoids are different, possibly due to differences in developmental states, media and culturing methods (2D adherent vs 3D matrix embedded).

Investigating media conditions reveals significant separation and correlation to cluster memberships (**Figure 1D**). Of the clusters without in vivo trophoblast, the largest block of cells belonging to clusters 0, 1, 4 and 6 contained undifferentiated ESC and various differentiated derivatives using protocols to generate trophoblast, including the BAP or BMP4 related protocols (**Figure 1D, Table 3**). The separation of these clusters from the trophoblast containing clusters 2 and 3 indicates that their bulk transcriptional profile is not similar to in vivo trophoblast, organoids or hTSC directly derived from the trophoblast. Clusters 2 and 3 containing the in vivo trophoblast samples and bona fide cell models primarily comprise trophoblast stem cell media conditions. Cluster 5 contains only a set of samples derived from ESC using trophoblast and ACE media.

Gene enrichments to clusters

We investigated the typical marker panel of genes used to evaluate trophoblast identity in cell-based differentiation experiments. Embryonic epiblast (ESC) markers such as NANOG and POU4F1 (OCT3-4) strongly label the ESC cluster 4 (**Figure 2**). A typical trophoblast marker panel of KRT7, CDX2, HAND1, GCM1, TFAP2C, ERVW-1, PGF, GATA3, ITGA2 and PAPPA produces irregular patterns with overlapping expression to clusters 0, 2, 3, 5 and 6 (**Figure 2**). The lack of specificity to the trophoblast containing clusters 2 and 3 questions the validity of these markers in the context of cell-based differentiation experiments.

We tested for differential expression to identify candidate markers that can distinguish bona fide trophoblast cells and models from ESC derivatives. We used a one versus all comparison and set significance at an FDR<0.05 and log2(fold change) > 0.5. We found 100s of enriched genes to each cluster with many genes displaying precise expression patterns to clusters 2, 3 or both clusters combined (**Figure 2, Table 4**). While untested, these candidate markers may be helpful to assess the quality of novel cell models relative to in vivo and bona fide cell sources.

To address the identities of the different clusters and their functional qualities, we applied gene set enrichment to the top significantly enriched genes in each cluster. The significant gene ontologies were grouped by overlapping gene sets and a consolidated title was generated using the Cytoscape plug in Enrichment Map (**Supplemental Table 3**). A potential problem with the human Gene Ontology is under-representing embryonic and especially extraembryonic structures and pathways.

Cluster 2, which contained the hTSC and ESC-derived cell lines, displayed enrichment to genes involved in blood vessel formation (angiogenesis), immune functions related to leukocytes and mast cells, and processes for locomotion and migration. In contrast, Cluster 3, which comprised organoids, villous and cytotrophoblast samples, and ESC-derived cells, was enriched in terms related to organism growth, morphogenesis and growth, signalling related to steroids, and stress protein kinase cascades. These ontological differences may relate to developmental stages represented by each model. The cluster 2 ontology enrichment contains early phenotypes of primary trophoblast post-implantation, such as the attraction of angiogenic processes and promotion of immune tolerance[26, 27]. Cluster 3 enriched ontologies likely represent processes of the established placental surface that transport nutrients and are elaborating the villous structures[28–30]. No terms specific to trophoblast or placenta were enriched in either cluster.

Cluster 5 was enriched in very few human gene ontologies. Most were related to cell movement regulation, inflammatory response and muscle structure development. The isolated clustering away from bonified trophoblast samples suggests that cluster 5 is not a trophoblast.

The remaining clusters (0, 1, 4 and 6) include the undifferentiated ESC samples (cluster 4) and differentiated samples from various methods. This sample cluster's connectivity suggests progressive development in the epiblast lineage. Chronologically, cluster 4 represents the undifferentiated ESC, containing both primed and naïve. Next, cluster 1 is highly enriched in terms related to neurogenesis and neuron projections. Cluster 6 showed a continued and expanded enrichment of ontology terms related to neurogenesis, including synapse formation and patterning of the telencephalon and forebrain. Additionally, cluster 6 was enriched in ontology terms related to cardiac muscle development and calcium mediated muscle action.

Cluster 0, the outermost group, is composed of endpoint BMP4 or BAP-based differentiated protocol samples (Days 5-10; **Table 3**). This cluster contained the largest number of differentially expressed genes and enriched ontologies. One of the biggest groupings of ontologies related to morphogenesis and branching in ureteric development contained 117 different ontology terms. Expectedly, a grouping of ontology terms related to BMP signalling is present. Continued enrichment of neurogenesis and cardiac developmental terms is also present. Of significant interest is the presence of enriched gene ontologies related to syncytium and plasma membrane fusion. Within the enriched ontologies are cardiac muscle cell fusion pathways[31] and share expression with trophoblast fusion[32]. Importantly, placental-related ontology terms were found as enriched, but these were significantly intersected with maternal tissue, such as decidua. These samples do not cluster with the hTSC, organoids or primary cytotrophoblasts, so they

are unlikely pure or enriched populations of trophoblast cells. However, we cannot exclude the possibility of trophoblast cell types being mixed with other mesodermal and neuronal cell types. Single-cell sequencing could resolve this complex mixture as it is important to determine if trophoblast cell fate can be reached from the naïve and primed ESC state.

Embryos, Blastoids and Stem Cells, oh my!

The most recent development in the derivation of the trophoblast is the production of blastoids from ESC, which appear to generate the trophectoderm, hypoblast (primitive endoderm) epiblast components in a cavitated structure[22–25]. A possible concern is that the TE structure is amnion, as there is a degree of gene expression overlap between these cell types and both grow as a monolayer tight epithelium. To determine if blastoids produce trophoblast similar to the blastocyst, we combined three published single-cell sequencing data sets on human blastoids [23–25]. We integrated these with a single-cell sequencing data set on human embryo development spanning totipotency to blastocyst formation[33]. Additionally, we integrated single-cell data of primed and naïve hESC, human dermal fibroblasts and hTSC derived from blastocysts.

In total, 6160 cells passed quality control filters and were statistically organized into 12 clusters and displayed in a UMAP (**Figure 4A**). The clusters showed good integration of independent data sets with no strong batch effects (**Figure 4B**). Clusters 6 and 11 represent cells of embryonic days 3 and 4 of human development, respectively (**Figure 4C**). Neither cluster contained significant numbers of cells from blastoids or cell lines (**Figure 4C**), suggesting that likely none of the published blastoid conditions achieve a totipotency state. Day 5 human embryos and 24-60-hour blastoids produced cells integrated to clusters that spanned epiblast, hypoblast (primitive endoderm) and trophoblast (**Figure 4C** and **D**). Additionally, the three blastoid data sets produced clusters of cells (clusters 4, 8 and 9) not generated by human embryos (**Figure 4C**).

As the focus of this review is the trophoblast, we will only briefly address the epiblast and hypoblast-associated clusters. Using known markers, the epiblast lineages spanned clusters 0, 2 and 10 and contained blastoid, human embryo and naïve ESC cells in each (**Figure 4D and E**). Cluster 8 is composed of cells from naïve ESC and blastoids but does not contain embryo-derived epiblast (**Figure 4D**). The primed ESCs in cluster 4 contain some blastoid cells but not naïve ESC or human embryo samples (**Figure 4D**). This suggests that blastoids generate appropriate embryo cell types (clusters 0 and 2), although the blastoid epiblast appears dominated by cluster 0 while the embryo is cluster 2. This proportion is similar to that observed in the naïve cell culture sample.

Cluster 5 is the hypoblast and contains cells of the blastoids and human embryo but not naïve or primed hESC (**Figure 4A, C and F**). This suggests that blastoid differentiation creates a developmental state that can access hypoblast fate without using totipotency.

The trophoblast from the human embryo falls into multiple clusters (1, 3 and 7, **Figure 4A, C and G**). Cluster 3 trophoblast expressed marker genes including CDX2, HAND1 and TFAP2C. Cluster 2 trophoblast

expressed gene associated with later developmental phenotypes, including GCM1, KRT7, the hormones PGF and CGA and the syncytialization marker ERVW-1 (Syncytin). The human TSC associate completely with cluster 7 along with cells from the blastoids of Yu and colleagues [24] and Yanagida and colleagues [23] (**Figure 4C**). The blastoids of Kagawa and colleagues [25] did not create cells that belong to cluster 7. Chronologically, the human embryo first produced cells of cluster 3 followed by 2 and 7 (**Figure 4C**). The blastoid time-series data from Kagawa and colleagues [25] similarly produced cluster 3 cells followed by 2 but did not produce cells of cluster 7 as stated above (**Figure 4C**). Time series data from Yanagida and colleagues [23] produced clusters 3 and 2 concurrently, followed by cluster 7 (**data not shown**). Our analysis indicates that blastoids produce a variety of trophoblast subtypes similar to those of the embryo but without access to totipotency. There is also a chronological similarity to trophoblast development between blastoids and embryos. There are differences in cellular proportions between the different blastoid methods although these are subject to technical bias and error.

Associated with the trophoblast clusters in cluster 9 produced by blastoids from Kanagawa and Yu [24, 25] and a small number of cells from naïve hESC (**Figure 4 A, C and D**). The absence of these cells from the human embryo suggests inappropriate development within the blastoids and a propensity of naïve hESC to spontaneously produce this cell type (**Figure 4C and D**). Overrepresentation analysis of human ontology using genes significantly enriched to cluster 9 (one versus all, FDR<0.05, log2(fold change) >0.5) revealed biological processes involved in adhesion and cell spreading, actin and cytoskeleton and epithelium morphogenesis that fit with the formation of a monolayered tight epithelium. Also present are terms related to muscle development and smooth muscle proliferation. Curiously, this is similar to ontological enrichments observed in cluster 5 from the ESC to trophoblast conversion models (**Figure 1 B**).

There are no GO terms annotated to the amnion, so this structure could not be detected using standard ontology databases. Markers of human amnion are not well known, and the few available are also markers of trophoblast (e.g., CDX2 and GATA3). A study of induced amnion from primed hESC provided a list of differentially expressed genes between amnion-like epithelial cells and hESC[15]. In a different approach, we imported validated amnion gene markers from the mouse EMAPA ontology (http://www.obofoundry.org/ontology/emapa.html). Of the 150 genes annotated with expression to the amnion in mice, 41 showed significant differential expression between the clusters. In a heat map of these 41 markers and the data set of merged single cells organized by cluster, 21 genes were specifically increased in cluster 9 relative to other clusters and several were significant two cluster 9 and other clusters (**Figure 5**). This is strong evidence of a propensity for naïve cells to make amnion under some blastoid conditions.

Discussion

The creation of cell models of the human trophoblast is fuelling a new interest in trophoblast tissue and the placenta from those traditionally not involved in placental and trophoblast research. This is an excellent opportunity to forge new collaborations between stem cell, tissue engineering and

trophoblast/reproductive biology research groups. However, caution is needed in this new burgeoning field, and a consensus is required on what constitutes trophoblast and how to assess cell lines and tissues derived from stem cells. Of concern is the propensity of cells in culture to find viable but not in vivo equivalent states[36]. Some recent articles have begun to address this[34, 35]. Our findings show that context is critical in interpreting and selecting markers. While classical trophoblast markers perform poorly in cell-based ESC to TSC differentiation/conversion models, they worked very well in blastoid models to identify trophoblast from epiblast hypoblast and amnion.

Lessons can be learned from past debates after creating human ESC and induced pluripotent stem cells on how to establish what constituted a “true” stem cell[37, 38]. Human cell models are challenged in contrast to the mouse cell lines that can be transplanted and assessed in vivo for their developmental potential. However, even in a “gold standard” host embryo transplantation assay, cells can be misinterpreted based solely on their location within the embryo or placenta[35]. Single markers can be misleading in cell culture, and even tissue definitive markers need to be reassessed in different experimental conditions outside of their in vivo context.

Our analysis of the human ESC to TSC derivation methods indicates that only some protocols produce large numbers of trophoblast cells. This was not an obvious conclusion. Our analysis is afforded a significant benefit of contrasting a variety of published data sets and methods with the recent creation of “gold standard” trophoblast stem cells lines and organoids derived directly from the trophoblast. While some methods produced cells with similar classical marker gene expression patterns, these same cells are revealed not to be transcriptionally similar to trophoblast when assessed by genome-wide methods. We present a consolidated data set of candidate markers to test for efficacy in cell type discrimination. However, including bona fide trophoblast cell models or data is likely a best practice to follow in cell culture experiments.

Blastoids are an exciting new cell culture model. As individually reported, our integrated analysis of three blastoid single-cell data sets indicated that each method produced cell types highly similar to the embryo [23, 25, 39]. Significantly, blastoids did not recapitulate early totipotent stages of embryogenesis yet somehow access trophoblast and hypoblast genes and developmental programs. We also observed embryo inappropriate and missing cell types within the epiblast and trophoblast lineages of blastoids. We found three time-dependent trophoblast populations produced by the embryo but that not each blastoid method produced them all. An open question is how these multiple trophoblast populations of the embryo and blastoids are spatial arranged, such as polar regions analogous to the mouse or other patterns. Do individual blastoids contain all of these cell types?

We observed that two blastoid models produced cells showing transcriptional profiles of amnion that group with the trophoblast clusters. Early trophoblast markers and amnion share common regulators such as CDX2 and GATA3. Given the connection of the trophoblast and amnion clusters, it is tempting to speculate on a developmental relationship. However, it is possible that this is not a normal developmental trajectory. As blastoids do not appear to use totipotency to reach trophoblast and hypoblast, a potential

explanation is that blastoids use amnion as a gateway. We observed that a few undifferentiated naïve hESC could produce rare cells clustered with blastoid amnion. The strong signalling environment of blastoid culture media may trans-differentiate amnion progenitors into trophoblast. However, one blastoid condition did not show the presence of amnion cells[23]. These blastoid culture conditions are possibly better at restricting amnion and channelling cells into the trophoblast. Alternatively, this data set contained fewer cells, and amnion may be missing due to under sampling. Recognition of these unintended cell types should help improve methods to better model in vivo development.

While we did not directly assess data on implantation models, care should be taken not to overinterpret decidualization. In the mouse, decidualization can be induced with lectin-coated beads and other growth factors[40, 41] in the complete absence of trophoblast. Better molecular characterization of implantation is needed to understand and differentiate generalized versus trophoblast-mediated decidual responses, such as models incorporating resident immune cell populations.

Important questions remain to be answered. We need a consensus on a panel of markers to assess a trophoblast cell's identity establish assays to determine a trophoblast stem cell's developmental capacity and functional similarity to in vivo counterparts. Lastly, how do epiblast cells access the extraembryonic lineages without using totipotency?

Methods

Identify data sets

NCBI's GEO was searched using the terms "trophoblast" and "placenta organoid". Data sets were filtered for a deposit date between July 1, 2018, and June 8, 2021, Homo sapiens, "expression profiling by array", and "expression profiling by high throughput sequencing". Samples with or of choriocarcinoma lines (BeWo, JEG3) and transfection-derived cell lines (HTR8/SVneo) were excluded. Array express was searched with the same criteria.

Data set processing

To ensure uniform analysis, bulk RNA sequencing data was obtained as raw FASTQ files from the Sequence Read Archive using NCBI's SRA toolkit Count tables were generated by aligning FASTQ files to the GRCh38.p12 human genome build using Kallisto pseudo alignment and importing the abundance files into R with Bioconductor's tximport by gene. Microarray raw data were directly downloaded from Array Express or GEO using Bioconductor's query package.

Dataset integration

Seurat was selected for downstream analysis for bulk RNA sequencing and microarray data due to its strength in integrating, clustering, and visually representing data. The relatively small sample size of bulk RNA sequencing and microarray data required manual integration of data sets before import and integration in Seurat. Bulk RNA sequencing data sets containing similar sample types were first filtered with limma's filterByExpr and normalized with calcNormFactors. We then merged data sets containing similar samples using the ribiosNGS package mergeDGEList function. Expression data were then log-transformed using limma's zoom function. The variation introduced due to the inclusion of different experiments was accounted for using RemoveBatchEffect with groups of highly similar samples across experiments assigned to eliminate covariate batches. Once four larger datasets were created, they were large enough to integrate with Seurat FindIntegrationAnchors and IntegrateData. The remainder of the analysis in Seurat involved scaling data, finding clusters, differential expression and visualizing the clusters with UMAP.

Single-cell sequencing

Count tables for single-cell sequencing were directly downloaded were available or generated from FASTQ files using HTSEQ2 aligned to the human genome. First, individual data sets were processed using Seurat package functions to clean and normalize data. Single-cell sequencing Seurate objects were integrated with Seurat FindIntegrationAnchors and IntegrateData. The Combined object was used for the remainder of the analysis in Seurat involved scaling data, finding clusters, differential expression and visualizing the clusters with UMAP.

Declarations

Acknowledgements

We thank Dr. Frances Wong for assistance with the integration of processing RNA-sequencing data.

Supplemental Table 1. List of all studies used in the analysis.

Supplemental Table 2. List of all samples used in the analysis.

Supplemental Table 3. Enrichment Map Ontologies for terminal clusters 0, 2, 3 and 5. Ontologies are based on the word frequency of nodes with significant overlap in gene members. All nodes were significant at an FDR corrected p-values of <0.1.

Conflict of interest declaration

The authors declare no conflicts of interest. The funding agencies played no role in the direction of the research.

Contributions

BC conceived the project. KN and BC analyzed data, generated figures and tables and wrote the manuscript.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

All authors agreed to the manuscript and conclusions.

Availability of data and material

Please see supplemental tables 1 and 2 for lists of accession IDs for data sets and specific samples used in this analysis.

Funding

BC is supported in part by a Tier II Canada Research Chair in maternal-fetal medicine. KN was supported in part by the Undergraduate Research Opportunity Fund from the Department of Physiology at the University of Toronto.

Authors' information (optional)

Not applicable.

References

1. Shi Y, Inoue H, Wu JC, Yamanaka S (2017) Induced pluripotent stem cell technology: A decade of progress. Nat Rev Drug Discov 16:115–130. https://doi.org/10.1038/nrd.2016.245

2. Aghazadeh Y, Poon F, Sarangi F, et al (2021) Microvessels support engraftment and functionality of human islets and hESC-derived pancreatic progenitors in diabetes models. Cell Stem Cell 28:1936-1949.e8. https://doi.org/10.1016/j.stem.2021.08.001

3. Mercuri ND, Cox BJ (2021) Meta-Research: A Poor Research Landscape Hinders the Progression of Knowledge and Treatment of Reproductive Diseases. bioRxiv 2021.11.16.468787. https://doi.org/10.1101/2021.11.16.468787

4. Tanaka S, Kunath T, Hadjantonakis a K, et al (1998) Promotion of trophoblast stem cell proliferation by FGF4. Science 282:2072–5

5. Kunath T, Yamanaka Y, Detmar J, et al (2014) Developmental differences in the expression of FGF receptors between human and mouse embryos. Placenta 35:1079–1088. https://doi.org/10.1016/j.placenta.2014.09.008

6. Okae H, Toh H, Sato T, et al (2017) Derivation of Human Trophoblast Stem Cells. Cell Stem Cell 22:50-63.e6. https://doi.org/10.1016/j.stem.2017.11.004

7. Sheridan MA, Fernando RC, Gardner L, et al (2020) Establishment and differentiation of long-term trophoblast organoid cultures from the human placenta. Nat Protoc 15:3441–3463. https://doi.org/10.1038/s41596-020-0381-x

8. Haider S, Meinhardt G, Saleh L, et al (2018) Self-Renewing Trophoblast Organoids Recapitulate the Developmental Program of the Early Human Placenta. Stem Cell Reports 11:. https://doi.org/10.1016/j.stemcr.2018.07.004

9. Xu R-H, Chen X, Li DS, et al (2002) BMP4 initiates human embryonic stem cell differentiation to trophoblast. Nat Biotechnol 20:1261–4. https://doi.org/10.1038/nbt761

10. Amita M, Adachi K, Alexenko a. P, et al (2013) Complete and unidirectional conversion of human embryonic stem cells to trophoblast by BMP4. Proc Natl Acad Sci 1–10. https://doi.org/10.1073/pnas.1303094110

11. Krendl C, Shaposhnikov D, Rishko V, et al (2017) GATA2/3-TFAP2A/C transcription factor network couples human pluripotent stem cell differentiation to trophectoderm with repression of pluripotency. Proc Natl Acad Sci 114:E9579–E9588. https://doi.org/10.1073/pnas.1708341114

12. Li Y, Moretto-Zita M, Soncin F, et al (2013) BMP4-directed trophoblast differentiation of human embryonic stem cells is mediated through a ΔNp63+ cytotrophoblast stem cell state. Development 140:3965–76. https://doi.org/10.1242/dev.092155

13. Bernardo ASASAS, Faial T, Gardner L, et al (2011) BRACHYURY and CDX2 Mediate BMP-Induced Differentiation of Human and Mouse Pluripotent Stem Cells into Embryonic and Extraembryonic Lineages. Cell Stem Cell 9:144–155. https://doi.org/10.1016/j.stem.2011.06.015

14. Roberts RM, Loh KM, Amita M, et al (2014) Differentiation of trophoblast cells from human embryonic stem cells: to be or not to be? Reproduction 147:D1-12. https://doi.org/10.1530/REP-14-0080

15. Zheng Y, Xue X, Shao Y, et al (2019) Controlled modelling of human epiblast and amnion development using stem cells. Nature 573:421–425. https://doi.org/10.1038/s41586-019-1535-2

16. Hayashi Y, Furue MK, Tanaka S, et al (2010) BMP4 induction of trophoblast from mouse embryonic stem cells in defined culture conditions on laminin. In Vitro Cell Dev Biol Anim 46:416–30. https://doi.org/10.1007/s11626-009-9266-6

17. Harun R, Ruban L, Matin M, et al (2006) Cytotrophoblast stem cell lines derived from human embryonic stem cells and their capacity to mimic invasive implantation events. Hum Reprod 21:1349–58. https://doi.org/10.1093/humrep/del017

18. Theunissen TW, Friedli M, He Y, et al (2016) Molecular Criteria for Defining the Naive Human Pluripotent State Resource Molecular Criteria for Defining the Naive Human Pluripotent State. 502–515

19. Cinkornpumin JK, Kwon SY, Guo Y, et al (2020) Naive Human Embryonic Stem Cells Can Give Rise to Cells with a Trophoblast-like Transcriptome and Methylome. Stem Cell Reports 15:198–213. https://doi.org/10.1016/j.stemcr.2020.06.003

20. Dong C, Beltcheva M, Gontarz P, et al (2020) Derivation of trophoblast stem cells from naïve human pluripotent stem cells. Elife 9:1–26. https://doi.org/10.7554/eLife.52504

21. Liu X, Ouyang JF, Rossello FJ, et al (2020) Reprogramming roadmap reveals route to human induced trophoblast stem cells. Nature 586:101–107. https://doi.org/10.1038/s41586-020-2734-6

22. Li R, Zhong C, Yu Y, et al (2019) Generation of Blastocyst-like Structures from Mouse Embryonic and Adult Cell Cultures. Cell 179:687-702.e18. https://doi.org/10.1016/j.cell.2019.09.029

23. Yanagida A, Spindlow D, Nichols J, et al (2021) Naive stem cell blastocyst model captures human embryo lineage segregation. Cell Stem Cell 28:1016-1022.e4. https://doi.org/10.1016/j.stem.2021.04.031

24. Yu L, Wei Y, Duan J, et al (2021) Blastocyst-like structures generated from human pluripotent stem cells

25. Kagawa H, Javali A, Khoei HH, et al (2022) Human blastoids model blastocyst development and implantation. Nature 601:600–605. https://doi.org/10.1038/s41586-021-04267-8

26. Zygmunt M, Herr F, Münstedt K, et al (2003) Angiogenesis and vasculogenesis in pregnancy. Eur J Obstet Gynecol Reprod Biol 110:10–18. https://doi.org/10.1016/S0301-2115(03)00168-4

27. Christiansen OB (2013) Reproductive immunology. Mol Immunol 55:8–15. https://doi.org/10.1016/j.molimm.2012.08.025

28. Carter AM (2012) Evolution of placental function in mammals: The molecular basis of gas and nutrient transfer, hormone secretion, and immune responses. Physiol Rev 92:1543–1576. https://doi.org/10.1152/physrev.00040.2011

29. Soncin F, Natale D, Parast MM (2014) Signaling pathways in mouse and human trophoblast differentiation: a comparative review. Cell Mol Life Sci 72:1291–1302. https://doi.org/10.1007/s00018-014-1794-x

30. Knöfler M, Pollheimer J (2013) Human placental trophoblast invasion and differentiation: a particular focus on Wnt signaling. Front Genet 4:190. https://doi.org/10.3389/fgene.2013.00190

31. Doherty KR, Cave A, Davis DB, et al (2005) Normal myoblast fusion requires myoferlin. Development 132:5565–75. https://doi.org/10.1242/dev.02155

32. Robinson JM, Ackerman WE, Behrendt NJ, Vandre DD (2009) While dysferlin and myoferlin are coexpressed in the human placenta, only dysferlin expression is responsive to trophoblast fusion in model systems. Biol Reprod 81:33–9. https://doi.org/10.1095/biolreprod.108.074591

33. Petropoulos S, Deng Q, Panula SP, et al (2016) Single-cell RNA-seq reveals lineage and X-chromosome dynamics in human preimplantation embryos. Cell 165:1012–1026

34. Lee CQEE, Gardner L, Turco M, et al (2016) What Is Trophoblast? A Combination of Criteria Define Human First-Trimester Trophoblast. Stem Cell Reports 6:257–272. https://doi.org/10.1016/j.stemcr.2016.01.006

35. Posfai E, Schell JP, Janiszewski A, et al (2021) Evaluating totipotency using criteria of increasing stringency. Nat Cell Biol 23:49–60. https://doi.org/10.1038/s41556-020-00609-2

36. Tonge PD, Corso AJ, Monetti C, et al (2014) Divergent reprogramming routes lead to alternative stem-cell states. Nature 516:192–197. https://doi.org/10.1038/nature14047

37. Maherali N, Hochedlinger K (2008) Guidelines and Techniques for the Generation of Induced Pluripotent Stem Cells. Cell Stem Cell 3:595–605. https://doi.org/10.1016/j.stem.2008.11.008

38. Hussein SMI, Elbaz J, Nagy AA (2013) Genome damage in induced pluripotent stem cells: Assessing the mechanisms and their consequences. BioEssays 35:152–162. https://doi.org/10.1002/bies.201200114

39. Yu L, Wei Y, Duan J, et al (2021) Blastocyst-like structures generated from human pluripotent stem cells. Springer US

40. Sakoff A (1993) Cell Reaction in Pseudopregnant Mice

41. Paria BC, Ma WG, Tan J, et al (2001) Cellular and molecular responses of the uterus to embryo implantation can be elicited by locally applied growth factors. Proc Natl Acad Sci U S A 98:1047–1052. https://doi.org/10.1073/pnas.98.3.1047

42. Io S, Kabata M, Iemura Y, et al (2021) Capturing human trophoblast development with naive pluripotent stem cells in vitro. Cell Stem Cell 28:1023-1039.e13. https://doi.org/10.1016/j.stem.2021.03.013

43. Guo G, Stirparo GG, Strawbridge SE, et al (2021) Human naive epiblast cells possess unrestricted lineage potential. Cell Stem Cell 28:1040-1056.e6. https://doi.org/10.1016/j.stem.2021.02.025

44. Mischler A, Karakis V, Mahinthakumar J, et al (2021) Two distinct trophectoderm lineage stem cells from human pluripotent stem cells. J Biol Chem 296:100386. https://doi.org/10.1016/j.jbc.2021.100386

45. Syrett CM, Sierra I, Berry CL, et al (2018) Sex-Specific Gene Expression Differences Are Evident in Human Embryonic Stem Cells and during in Vitro Differentiation of Human Placental Progenitor Cells. Stem Cells Dev 27:1360–1375. https://doi.org/10.1089/scd.2018.0081

46. Tan L, Lacko LA, Zhou T, et al (2019) Pre- and peri-implantation Zika virus infection impairs fetal development by targeting trophectoderm cells. Nat Commun 10:1–12. https://doi.org/10.1038/s41467-019-12063-2

47. Chhabra S, Liu L, Goh R, et al (2019) Dissecting the dynamics of signaling events in the BMP, WNT, and NODAL cascade during self-organized fate patterning in human gastruloids

48. Tsuchida N, Kojima J, Fukuda A, et al (2020) Transcriptomic features of trophoblast lineage cells derived from human induced pluripotent stem cells treated with BMP 4. Placenta 89:20–32. https://doi.org/10.1016/j.placenta.2019.10.006

Tables

Table 1. ESC derived TSC samples clustered with organoids

Sample ID	paper	cell_type
SRR11791787	Liu_20[21]	iTSC from naive iPSC
SRR11791791	Liu_20[21]	extravillous trophoblast from naive iPSC
SRR11791792	Liu_20[21]	extravillous trophoblast from naive iPSC
SRR11791793	Liu_20[21]	iTSC from naive iPSC
SRR11791794	Liu_20[21]	iTSC from naive iPSC
SRR11791797	Liu_20[21]	extravillous trophoblast from naive iPSC
SRR11791798	Liu_20[21]	extravillous trophoblast from naive iPSC
SRR11791801	Liu_20[21]	syncytiotrophoblast from naive iPSC
SRR11791802	Liu_20[21]	syncytiotrophoblast from naive iPSC
SRR10257237	Dong_20[20]	hTSC from naive iPSC
SRR10257238	Dong_20[20]	hTSC from naive hESC
SRR10257239	Dong_20[20]	extravillous trophoblast from hESC
SRR10257240	Dong_20[20]	extravillous trophoblast from iPSC
SRR10257241	Dong_20[20]	syncytiotrophoblast from hESC
SRR10257242	Dong_20[20]	syncytiotrophoblast from iPSC
SRR10911111	Dong_20[20]	hTSC from naive iPSC
SRR11050189	Io_21[42]	syncytiotrophoblast from hESC
SRR11050190	Io_21[42]	syncytiotrophoblast from hESC
SRR11050191	Io_21[42]	extravillous trophoblast from hESC
SRR11050192	Io_21[42]	extravillous trophoblast from hESC
SRR13744481	Guo_21[43]	trophectoderm from naive iPSC
SRR13744482	Guo_21[43]	trophectoderm from naive iPSC
SRR13744486	Guo_21[43]	trophectoderm from naive iPSC
SRR13663546	Guo_21[43]	cytotrophoblast from hESC
SRR13663547	Guo_21[43]	cytotrophoblast from hESC
SRR10100464	Mischler_21[44]	hTSC from hESC

Table 2. ESC derived TSC clustered with trophoblast derived hTSCs

Sample ID	paper	cell_type
SRR10257236	Dong_20[20]	hTSC from naive hESC
SRR10911110	Dong_20[20]	hTSC from naive hESC
SRR13663510	Guo_21[43]	hPSC
SRR13663548	Guo_21[43]	cytotrophoblast from hESC
SRR13663549	Guo_21[43]	cytotrophoblast from hESC
SRR13663562	Guo_21[43]	cytotrophoblast from hESC
SRR13663563	Guo_21[43]	cytotrophoblast from hESC
SRR13663564	Guo_21[43]	cytotrophoblast from hESC
SRR13663565	Guo_21[43]	cytotrophoblast from hESC
SRR10100455	Mischler_21[44]	hTESC from hESC
SRR10100461	Mischler_21[44]	hTSC from hESC
SRR10100462	Mischler_21[44]	hTSC from hESC
SRR10100463	Mischler_21[44]	hTSC from hESC
SRR10100465	Mischler_21[44]	hTSC from hESC

Table 3. Cluster membership of samples generated by BMP derivation methods

Sample ID	paper	treatment_time	media	seurat_clusters
SRR7357227	Syrett_18[45]	Day 5	BMP4	0
SRR7357229	Syrett_18[45]	Day 5	BMP4	0
SRR7357231	Syrett_18[45]	Day 5	BMP4	0
SRR7357233	Syrett_18[45]	Day 5	BMP4	0
SRR7357235	Syrett_18[45]	Day 5	BMP4	0
SRR7357237	Syrett_18[45]	Day 5	BMP4	0
SRR9606444	Tan_19[46]	Day 4	BMP4	0
SRR9606445	Tan_19[46]	Day 9	BMP4	0
SRR10123845	Chhabra_19[47]	Day 2	BMP4	0
SRR10123846	Chhabra_19[47]	Day 2	BMP4	0
SRR10123847	Chhabra_19[47]	Day 2	BMP4_IWP2	0
SRR10123848	Chhabra_19[47]	Day 2	BMP4_IWP2	0
GSM3555684	Tsuchida_20[48]	Day 10	BMP4	1
GSM3555685	Tsuchida_20[48]	Day 10	BMP4	1
GSM3555688	Tsuchida_20[48]	Day 10	BMP4	1
GSM3555689	Tsuchida_20[48]	Day 10	BMP4	1
GSM3555692	Tsuchida_20[48]	Day 10	BMP4	1
GSM3555693	Tsuchida_20[48]	Day 10	BMP4	1
GSM3555696	Tsuchida_20[48]	Day 10	BMP4	1
GSM3555697	Tsuchida_20[48]	Day 10	BMP4	1
SRR7357226	Syrett_18[45]	Day 0	BMP4	6
SRR7357228	Syrett_18[45]	Day 0	BMP4	6
SRR7357230	Syrett_18[45]	Day 0	BMP4	6
SRR7357232	Syrett_18[45]	Day 0	BMP4	6
SRR7357234	Syrett_18[45]	Day 0	BMP4	6
SRR7357236	Syrett_18[45]	Day 0	BMP4	6
SRR10123849	Chhabra_19[47]	Day 2	BMP4_SB431542	6
SRR10123850	Chhabra_19[47]	Day 2	BMP4_SB431542	6

Table 4. Top 10 gene marker candidates for each cluster

gene	cluster	avg_log2FC	p_val_adj
VTCN1	0	12.0	2.7e-32
LUM	0	10.4	1.6e-29
TRIM55	0	8.8	3.1e-29
ITGB6	0	10.0	4.5e-29
PAPPA2	0	9.1	9.5e-29
MTUS2	0	9.9	4.6e-26
COL5A1	0	9.4	1.8e-23
COL3A1	0	11.1	3.4e-21
COL6A3	0	9.5	7.6e-19
ACTA2	0	9.4	2.4e-18
IGFBP6	1	8.3	1.5e-10
ANGPTL2	1	6.0	1.2e-08
THY1	1	9.2	3.5e-07
ITGA4	1	6.5	4.5e-06
RARRES2	1	6.9	2.1e-05
DDR2	1	6.3	1.3e-04
TMSB15A	1	6.5	5.5e-04
CD248	1	9.5	5.0e-03
SDC1	2	8.9	1.2e-22
MARK	2	6.0	3.0e-22
TREML2	2	8.1	8.6e-21
CST6	2	5.9	3.3e-20
CLDN1	2	6.9	4.6e-20
EBI3	2	6.6	6.5e-20
COL21A1	2	6.9	7.7e-20
DIO2	2	9.3	1.0e-19
MFSD2A	2	6.3	1.6e-18
SLCO4A1	2	6.4	1.6e-16

gene	cluster	avg_log2FC	p_val_adj
GPX3	3	9.4	8.9e-24
GIPC2	3	7.5	3.1e-23
FAM3B	3	7.9	4.8e-23
ACSS1	3	7.4	1.4e-22
THBD	3	7.1	1.6e-22
NRCAM	3	6.5	3.7e-22
NRN1	3	8.1	5.9e-22
EGR1	3	12.0	6.3e-21
FEB	3	8.1	9.9e-21
GALNT15	3	7.1	4.6e-20
PRDM14	4	5.6	3.0e-17
DPPA2	4	4.6	4.1e-17
POU5F1	4	5.4	6.4e-17
NANOG	4	5.8	6.7e-17
ERG	4	4.5	2.1e-16
TDGF1	4	4.7	2.6e-16
TNFSF9	4	4.3	5.1e-15
MT1G	4	5.1	2.1e-13
ZNF492	4	6.4	9.3e-06
BIN2	5	9.9	2.3e-10
DSC3	5	5.2	2.9e-10
LINC00668	5	6.4	1.2e-09
TNS4	5	5.6	1.5e-09
APR	5	6.4	3.7e-09
ITGB4	5	6.3	4.7e-09
FOLR1	5	6.6	1.3e-08
GJA5	5	5.5	8.0e-08
SLC12A3	5	8.0	3.7e-07

gene	cluster	avg_log2FC	p_val_adj
TLR7	5	5.3	2.7e-06
GRIK2	6	5.5	5.6e-09
YPEL1	6	4.2	8.0e-09
RIPPLY3	6	6.1	1.2e-08
APBA2	6	3.9	2.0e-08
DPYSL5	6	4.2	3.5e-08
MAP	6	3.9	3.6e-08
C6orf141	6	3.8	4.3e-08
SLC16A14	6	4.2	4.7e-08
CDH6	6	4.4	1.7e-07
SCUBE1	6	4.1	2.4e-07

Figures

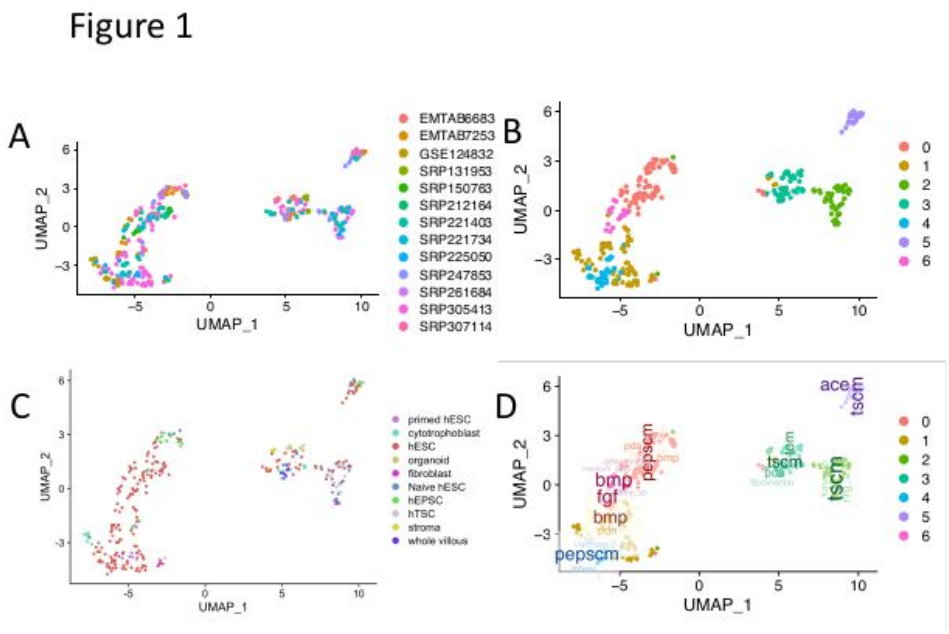

Figure 1

UMAP projections of integrated data sets of ESC to TSC/trophoblast conversion models. A) Samples coloured by data set origin. B) Samples coloured by statistical cluster numbered from 0-6 in order of the number of samples per cluster. C) Samples coloured by source material indicate clusters 2 contains hTSC and 3 holds organoids and in vivo isolates. Cluster4 is undifferentiated primed and naïve hESC and iPSC. D) Word cloud overlay of media used in cell culture/derivation for samples.

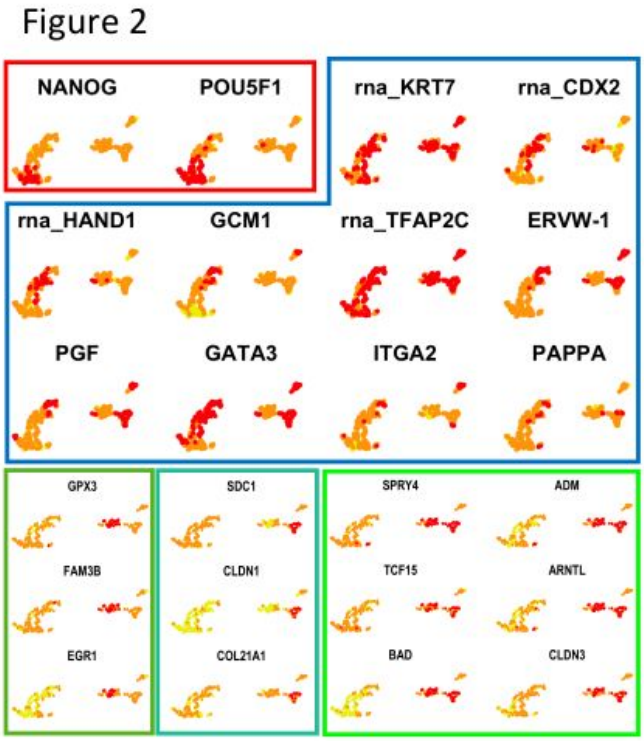

Figure 2

Ontology identifiers from EnrichmentMap analysis of differentially expressed genes to each cluster. Genes identified with increased expression (FDR<0.05 and a log2(fold change)>0.5) to each cluster in a one versus all calculation were assessed against the human Gene Ontology Biological Process.

EnrichmentMap organized significantly enriched ontologies (FDR<0.05) to consolidate ontologies with significant overlap in gene membership. A complete list of ontologies is in Supplemental Table 2.

Figure 3

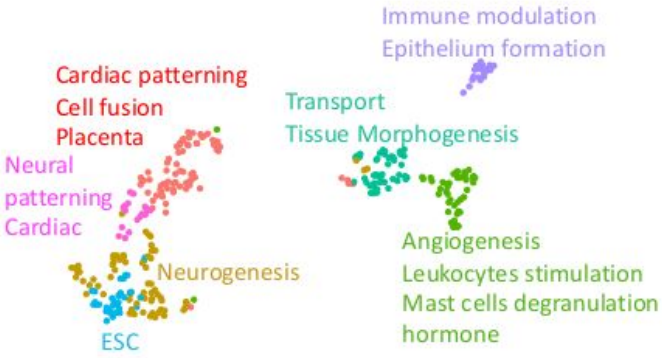

Figure 3

UMAP projections of samples coloured by expression of key markers. Expression of individual genes is colour scaled from low to high (yellow-orange-red). Markers of pluripotency show specificity to clusters 4 and 1 (NANOG, POU5F1) are boxed in red. Markers of trophoblast fate (blue box) show inconsistent patterns that encompass clusters 0, 2, 3 and 5, where only clusters 2 and 3 contain in vivo and in vitro trophoblast samples. Genes calculated to have enriched expression to trophoblast clusters 2, 3 or both clusters are boxed in green shades.

Figure 4

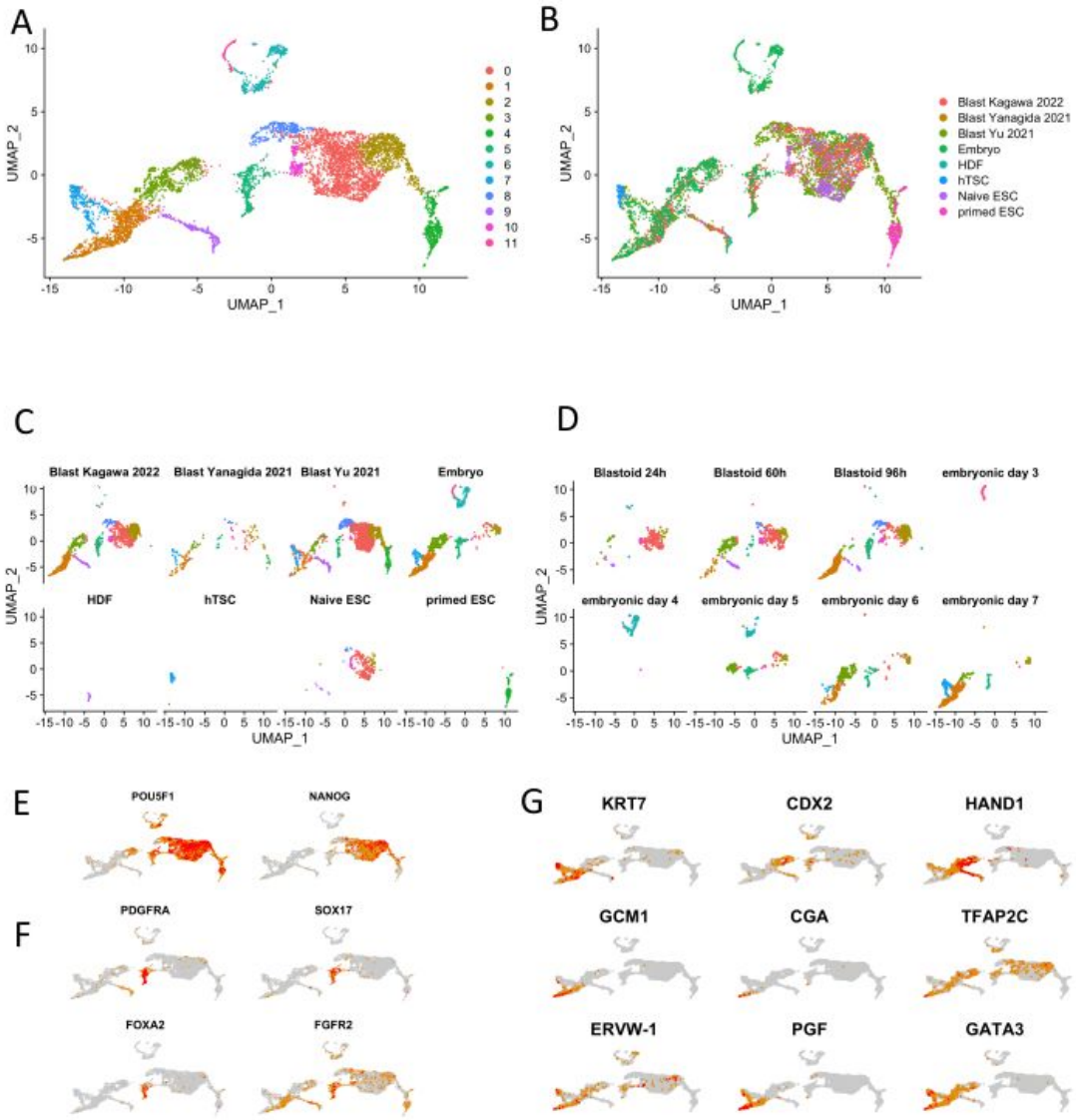

Figure 4

UMAP projections of aggregated single-cell sequencing data of blastoids, human embryos and stem cell lines. Each point represents an individual cell. A) Cells are coloured by cluster membership. Clusters 0-11 are numbered in order of largest to smallest. B) Cells are coloured by sample origin and show some sample unique or enriched arrangements, such as embryo populations, primed hESC, hTSC. C) Multi-panel display of UMAP projection coloured by cluster membership as in A and separated by sample origin as in B. Comparison of blastoid cells to embryo cells (top row) highlights contributions to similar clusters. Cell lines (bottom row) show tight aggregations typically to a single cluster. D) blastoid cells from Kagawa 2022 separated by culture time (24, 60 and 96 hours) compared to human embryos separated by time show a similar arrangement of cluster memberships from day 5 through 7. E-G are cells coloured by gene expression from low to high (grey-orange-red). E) Pluripotency genes. F) Hypoblast or primitive endoderm markers. G) Trophoblast markers show different patterns. CDX2 and HAND1 are expressed in clusters that initiate on day 5 of embryo development and are reduced in other clusters. TFAP2C and GATA3 show general expression through all clusters. GCM1, CGA, PGF and ERVW-1 show expression beginning in later clusters associated with embryonic days 6 and 7.

Figure 5

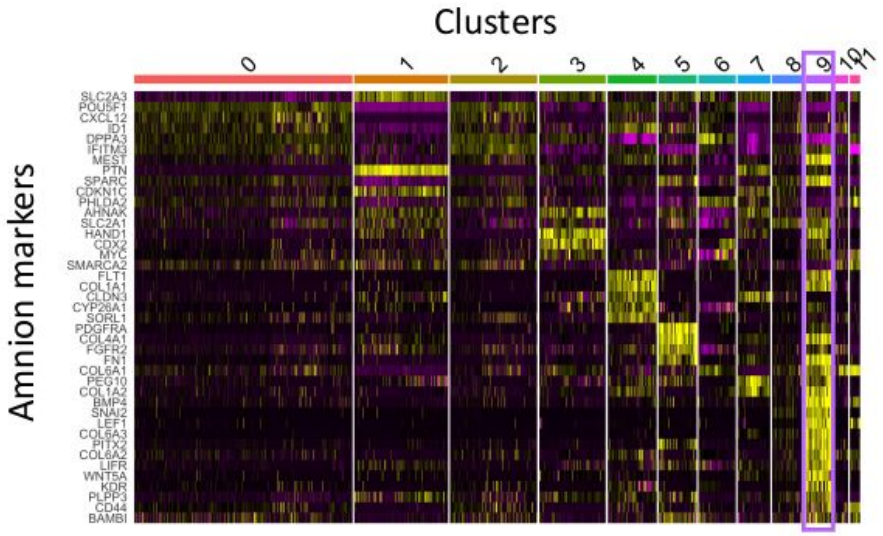

Figure 5

Heatmap of mouse embryonic amnion markers in cells organized by cluster membership. Cluster 9 (boxed in purple) is only found in blastoid samples but associated with trophoblast cell clusters found in embryo and blastoid samples. Cluster 9 shows high enrichment of differentially expressed amnion genes.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaltable1datasets.csv
- Supplementaltable2samples.csv
- Tableofontologyenrichments.xlsx