Single-cell analysis of human trophoblast stem cell specification reveals activation of fetal cytotrophoblast expression programs including coronavirus associated host factors and human endogenous retroviruses

Authors:
Ethan Tietze*[a], Andre Rocha Barbosa*[a,b,c], Veronica Euclydes*[a,b,h], Hyeon Jin Cho[a], Yong Kyu Lee[a], Arthur Feltrin[a,d], Joyce van de Leemput[a,$], Pasquale Di Carlo[a,i], Tomoyo Sawada[a,b], Kynon J. Benjamin[a,b,f], Helena Brentani[c,h], Joel E. Kleinman[a,f], Thomas M. Hyde[a,b,f], Daniel R. Weinberger[a,b,e,f,g], Gianluca Ursini[a,e], Ronald McKay[a,e], Apua C.M. Paquola[a,b], Joo Heon Shin[a,b], and Jennifer A. Erwin[#a,b,e]

[a] Lieber Institute for Brain Development, Baltimore, Maryland, USA
[b] Department of Neurology, Johns Hopkins School of Medicine, Baltimore, Maryland, USA
[c] Inter-institutional Graduate Program on Bioinformatics, University of São Paulo, São Paulo, SP, Brazil
[d] Center for Mathematics, Computation and Cognition, Federal University of ABC, Santo André, SP, Brazil
[e] Department of Neuroscience, Johns Hopkins School of Medicine, Baltimore, Maryland, USA
[f] Department of Psychiatry & Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, USA
[g] Department of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA
[h] Department of Psychiatry, University of Sao Paulo, Medical School, São Paulo, Brazil
[i] Department of Basic Medical Science, Neuroscience, and Sense Organs – University of Bari Aldo Moro, Bari, Italy
*Equal contributor
[#] Lead contact and corresponding author: Jennifer.Erwin@libd.org

[$] Present address: Center for Precision Disease Modeling and Division of Endocrinology, Diabetes and Nutrition, Department of Medicine, University of Maryland School of Medicine, 670 West Baltimore Street, Baltimore, MD, 21201, USA

Summary

The human placenta is increasingly a focus of research related to early child development and the impact of maternal hyperimmune states. The ability to model human trophectoderm disease states from human pluripotent stem cells, the nature of human pluripotent stem cell potency and the mechanisms regulating human trophectoderm specification remains poorly understood. Recent work suggests that only the naive state can give rise to trophectoderm and that primed iPSC generate mixed amnionic and mesoderm lineages. Here we identify conditions that efficiently drive the specification of primed iPSC to trophectoderm, named Trophoblast Stem Cell (TSC). iPS-derived-TSC share transcriptional, morphological and functional characteristics with human *in vivo* cytotrophoblasts including activation of human endogenous retroviruses, expression of COVID-19 associated host factors and generation of multinucleated syncytiotrophoblasts with a large fusion index. At high densities in 5% O_2, iPS-derived-TSC form villi-like structures and express extravillous and syncytiotrophoblast proteins HCG-β and HLA-G. Using temporal single cell RNAseq, we define the molecular changes associated with specification under three separate conditions: 1) BMP4, 2) BMP4 and inhibition of WNT, 3) activation of EGF and WNT, inhibition of TGFbeta, HDAC and ROCK signaling (named TSC). With 9,821 high-quality single cell transcriptomes, we find that BMP4 gives rise to mesenchymal cells while TS conditions lacking exogenous BMP4 generate a stable proliferating cell type that is highly similar to six week placenta cytotrophoblasts. TFAP2A programs the specification of primed iPS cells to TSC without transitioning through a naive state. TSC specification independent of exogenous BMP4 will allow for robust and reproducible studies of the cytotrophoblast component of human placenta.

Introduction

The first cellular specification occurs early in human development when blastomeres divide asymmetrically to generate the outer trophectoderm and the inner cell mass of the pre-implantation embryo. Upon implantation, primitive syncytial trophoblast invades and remodels the maternal decidual environment before specifying into the cytotrophoblast and syncytiotrophoblast layers found in placenta villi. Currently, there is a limited understanding of the early developmental trajectories of human trophoblast differentiations (Rossant, 2018). Human models are essential to understanding how the human placenta is built. Placental structures and interaction of placental cells with the maternal environment vary tremendously across mammals. Comparative genome analysis demonstrates that genes involved in placenta and reproduction are amongst the most divergent between human and mouse (Mouse Genome Sequencing Consortium et al., 2002). The evolutionary tinkering between species is driven by endogenous retroviruses (ERVs) ((Johnson, 2019)). Now comprising 8% of the human genome, human endogenous retroviruses are retroviruses acquired by ancestral humans which then integrate into the genome and are passed onto the next generation (Lander et al., 2001). ERV-derived proteins have been repeatedly co-opted for placental functions to promote cell–cell fusion and formation of syncytial trophoblast by endogenous retrovirus group W member 1 (*ERVWE1* or *syncytin-1*) (Blond et al., 1999, 2000; Cornelis et al., 2013) and *ERVFRD-1* (Blaise et al., 2003). ERVs harbor regulatory sequences that function as enhancers to control placental

gene expression, such as a primate-specific ERV which functions as a placenta-specific enhancer for corticotropin-releasing hormone (CRH), a hormone linked to the control of birth timing in humans (Blond et al., 2000; Dunn-Fletcher et al., 2018). Therefore, human cellular systems are crucial to understand the many human and primate-specific properties of human placenta.

Human induced pluripotent stem cells are powerful models to understand causal genetic mechanisms of human disease. For organs such as the brain, rapid progress has led to iPS models with increasing complexity and utility (Mansour et al., 2020). For the human placenta, which is increasingly implicated as a risk factor for developmental disorders associated with a maternal hyperimmune state, the ability to faithfully model human trophectoderm disease states from human pluripotent stem cells is a matter of debate. Several model systems have been developed to differentiate primed and naive iPSC to trophectoderm. Soon after the derivation of the first human stem cells, Bone Morphogenetic Protein 4 (BMP4) was shown to differentiate human stem cells to syncytial cytotrophoblasts (Xu et al., 2002). Since then, various BMP4-based methods have been described including the BAP protocol (BMP4, A83-01 an alk-5 inhibitor and PD173074 and FGF and VEGF inhibitor)(Roberts et al., 2018) and a combination of BMP4 and IWP2 (Horii et al., 2019) , an inhibitor that blocks the biosynthesis of mature WNT proteins (Chen et al., 2009). Primed iPS differentiations have almost exclusively included the use of BMP4, despite the fact that single cell sequencing of the peri-implantation human embryo found the BMP signaling pathway and SMAD protein signal transduction to be enriched in primitive endoderm, not trophectoderm (Blakeley et al., 2015; Petropoulos et al., 2016; Stirparo et al., 2018). BMP4 induces differentiation of mesendoderm of the embryo proper (Goldman et al., 2009; Schneider et al., 2003; Yang et al., 2008; Zhang et al., 2008) and also extraembryonic endoderm (Vallier et al., 2009) in addition to upregulating some genes expressed in trophoblasts. TB differentiations using these model systems generate heterogeneous cells with mesodermal contamination and fail to maintain a trophoblast stem cell population (Bernardo et al., 2011; Drukker et al., 2012; Erb et al., 2011; Yabe et al., 2016). Models inconsistently manipulate morphogens including BMP and WNT to generate trophoblast-like cells that are characterized with a limited set of genes. Recent single cell analyses of the early placenta have characterized the expression of these genes in multiple placental cell types (Liu et al., 2018; Suryawanshi et al., 2018; Vento-Tormo et al., 2018). As many of these genes are either expressed in many different cell types or not detected at all, their findings call into question the confidence with which specific iPS derived placental cell types can be identified. Therefore, it is a matter of debate which pluripotent states can give rise to trophectoderm and how similar these trophectoderm cells are to *in vivo* placenta.

Understanding if and how induced pluripotent stem cells can generate human trophectoderm is central to understanding human pluripotent stem cell potency and the mechanisms regulating human trophectoderm specification. In the present study, we first ask if BMP4 is required for primed pluripotent stem cell specification to trophectoderm. In the absence of BMP4, we find that activation of wingless/Integrated (Wnt) and epidermal growth factor (EGF) while inhibiting Transforming Growth Factor Beta (TGFb) histone deacetylase (HDAC) and Rho-associated protein kinase (ROCK) efficiently specify multiple primed pluripotent stem cell lines to

trophectoderm. These factors were previously discovered to derive and grow trophoblast stem cells from blastocysts and the early placenta (Okae et al., 2018). We demonstrate that iPS-derived TS cells are capable of self-renewal for at least 30 passages, differentiate into syncytial cytotrophoblasts and villus cytotrophoblasts and generate villi-like structures in low oxygen. We used temporal single-cell sequencing analysis to elucidate the transcriptional landscape of this specification compared to previously implemented protocols involving BMP4. We find that BMP4 conditions alone largely generate extraembryonic mesoderm cells. BMP4 combined with inhibition of WNT shifts differentiation away from a mesodermal fate while upregulating trophoblast associated lineage markers but cells are transcriptionally distinct from *ex vivo* cytotrophoblasts. Trophoblast cells generated by TS conditions are transcriptionally highly similar to *ex vivo* cytotrophoblasts and uniquely upregulate endogenous retroviral genes. Specification of primed iPS cells to TS occurs without entering a naive state and *TFAP2A* and *ESRRG* are central genes regulating the specification. Defining the molecular mechanisms specifying the trophectoderm is key to understanding placental associated diseases and improving iPS-derived models.

Results:

Derivation of trophoblast stem cells (TS) from primed human embryonic stem cells and induced pluripotent stem cells without BMP4

In search of alternative methods to specify iPS cells to trophectoderm, we reasoned that primed human pluripotent stem cells might be capable of specification to trophoblast stem cells directly by activating wingless/Integrated (WNT) and epidermal growth factor (EGF) and inhibiting Transforming Growth Factor Beta (TGFb) histone deacetylase (HDAC) and Rho-associated protein kinase (ROCK). While previously described primed iPS differentiations to trophectoderm have added BMP4, BMP4 is not expressed in the early pre-implantation trophoblast and it is not known if iPS can differentiate to trophectoderm in the absence of BMP4 ((Blakeley et al., 2015; Petropoulos et al., 2016; Stirparo et al., 2018). Activation of WNT and EGF and inhibition of TGFb, HDAC and ROCK enable the long-term culture of human trophoblast stem cells derived from primary embryos (Okae et al., 2018). 24 hours after clump passage of the human embryonic stem cell line H1 (Thomson et al., 1998) in primed feeder-free stem cell conditions, ES colonies demonstrated domed tight packing with defined edges characteristic of primed cells (Figure 1A and B). 48 hours post passage, cells were switched directly to TSC media. By brightfield imaging over the subsequent 6 days, we observed that cells proliferated rapidly and adopted a flatter appearance with some cells adopting a cobblestone appearance (Figure 1A-D). After passaging, two morphologically distinct populations emerged. Circular colonies with an epithelial-like appearance could be seen surrounded by phase bright fibroblastic cells (Figure 1E,F). The inner epithelial-like cells continued to proliferate and appeared morphologically similar to TS cells derived from primary human embryos (Figure 1G). Immunofluorescence revealed that subpopulations of the colony cells expressed TP63 and/or KRT7 and minimally expressed VIM while surrounding fibroblastic cells strongly expressed VIM (Figure 1H). Over the course of several passages, the epithelial-like cells preferentially expand and reach purity between passage 5-7. These TSC widely express both TP63 and KRT7 (Figure 1I-K,

167 Supplemental Figure 1) without expressing the pluripotency genes SOX2 and NANOG
168 (Supplemental Figure 1). In summary, we find that activating wingless/Integrated (Wnt) and
169 epidermal growth factor (EGF) and inhibiting Transforming Growth Factor Beta (TGFb) histone
170 deacetylase (HDAC) and Rho-associated protein kinase (ROCK) directly in primed iPS lines
171 results in efficient specification of trophoblast stem cells.
172
173 Pluripotent stem cell lines vary in their propensity to differentiate to different cell types (Osafune
174 et al., 2008). Therefore, we asked if TSC conditions without exogenous BMP4 could specify a
175 variety of primed human stem cell lines. We confirmed TS specification by VIM, KRT7 and TP63
176 immunofluorescence on iPS lines reprogrammed from dermal fibroblasts by sendai, named
177 2014.06 (Supplemental Figure 2), and two postmortem dura fibroblast-derived iPS cell lines
178 reprogrammed by episome, named LIBD7c6 and LIBD1c8 (Sawada et al., 2020)(Figure 1 H-K).
179 To confirm the ability to proliferate as TS-like cells, iPS-derived TS lines were passaged for up
180 to 32 passages and maintained KRT7 and TP63 expression (Figure 1K).
181
182 **Hypoxia induces differentiation of iPS-derived TS cells into villi-like structures composed**
183 **of differentiated syncytial cytotrophoblasts and villus cytotrophoblasts**
184
185 To investigate the maturation potential of iPS-derived-TSC, we evaluated the ability of iPS-
186 derived TS cells to differentiate to mature trophectoderm cell types. Proliferative trophoblast
187 stem cells are bipotential stem cells with the capacity to differentiate to syncytial
188 cytotrophoblasts and extravillous cytotrophoblasts (Okae et al 2018). Syncytial cytotrophoblast
189 is a terminally differentiated multinucleated epithelial layer that infiltrates the maternal
190 endometrium. The multinucleated cell forms from multiple cytotrophoblast cell fusions. When TS
191 cells were grown to high density, we observed the spontaneous formation of multinucleated
192 cells by brightfield microscopy. We confirmed that multinucleated syncytial cells are Ki-67
193 negative nuclei surrounded by the tight-junction protein ZO-1 positive cells (Figure 2A and B).
194 We observed large areas of syncytial formation fused nuclei (Figure 2A and S3).
195 Syncytiotrophoblasts produce and secrete high levels of human chorionic gonadotropin
196 (hCGB3), a hormone essential for cytotrophoblast differentiation, immuno-suppression, and
197 placental invasion (Cole, 2010). We observe expression of hCGB3 in regions of the
198 differentiated cytotrophoblast cultures, indicating differentiation to syncytiotrophoblasts (Figure
199 2C). Extravillus trophoblasts express human leukocyte antigen G (HLA-G), the main molecule
200 involved in maternal immune system evasion upon placental invasion of the endometrium. We
201 also readily detected regions of HLA-G positive cells, indicating iPS-derived TS cells are
202 capable of differentiation to extravillous cytotrophoblast (Figure 2D). Expression of hCGB3 and
203 HLA-G was specific to iPS-derived TS cells and not detected in iPS cells (Supplemental Figure
204 1I).
205
206 To further induce differentiation of the iPS-derived-TSC, we tested if cells will respond to altering
207 oxygen concentrations. Hypoxia is a natural cue that stimulates syncytial cytotrophoblast
208 differentiation *in vivo* (Wakeland et al., 2017), therefore we transitioned a high density TS cell
209 culture from atmospheric oxygen of 20% to 5%. In 5% oxygen, these cells demonstrate minimal
210 contact inhibition and form 3D villi-like structures that protrude from the monolayer surface of

211 the culture dish (Figure 2 E-H). When imaged by confocal microscopy, these villi-like structures
212 are approximately 50 microns in height. In summary, iPS-derived-TSC readily differentiate into
213 both HLA-G positive villus cytotrophoblasts and hCGB3 positive syncytial cytotrophoblasts.
214
215 **Single cell RNA sequencing reveals unique transcriptional programs for iPS-derived TS**
216 **specification**
217
218 To define the molecular events involved in the specification of iPS cells to trophectoderm, we
219 performed temporal single cell RNAseq analysis of human iPS cells specified to placental
220 trophoblast cells by three separate conditions: 1) addition of BMP4 (Xu et al., 2002), 2) addition
221 of BMP4 and inhibition of Wnt with the small molecule inhibitor IWP2 (Horii et al., 2019), 3)
222 Trophoblast Stem Cell condition (TSC) activation of EGF and Wnt, inhibition of TGFbeta, HDAC
223 and ROCK signaling to generate trophoblast stem cells (Figure 3A). Sequencing was generated
224 for iPS cells before differentiation (day 0), for BMP4 on days 4 and 6, for BMP4+IWP2 on days
225 2, 4 and 6, and on days 2, 4, 6 and 8 for TS condition, resulting in 10 single-cell transcriptomes
226 (Figure 3A). In TS condition, the day 8 is a timepoint two days after the first passaging
227 annotated as p1. In the BMP4 and BMP4 + IWP2 conditions, no proliferative cells survived
228 passaging, therefore the p1 time point was not sequenced. Using the highly parallel droplet
229 based single cell sequencing method Drop-Seq (Macosko et al., 2015), 9,821 high-quality cells
230 were obtained after removing cells with less than 1000 genes detected and more than 20% of
231 mitochondrial mapping. The average ratio of mapping reads for mitochondrial genes in all cells
232 were 3.9%, indicating good viability (Supplemental Figure 4A). The total number of genes
233 detected ranged between 21804 for BMP4 + IWP2 day 4 to 27134 for TS day 4 (Supplemental
234 Figure 4B) time point with between 563 to 1572 high quality cells per timepoint (Supplemental
235 Figure 4C).
236
237 In order to explore the data in an unbiased manner, we performed dimensionality reduction for
238 the most variable genes across all cells with Principal Component Analysis (PCA) and Uniform
239 Manifold Approximation and Projection (UMAP) embedding. Generally, single cells separated by
240 differentiation stage and condition, with UMAP dimension 2 generally capturing the
241 differentiation condition (Figure 3B). To identify transcriptionally similar groups of cells, we
242 performed a graph-based clustering analysis of k nearest neighbors using the euclidean
243 distance of the 30 first PCs in Seurat (Stuart et al., 2019). We identified 19 clusters, and clusters
244 contained mostly a single time-point and media condition (Figure 3B-D and S4D and Table S1).
245 IPS clusters 1, 2 and 3 mostly contained IPS cells with slight contributions from the first days of
246 all three differentiation conditions. Cells from the BMP4+ IWP-2 condition were found
247 predominantly in a group of five heterogeneous clusters termed BR (BMP responsive). BR1 was
248 composed of roughly equal parts iPS cells and BMP4 + IWP-2 day 2 treated cells. BMP4+IWP2
249 cells made up the majority of the cells in the other BR clusters alongside a smaller population of
250 BMP4 cells. BR2 and BR3 were predominantly derived from BMP4+IWP2 day 2 cells, while day
251 4 contributed to BR4, and day 6 contributed to BR. Cells from the BMP4 condition almost
252 exclusively formed four clusters termed SC (Stromal cell). SC1 and SC3 were predominantly
253 composed of cells from day 4 of the differentiation while SC2 and SC were predominantly from
254 day 6 cells. In addition, cluster SC contained a small contribution of cells originating from the

255 final differentiation day of the TS condition termed TSP1 (TS passage 1). Cells from the TS
256 differentiation almost exclusively populated 7 clusters termed TS. TS1 was exclusively
257 composed of Day 2 cells and TS 2-5 were composed of a mixture of cells from days 2-6. The
258 final TS day (TSP1) separated into two clusters named TS and TS6, which confirms the
259 brightfield observations of 2 cell populations. Interestingly, 14 cells from BMP4+IWP2 d6 and 4
260 cells from the BMP4 d6 contribute to the final TS cluster (Figure 3D, Table S1).
261
262 To understand how transcriptional profiles change during the differentiation, we asked which
263 cells express canonical pluripotent, trophectoderm and mesoderm specific genes. We find that
264 the stem cell transcription factor *SOX2* is highest expressed in the iPS stage and is absent from
265 the most differentiated BR, TS and SC clusters (Figure 3E). *SOX2* expression is maintained at a
266 higher level in the initial days of the TS conditions compared to BMP WNTi and BMP treatment
267 and absent from the most differentiated TS cells. The Paired-like homeodomain transcription
268 factor 2 PITX2 is highly expressed in the later stages of BMP4 differentiation, suggesting that
269 BMP4 conditions generate placental stromal cells (Figure 3F). In single cell RNAseq of the
270 human placenta, *PITX2* was identified as the top marker gene for a stromal fetal communicating
271 cell type F1 by Vento-Tormo (Liu et al., 2018; Vento-Tormo et al., 2018). Keratin 7, *KRT7*, and
272 human chorionic gonadotropin subunit alpha, *CGA*, are expressed in both stem cell derived
273 cytotrophoblasts and *in vivo* placenta (Yabe et al., 2016). Indicating specification to
274 trophectoderm transcriptional programs, we find both *KRT7* and *CGA* are expressed in the later
275 stages of the BMP+ WNTi condition and TS condition (Figure 3G-H). Overall, both TS and
276 BMP4+ WNTi treated cells demonstrated an increased expression for most TB markers
277 including *GATA3* and *TFAP2C* (Figure S5A) while the BMP4 alone conditions demonstrated a
278 more mesoderm associated expression of *GATA4*, *TBXT* and *PDGFRA* (Figure S5B). In mouse,
279 the Caudal-type homeobox 2 (CDX2) is involved in the segregation of the ICM and
280 trophectoderm lineages at the blastocyst stage by repressing OCT4/POU5F1 and NANOG in
281 the TE. Overexpression of CDX2 in mouse embryonic stem cells causes differentiation to
282 trophoblast stem cells (Erwin et al., 2012; Niwa et al., 2005). In humans, *CDX2* expression is
283 initiated after blastocyst formation and has variable expression patterns in trophectoderm
284 (Niakan and Eggan, 2013). Similar to TS cells derived from primary human embryos, *CDX2* is
285 not expressed in the differentiated TS state, and there is sparse expression of *CDX2* at d4 of all
286 differentiations (Figure S5A).
287
288 Next we asked which genes are preferentially expressed in each cluster by non-parametric
289 Wilcoxon rank sum test (adj.p-value < 0.05; logFC > 0.25)(Figure 3I, Table S2). We defined cell-
290 population markers by comparing gene expression of each cell population compared to all other
291 cells. *DUSP6* and *THY1* are primed stem cell state genes and are among the top marker genes
292 identified in the iPS cells. The SC cluster marker genes includes WNT signaling genes *RSPO2*
293 a WNT agonist (Kazanskaya et al., 2004) and *PITX2* which interacts with both WNT signaling
294 and regulates collagen expression during development (Hendee et al., 2018). Markers of the BR
295 cluster include *PRSS23* a serine protease involved in Snail-dependent Epithelial to
296 Mesenchymal Transition (Chen et al., 2013) and regulators of Insulin-like growth factor signaling
297 *IGFBP7* (Oh et al., 1996) and *PAPPA2* (Overgaard et al., 2001). Markers of the TS clusters
298 include genes involved in EGFR signaling *NPNT* (Arai et al., 2017), *IFI6* (Meyer et al., 2015),

and *MUC15* (Wang et al., 2013). Additionally, *ARHGAP18* is a Rho GTPase activating protein that functions as a downstream effector of YAP signaling (Porazinski et al., 2015) recently implicated in the specification HiPS to bipotent trophoblast stem cells (Mischler et al., 2019).

In sum, TSC, BMP4 and BMP4 with inhibition of WNT pattern different cell types. Activation of WNT and EGF and inhibition of TGFb, HDAC and ROCK specify primed iPSC to cytotrophoblasts, BMP4 alone induces mesenchymal fates and inhibition of WNT in the context of BMP4 generates a small percentage of TS cells. For use by the community, we provide a web browser of the single cell expression profiles (http://erwinpaquolalab.libd.org/publications/singlecellTS) and the differentially expressed genes for all pairwise cluster comparisons (Table S3).

iPS-derived trophoblast stem cells resemble first trimester placental cytotrophoblasts while BMP4 differentiation generates placental mesenchymal cells

We next asked how similar are the iPS-derived placental cells to *ex vivo* cells found in human placenta. There is considerable debate about the authenticity of iPS derived trophoblast cells and primary trophoblast stem cells, therefore we set out to directly compare the single cell transcriptional profiles to *ex vivo* datasets. To characterize the expression profiles of the 19 clusters, we analyzed the overlap with cell type-specific genes identified by two published single-cell studies of early human placenta (Liu et al., 2018; Vento-Tormo et al., 2018). We performed gene set enrichment analyses using Cell-Specific Expression Analysis (CSEA) (Xu et al., 2014). We identified cell type specific genes that were preferentially expressed in each cell type/cluster of the 19 clusters from iPS-derived cells by the specificity index probability (pSI) statistic at thresholds of $p< 0.05$. Next, we tested whether cell-type specific genes previously identified by single cell studies of human placenta are over-represented in the cell-type specific genes from our iPS specification by hypergeometric test and applied the Bonferroni correction for multiple comparisons, considering all the tested gene lists [$\alpha = 0.05/ (19\times (38+14)) = 5.1\times10^{-5}$]. Vento-Tormo identified a set of 38 cell types from maternal (decidua and blood) and fetal tissues (placenta) of first trimester placenta and reported 30 cell-type specific genes per cell type (Vento-Tormo et al., 2018). We found that the transcriptional profile from all iPS clusters and clusters from d2-d4 timepoints did not have significant enrichment for any of the placental cell type clusters (Figure S6A). In contrast, the most mature iPS-derived TS cluster was highly enriched for the syncytial cytotrophoblast (SCT) and villous cytotrophoblast (VCT) cluster ($p=\log_{10}^{-32}$ and $\log_{10}^{-17}$) (Figure 4A). The expression profile of the iPS-derived TS cluster was highly specific for SCT and VCT with no significant enrichment for any other cell-type in the maternal fetal interface. The D6 BMP4 cluster SC was most enriched in the two fibroblast clusters (F1 and F2) ($p=\log_{10}^{-26}$ and $p=\log_{10}^{-13}$). Vento-Tormo described these placental fibroblasts as mesenchymal stromal cells of fetal origin that derive from the primitive endoderm expressing GATA4, GATA6, PDGFRa, and SOX17 (Figure S6 and S7). The BMP4 + IWP2 conditions (BR) were not strongly enriched for any specific placental cell type and instead had weaker enrichment for several clusters including syncytial cytotrophoblast (SCT), and maternal decidua derived F1 endodermal cell type and dS1. In sum, iPS-derived TS cells are highly

342 enriched for expression of genes specific to syncytial cytotrophoblast and villous cytotrophoblast
343 of *ex vivo* human placenta.
344
345 Next we compared the expression profiles to early placental dataset from the 8 week and 24
346 week placenta (Liu et al., 2018) (Figure 4B). Again, we found strong enrichment for the iPS-
347 derived TS cells in the eight week fusion competent CTB1 (cytotrophoblast, log_{10}^{-9}). BMP4 alone
348 and BMP4 + IWP2 conditions were enriched for two different mesoderm cells identified by Liu et
349 al (log_{10}^{-23} and log_{10}^{-4}). To visualize placental cell-type specific expression patterns among the
350 iPS-derived cell types, we plotted a heatmap of the top unique markers of the fetal cell types
351 identified in Vento-Tormo and polar trophectoderm markers identified in (Petropoulos et al.,
352 2016) (Figure 4C). Consistent with the Gene Set Enrichment Analysis, we observed high
353 expression for a majority of unique marker genes previously identified in polar trophectoderm
354 villus cytotrophoblasts (VCT) and syncytial cytotrophoblasts (SCT). We also compared the
355 expression profiles of all iPS-derived cellular clusters to previously identified lineage and marker
356 genes from pre-implantation datasets (Stirparo et al 2018, Figure S7A). Again, we found the
357 highest expression of trophoblast genes KRT18, TEAD3, GATA3 and GATA2 in iPS derived TS,
358 with moderate expression in BR cells. Primitive endoderm genes SOX17, GATA4, GATA6 were
359 highest expressed in SC cells. When compared to primary trophoblast cells (Suryawanshi et al
360 2019, Figure S7B), we confirmed that TS cells expressed the highest level of cytotrophoblast
361 and EVT specific genes. In summary, iPS cells specified via TS conditions have a
362 transcriptional profile highly homologous to cytotrophoblasts taken directly from human placenta
363 while BMP4 differentiation generates placental mesenchymal cells.
364
365 **Primed iPS specify to trophoblast by activation of TFAP2A without activation of naive**
366 **stem cell programs**
367
368 It has been reported that primed stem cells are restricted in their potency and unable to
369 differentiate to TSC while naive stem cells readily differentiate to TSC because they are poised
370 to specify to trophectoderm (Dong et al., 2020). Therefore, we asked if the primed iPS or TSCM
371 treated cells adopt a naive stem cell program during differentiation. For 12 previously identified
372 naive marker genes (Messmer et al., 2019), naive transcripts were nearly absent. Five genes
373 HORMAD1, ALPPL2, KHD3L, TRIM60, and HYAL4 had no observed expression in any cell
374 (Figure 5A, Table S4). Less than 1% of cells for any condition had detectable expression for six
375 other marker genes ALPP, OLAH, LYZ, MEG8, KDC1L, and FAM151A. ZNF729 was the only
376 gene expressed in 0.51% of TSp1 cells to 14.69% of IWP2d2. In contrast, primed marker genes
377 were consistently expressed in iPS clusters and often throughout the differentiation state. For
378 example, DUSP6 was detected in 74.87% of the iPS cells, indicating that the iPS initiated in the
379 primed state. In sum, primed iPS transition to cytotrophoblasts without transitioning through a
380 naive state.
381
382 We asked how BMP signaling contributes to the differentiation by evaluating BMP family
383 signaling in all samples. Low percentages of cells expressed BMP4 in TS conditions, while
384 BMP4 was significantly upregulated in several cell clusters including Day 4 BMP4 and BMP4 +
385 IWP2 treated cells compared to iPS and TS cell clusters (Figure 5B, Table S4). BMP5 and

386 BMP7 were expressed in a larger portion of TSCM treated cells compared to iPS (BMP5: 0-
387 0.67% iPS vs 24.72% TS1 cells and BMP7: 1.01-4.71% iPS vs 29.06% TS1 cells). The BMP
388 effector proteins SMAD1,4, and 5 is expressed in the first days of all conditions and is absent
389 from the mature TS cluster. We conclude that BMP5 and BMP7 signaling patterns iPS-derived
390 TS specification instead of BMP4.
391
392 To identify the most significant gene expression changes associated with the initial stages of TS
393 specification, we studied a subset of the sample using only cells from the iPS and TS day 2
394 samples and applied a single-cell trajectory inference and pseudotime estimation (STREAM) to
395 the single cell transcriptome data (Chen et al., 2019). Briefly, single cells were ordered along
396 probabilistic trajectories and a numeric value referred to as pseudotime was assigned to each
397 cell to indicate how far it progresses along the dynamic differentiation. STREAM identifies an
398 initial branch composed of both iPS and TS day2 cells (S3-S0). This level of analysis also
399 identified branches containing IPS cells (S0-S2, S0-S4), reflecting the pluripotency continuum,
400 as well as a TS day 2 (S0-S1) branch reflecting the exit from pluripotency (Figure 5C and Table
401 S5). On branch S3-S0, node S3 was predominantly composed of iPS cells while TS day 2 cells
402 are more abundant towards S0 (Figure 5C, S8A,B). We next identified the transcriptional states
403 that contribute to the specification by identifying highly expressed genes within a branch that are
404 correlated with the transition along the specification branches (see methods). Top positively
405 correlated transition genes along branch S3-S0 include *LRRN3* which amplifies MAPK signaling
406 through EGFR (Fukamachi et al., 2002), *MEIS2* which codes for a TALE homeobox protein that
407 promotes tumor progression through the combined action of two isoforms increasing YAP
408 nuclear translocation and activating the WNT/b-catenin pathway (Guan et al., 2019), *BMP5*,
409 *BMPR1B* and *TFAP2A* a transcription factor key to the suppression of pluripotency and
410 expression of trophoblast associated genes (Krendl et al., 2017) (Figure 5C, S8C). These genes
411 were expressed not only along branch S3-S0 but also along branch S0-S1 (Figure S8C). The
412 branch S0-S1 captures the transition out of the pluripotent state and into the TS state with
413 upregulation of cytokeratins (*KRT18*, *KRT19*, *KRT8*) and calcium binding proteins (*S100A11*,
414 *S100A10*) known to be expressed in trophectoderm (Abd El-Aleem and Dekker, 2018; Liu et al.,
415 2006) (Figure S8D). Also found as a top transition gene, *TMSB4X* is implicated in stemness of
416 progenitor trophoblast cells of first trimester human placentas by increasing NOTCH1 activity
417 (Huang et al., 2016; Lee et al., 2018; Lv et al., 2013). In summary, trophectoderm transcriptional
418 programming occurs early in the specification and utilizes previously identified trophectoderm
419 associated regulatory factors including *TMSB4X*, *YAP*, *BMP5*, and *TFAP2A*.
420
421 **STREAM pseudotime analysis of trophoblast differentiations reveals upregulation of**
422 **stromal mesenchyme upon BMP4 treatment and trophectoderm transcriptional**
423 **programming upon TS treatment.**
424
425 To understand the transcriptional changes associated with the full trajectory of iPS to
426 trophectoderm specification, we repeated the STREAM analysis with the complete iPS to d8
427 dataset. STREAM identifies 4 main branching points among all single cells (Figure 6A). The
428 early differentiation days for all conditions were mixed in pseudotime demonstrating a similar
429 transcriptional state at the initial phases of IPS commitment when compared in the context of d6

430 and d8 cells. Similar to the UMAP projection (Figure 3B,C), STREAM inferred mixing of two iPS
431 stages with cells from the d2 treatments. STREAM inferred three main branching points which
432 separated based on the most differentiated cells from each condition. To reveal the
433 transcriptional states that contribute to the branching, we carried out differential expression
434 analysis to identify highly expressed genes within a branch that are differentially expressed
435 compared to all other branches, termed leaf gene detection analysis in STREAM (see methods)
436 (Figure 6B, Table S6). For the zero IPS branch point, pluripotent-specific genes including
437 *DNMT3A* and *POU5F1/OCT4* were highly differentially expressed. For the branch of cells
438 exposed to BMP4 (SC), mesoderm markers including *PITX2* were identified as differentially
439 expressed. For the BMP+IWP2 branch, placental associated genes were identified including
440 *RAB31*, a small GTPases highly expressed in placenta that regulates intracellular membrane
441 trafficking of ligand bound *EGFR* (Chua and Tang, 2014). For the TS branch, significantly
442 differentially expressed genes were associated with trophectoderm commitment and
443 implantation and preeclampsia including *GATA3*. We performed a Gene Ontology enrichment
444 analysis for biological processes using the top 300 significantly differentially expressed leaf
445 genes for the TS branch. Significantly overrepresented GO categories included placental
446 development, female pregnancy, viral processes (Figure 6 C,D).
447
448 **Human endogenous retrovirus derived genes participate in a regulatory subnetwork**
449 **specifically within iPS-derived trophoblast stem cells**
450
451 We found the set of genes involved in viral-like processes are upregulated in the later TS
452 differentiation timepoints (Figure 6C). This echoes an increasing body of evidence
453 demonstrating that endogenous retroviruses play a fundamental role in the evolutionary
454 diversification of the mammalian placenta (Johnson, 2019). Therefore, we asked what is the
455 expression of human ERV-derived genes during trophectoderm specification? We found specific
456 upregulation of five ERV-derived genes in the most differentiated TS cell type (Figure 7A).
457
458 To understand the place of ERV-derived genes in the regulatory TS regulatory network, we
459 analyzed the participation of ERV-derived genes in a gene regulatory network. We built a
460 transcription factor and target gene network model using the Passing Attributes between
461 Networks for Data Assimilation (PANDA) algorithm to all clusters among the conditions. PANDA
462 integrates information from TF-sequence-motif data, gene expression and protein-protein
463 interaction (PPI) in a message-passing approach (Glass et al., 2013). Briefly, PANDA's
464 algorithm makes the assumption that genes targeted by the same TF are more likely to be co-
465 expressed than genes that are regulated by a different TF. Using a TF-sequence motif as initial
466 network model, the algorithm refines and updates the model by computing the congruence
467 between two measures: 1) Concordance in co-expression of genes presenting motif for a same
468 TF, and 2) Concordance in the regulatory profile of a set of TFs that interact with each other via
469 Protein-protein interaction data. Therefore, the likelihood of interaction between a TF and a
470 target gene is given by its edge weight, which states for the coordination of targets' expression
471 by a TF complex. PANDA outputs Z-score normalized edge weights, therefore positive and
472 negative weights can be seen as likely and unlikely interactions respectively. One network was

473 created for each previously identified cell cluster based on the expression profile of the top 100
474 cells in each cluster (based on PC1) with genes expressed in at least 10% cells.
475 We found that ERV interactions (Transcription Factors + ERV-derived genes) are prominent in
476 the most differentiated TS cluster and in no other cell clusters. We found regulatory interactions
477 for five ERV-derived genes (*ERVH48-1*, *ERV3-1*, *ERVMER34-1*, *ERVW-1* and *ERVFRD-1*) are
478 highly unique to the TS cluster network, while only 0 or 1 ERV-derived gene was found in the
479 other cell clusters (Figure 7B, S9A). Clusters SC and SC 3 (BMP4 cells) exhibited interactions
480 for gene *ERVH48-1*, and clusters BR 1, BR 3 and BR 4 (BMP4+IWP2 cells) presented the gene
481 *ERV3-1*. Interestingly, genes encompassed in ERV regulatory interactions of iPS derived TS
482 cells were enriched in biological pathways related to hormone metabolism, cell differentiation
483 and the immune system (adj. p value < 0.01) (Figure 7C, Table S7). In addition, we observed
484 ERV regulatory interactions are also enriched for categories such as placenta development,
485 trophectodermal cell differentiation and syncytium formation (Table S7). To understand the
486 significance of the ERV-subnetwork interactions, we next asked what is the strength of ERV
487 edges in the network context of the TS cluster? We found that weights of ERV regulatory
488 interactions were highly ranked among all network connections (Wilcoxon Rank Sum
489 pval=$1.6778e^{-31}$, permuted p value=0.001). In sum, we find that ERV-derived genes participate
490 in a regulatory subnetwork specifically within TS cells.
491
492 **Coronavirus associated host factors are expressed in iPS-derived TS cells and could**
493 **potentially model cytotrophoblast involvement in infection**
494
495 We next asked if these iPS-derived cells could be relevant to model the cytotrophoblast
496 involvement in infection. Infections of pregnant women can severely impact both the developing
497 fetus and pregnant women. Trophoblast cells at the maternal-fetal interphase play a crucial role
498 in the impact of these infections. Syncytiotrophoblasts are the barrier that protects the fetus from
499 some infections (Parry et al., 1997). The permissiveness of trophoblast cells to a specific
500 infection largely determines if an infection reaches the fetus. Pathogens such as Toxoplasma
501 gondii, Rubella, herpesvirus, cytomegalovirus, and Zika virus can cross the placenta, infect the
502 fetus and cause congenital disease (Stegmann and Carey, 2002; Coyne and Lazear, 2016). It is
503 still too soon to know the complete impact of the severe acute respiratory syndrome coronavirus
504 2 (SARS-CoV-2) and the associated coronavirus disease 2019 (COVID-19) on pregnant women
505 and fetus. But it has been established that the placenta can be infected by SARS-CoV-2
506 (Vivanti et al., 2020). Emerging data suggest that placental SARS-CoV-2 infections are rare
507 (Vivanti et al., 2020), perhaps because the CoV-2 host entry substrate angiotensin-converting
508 enzyme 2 (ACE2) and the serine protease TMPRSS2 are minimally expressed in the placenta
509 (Pique-Regi et al., 2020). Although ACE2 and TMPRSS2 co-expression are rare in the maternal
510 fetal interface, alternative entry receptors such as BSG (Aguiar et al., 2020). When considering
511 expression analysis of confirmed and potential SARS-CoV-2 host factors, Singh et al reported
512 that the cytotrophoblast is the cell type most susceptible to coronavirus infection in the first
513 trimester placenta while confirming that *ACE2* and *TMPRSS2* are lowly expressed (Singh et al.,
514 2020).
515

To evaluate the potential tropism of SARS-CoV-2 in iPS-derived TS cells, we profiled SARS-CoV-2 and coronavirus associated factors. Overall, iPS-derived TS cells expressed coronavirus associated factors in similar patterns compared to ex vivo cytotrophoblasts (Singh et al., 2020), (Pique-Regi et al., 2020). *ACE2* and *TMPRSS2* are minimally expressed (Figure 8A, Table S8) and putative alternative entry receptors *BSG*, *HSPA5* and *DPP4* are more expressed. Other potential entry proteases *CTSB*, *CTSL* and *FURIN* are all expressed at a higher level (Figure 8B, Table S8). The placenta also protects the fetus from infection by the active transfer of maternal antibodies to the fetus. Syncytiotrophoblasts transfer protective humoral maternal IgG to the fetus via the neonatal Fc receptor encoded by the FCGRT gene (Leach et al., 1996). Consistent with the findings that iPS-derived TS cells are transcriptionally similar to cytotrophoblasts and syncytiotrophoblasts, iPS-derived TS cells also express FCGRT. In sum, iPS-derived TS cells are highly transcriptionally similar to *ex vivo* cytotrophoblasts and syncytiotrophoblasts with respect to placental host factors involved in infection. Therefore, iPS-derived TS cells may be useful to model the involvement of human cytotrophoblasts in infection.

Discussion

Here we demonstrate that primed iPS cells are specified to trophectoderm without BMP4 and instead with EGF, TGF-β inhibitors, a WNT activator and Valproic acid induced HDAC inhibition. IPS-derived TS cells are highly transcriptionally similar to cytotrophoblasts found in human placenta. TS cells proliferate rapidly, display very variably sized nuclei, readily become multinucleated as observed in early post-implantation trophoblast and strongly express extravillous and syncytiotrophoblast associated proteins (HLA-G, HCG-β) in sub populations. After eight days of specification, iPS-derived TS cells display a uniquely active transcriptional network of human endogenous retroviruses similar to *in vivo* trophoblast. We demonstrate that in hypoxic conditions, iPS-derived TS cells create 3D villi-like structures, conditions which will likely be useful for future 3D organoid models.

The ability of primed human stem cells to differentiate into true trophectoderm lineages is under debate. Here we demonstrate that primed human stem cells retain a broad potency to generate extra embryonic lineages which are transcriptionally highly similar to placenta. Our results are consistent with the model that the primed state is on a continuum of potency, rather than a fixed and restricted state (Weinberger et al., 2016). Methodological differences and the frequent use of BMP4 likely underlie previously reported differences in the ability of primed and naive pluripotent stem cells to generate TSCs. We confirm that the BMP4 model can generate cells with key features of trophoblast, but in the defined conditions used here BMP4 specified cells are less similar to *in vivo* cytotrophoblasts. Following the recent derivation of human TSCs from primary cells (Mischler et al., 2019; Okae et al., 2018), several attempts have been made to derive TSCs from ES/iPS from a range of pluripotent states. Mischler et al. 2019 used primed ES cells with a hybrid protocol that combined a three day treatment of BMP4 and SB431542 with long term culture in TM4 media (CYM554 a S1PR3 agonist, CHIR99021, A83-01, and FGF10) to generate a CDX2+ TSC-like population (Mischler et al., 2019; Okae et al., 2018). When transferred to TSCM media, these cells lost CDX2 expression and adopted characteristics of TSCs. Omitting TM4 media and directly transferring BMP4 treated cells to

TSCM resulted in a gradually enriching population of epithelial cells comparable with their CDX2- TSC-like population. Interestingly, this process mirrors the differentiation progression we observed when we directly treated primed iPS with TSCM. Our data suggest that BMP4 is not involved in this transition. TSC-like cells have also been generated from primed iPS cells using a micromesh culture system with primed pluripotent stem cell media followed by culture in TSCM (Li et al., 2019) although the comparison of these cells to *in vivo* placental cells is less extensive. In contrast, Dong et al. 2020 reported that only naive and not primed ES could generate TSC-like cells by direct exposure to TSCM (Dong et al., 2020). Human expanded potential stem cells (EPSCs) can also generate TSC-like cells when differentiated in TSCM (Gao et al., 2019). Interestingly, this group was unable to generate TSC-like cells using naive or primed ES cells under similar conditions. Here, we demonstrate that the continuum of pluripotency present in the primed state readily allows for specification to trophectoderm.

The robust differentiation of primed iPS cells to trophectoderm allows for modeling of placental diseases from patient-specific cells without first trimester primary tissue. Placental adaptive responses to insults can alter fetal development and program the fetus for later disease including diabetes and neurodevelopmental disorders (Myatt, 2006). As both the protective barrier and a component of the innate immune system, cytotrophoblasts and their progeny play important roles in the outcome of insults (Guleria and Pollard, 2000). Recent evidence that Covid19 leads to a hyperimmune state has implications for placental biology and fetal development, potentially analogous to the effects of other hyperimmune states during pregnancy (Al-Haddad et al., 2019). Genetic variation among individuals influences how a pathogen affects the fetus. iPS models of this genetic variation could be harnessed for therapeutic intervention. We anticipate that iPS-derived TS cells will be a powerful model to study the placental contribution to disease and the genetic regulation of human placental evolution.

Acknowledgements

Supported by funding from the Lieber Institute for Brain Development and a NARSAD Young Investigator Grant from the Brain & Behavior Research Foundation to J.A.E. The authors are grateful for the financial support of the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Financing code 001 (Andre Barbosa, Veronica Euclydes and Arthur Feltrin). We thank Bill Ulrich for the single cell RNAseq browser.

Declaration of Interests

The authors declare no competing interests.

598 **Figures and Figure Legends**
599
600 **Figure 1: Derivation Of Trophoblast Stem Cell Like (TS) Cells From Human Pluripotent**
601 **Stem Cells.**
602 **A.** Schematic of protocol used to derive TS cells from human pluripotent stem cells (PSCs) by
603 activating wingless/Integrated (WNT) and epidermal growth factor (EGF) and inhibiting
604 Transforming Growth Factor Beta (TGFb) histone deacetylase (HDAC) and Rho-associated
605 protein kinase (ROCK).
606 **B-G.** Representative phase contrast images of differentiating TS cells (ES line H1). (B) Images
607 of differentiating cells on days 2 (B), 4 (C), and 6 (D) after addition of differentiation (TS) media.
608 **E,F.** Images of TS cell morphology 24 (E) and 72 (F) hours following passage (P1) on day 6. (G)
609 Mature TS lines generated from ES line H1 after 10 passages.
610 **H.** Immunofluorescence for VIM (magenta), KRT7 (red), TP63 (green), and Hoechst 33342
611 (blue) in TS cells after 1 passage (P_0)(iPS line LIBD7c6).
612 **I-K.** Confocal images of KRT7 (red), TP63 (green), and DAPI (blue) for TS cells generated from
613 passage 11 H1 (I) and 2014.06 (J) and passage 32 LIBD1C8 (K).
614 All phase contrast images scale bars indicate 100 µm. All immunofluorescence images scale
615 bars indicate 50 µm.
616
617 **Figure 2: iPS-derived TS Cells Differentiate Into Syncytial Cytotrophoblast, Extra Villous**
618 **Cytotrophoblasts and Villi-like Structures.**
619 **A,B.** Immunofluorescent images of mature TS line LIBD1C8 stained for ZO-1 (red), KI-67
620 (green), and DAPI (grey).
621 **C,D.** Immunofluorescent images of mature TS line LIBD1C8 stained for DAPI (red) and hCG (C)
622 or HLA-G (D) (grey).
623 **E-H.** Villi like structures spontaneously form when cultured in 5% O_2. (E) Three dimensional
624 reconstruction of a "villi" from mature TS line 2014.06 using immunofluorescence images of
625 nuclei visualized with Hoechst 33342. Virtual sections at the base (H), midpoint (G), and top (F)
626 of the structure.
627 All scale bars indicate 50 µm.
628
629 **Figure 3: Temporal Single Cell RNA Sequencing Of iPS To Trophoblast Differentiation**
630 **A.** Experimental schematic of trophoblast differentiation experiment. HiPS cells were plated and
631 allowed to adhere to culture dishes for 24 hours before addition of differentiation media. Cells
632 were collected for single-cell RNA sequencing at indicated time points.
633 **B,C.** Dimensionality reduction for the most variable genes across all cells with Principal
634 Component Analysis (PCA) and Uniform Manifold Approximation and Projection (UMAP)
635 embedding of 9,821 single cell transcriptomes from three distinct media treatments with
636 groupings based on (B) sample identity or (C) gene expression clusters calculated by k nearest
637 neighbors using the Euclidean distance of the 30 first PCs which identifies 19 clusters. Cells
638 from the BMP4 condition almost exclusively formed four clusters termed SC (Stromal cell
639 indicated in blue). Cells from the BMP4+ IWP-2 condition were found predominantly in a group
640 of five heterogeneous clusters termed BR (BMP responsive indicated in purple). iPS indicated in
641 green and TS clusters are indicated in orange.

D. Plot of the proportion of cells from each differentiation condition which are present in the final k-means clusters TS, TS6, BR and SC. Color represents the differentiation condition scale in 3B BMP4+IWP2: purple, BMP4: blue, iPS: green, TS: orange. Cell proportions for all clusters can be found in supplemental table S1.
E-H. Single cell visualization of gene expression of selected markers of HiPS cells (E), placental stromal cells (F), and trophoblasts (G,H) on UMAP projections. Grey dots indicate cells with no expression and blue scale indicates normalized expression.
I. Dot plot displaying differentially upregulated marker genes for the most mature cell clusters from each treatment SC (BMP4), BR (BMP4+IWP2), and TS 6, TS (TS media) compared to the average expression of all other clusters. Size of dot indicates the percent of cells with expression and intensity of the blue indicates the average expression level.

Figure 4: Trophoblast Differentiations Are Enriched For First Trimester Placental Cell Types
A,B. Gene set enrichment analysis of most mature iPS-derived cell clusters (SC (Stromal Cell, BMP4), BR (BMP responsive, BMP4+IWP2), and TS 6, TS (Trophoblast stem cell media) to cell types identified in previous single cell sequencing of first trimester placenta. Cell types
C. Composite heatmap of normalized gene expression of the top unique markers of placental cell types of fetal origin identified in Vento-Tormo et al. 2018 and polar trophectoderm markers (top 50) identified in Petropolus et al. 2016.

Figure 5: Primed iPS Specification To Trophoblast Stem Cells Initiates With TFAP2A Rewiring Without Activation Of Naïve Stem Cell Programs
A. Dotplot displaying BMP signaling in each cellular cluster organized by the treatment identity of the majority cells. See table S3 for source data.
B. Dotplot displaying markers of naïve and primed human pluripotent stem cells for BMP4, BMP4 + IWP2, and TS cell differentiations by sample. See table S3 for source data.
C. STREAM pseudotime trajectory of HiPS differentiation for HiPS and TS D2 cells visualizing cells proportion along the smoothed pseudotime axis (orange represents TS, green represents hIPS). Nodes are labeled with numbers 0-4. Branches are defined as the cells between 2 nodes (Branch 1,4). Bottom: Top ranked (by STAT) genes with expression significantly correlated with the pseudotime transition of the branches. MEIS2, TFAP2A, and less differentiated TS d2 to more differentiated TSd2 (S0-S1) KRT18 (STAT 0.692). Ranks 2-4 for these branches is found in figure S4 C-D. Additional graphs showing the branch assignment with cells and a cell level representation of the pseudotime with the clusters Figure S9 and table S4.

Figure 6: STREAM Pseudotime Analysis Of Trophoblast Differentiations Reveals Upregulation Of Stromal Mesenchyme Upon BMP4 Treatment and Trophectoderm Transcriptional Programming Upon TS Treatment.
A. STREAM pseudotime trajectory of HiPS differentiation including all samples highlighting the distribution of each previously described cell cluster along pseudotime.
B. Average gene expression of example genes which are significantly differentially expressed for each branch compared to other branches. S0 initial differentiation branch: 1284 significant genes, POU5F: p value = 1.032932e-297 by Kruskal–Wallis H-test. S4 branch dominated by BR

686 (BMP4+IWP2 cells): 538 significant genes, RAB3 p value = 0 by Kruskal–Wallis H-test. S2
687 branch. S2 branch dominated by SC (BMP4+ cells): 1395 significant genes, PITX2 p value = 0
688 by Kruskal–Wallis H-test and top ranked. S3 branch dominated by TS cells. 3855 significant
689 genes, GATA3 p value = 0 by Kruskal–Wallis H-test and top ranked gene. The complete list of
690 significant genes see Supplemental Table S5.
691 **C.** Upset plot of shared genes within significant Gene Ontology enrichment terms generated
692 from the top 300 leaf genes for branch 1-3 (specification to TS). Gliogenesis p=7.92e-06,
693 Leukocyte activation involved in immune response p=1.72e-05, Reproductive system
694 development p=4.05e-05, Female pregnancy p=7.95e-05, Viral Process p=.0004, Blood Vessel
695 Morphogenesis p=.003, Placenta Development p=.008
696 **D.** Gene expression of CTSB along the pseudotime trajectory. CTSB is a significant leaf gene
697 for branch 1-3 that is annotated as associated with all GO processes listed in (6C).
698
699 **Figure 7: Endogenous Retrovirus-derived Genes Participate In A Gene Regulatory**
700 **Subnetwork Uniquely In TS Cells.**
701
702 **A.** Dotplot of HERV expression in most mature cell clusters from each treatment and the HiPS
703 cluster 1. Size of ball indicates the percent of cells expressing the gene and color represents the
704 z-normalized expression value.
705 **B.** Cytoscape representation of the top transcription factor partners of ERVs detected in TS and
706 SC. 164 TFs and 6 Targets. See Supplemental table S7.
707 **C.** Visual Representation of GO terms from most significant partners of ERVs shown in panel B
708 (p<.0001). Specifics of programs and settings to obtain these results. Cytoscape 3.7.2, cluego
709 v2.5.4, cluepedia version v1.5.4, go term fusion, levels 3-8, kappa .35, graph of highest
710 significance within each group, input of 164 TFs from panel B supplemental table with analysis
711 output and details of GO analysis includes all go terms off the
712
713 **Figure 8: Coronavirus Associated Host Factors Are Expressed In iPS-derived TS Cells**
714
715 Dotplot of Coronavirus associated host factor expression in the iPS-derived TS cell clusters TS
716 and TS6. Size of ball indicates the percent of cells expressing the gene and color represents the
717 average expression value.
718 **A.** Coronovirus associated host factors expressed in 0-10% of cells.
719 **B.** Coronavirus associated host factors expressed in >10% of cells.
720
721 **Supplemental Tables**
722 Table S1: Number of cells from each condition/day that are assigned to the cluster
723 Table S2: Differentially expressed genes in each cluster where a single cluster is compared to
724 all other clusters
725 Table S3: Differentially expressed genes analyzed by pairwise comparisons of one cluster
726 compared to another cluster
727 Table S4: Expression analysis of naive, primed and BMP signaling during differentiation
728 Table S5: Stream pseudotime analysis of transition and differentially expressed leaf genes for
729 iPS to TSd2

730 Table S6: Stream pseudotime analysis of transition and differentially expressed leaf genes for
731 all cells
732 Table S7: ERV gene regulatory network and associated GO enrichment.
733 Table S8: Expression analysis of coronavirus associated factors.
734
735

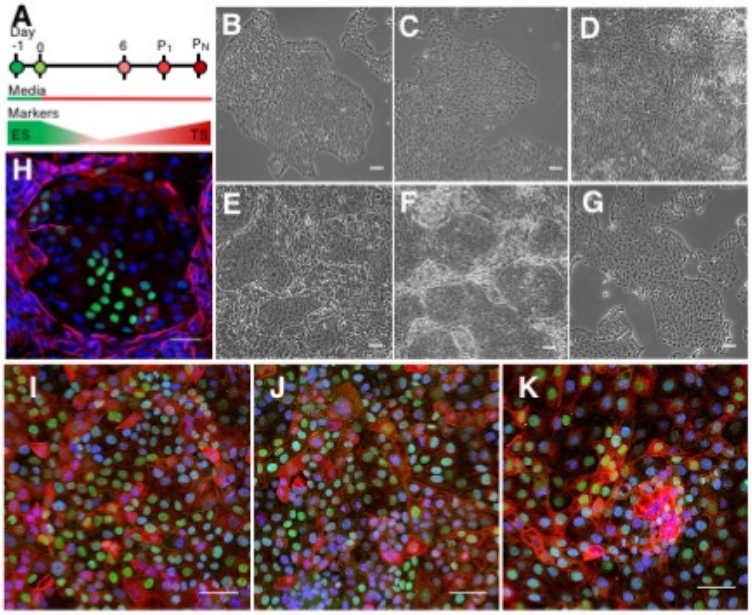


Fig. 1: Derivation of trophoblast stem cell like (TS) cells from human pluripotent stem cells.

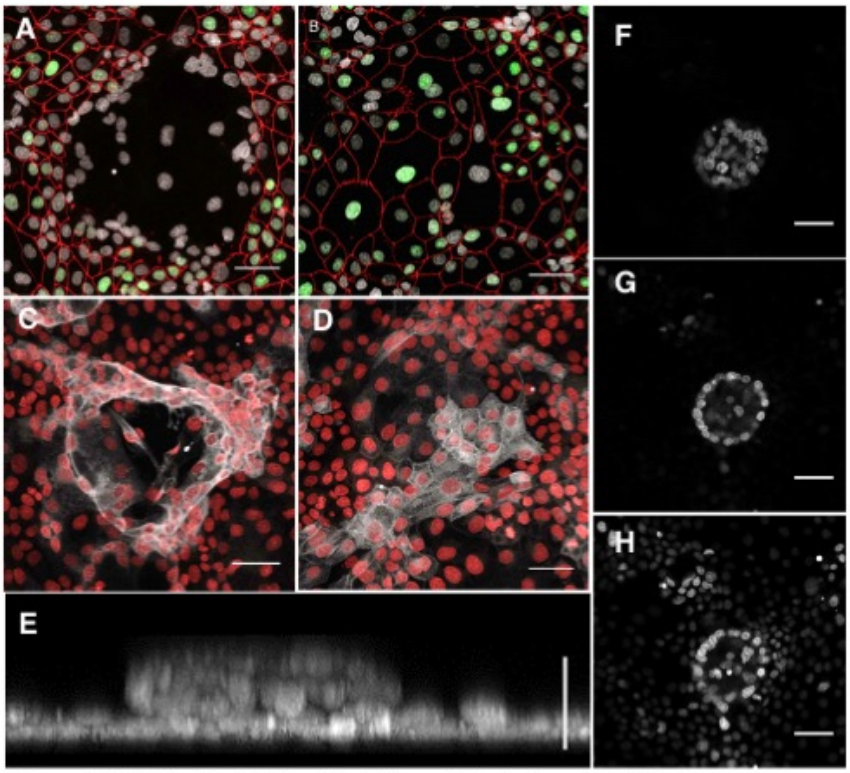


Figure 2: iPS-derived TS cells differentiate into syncytial cytotrophoblast, extra villous cytotrophoblasts and villi-like structures.

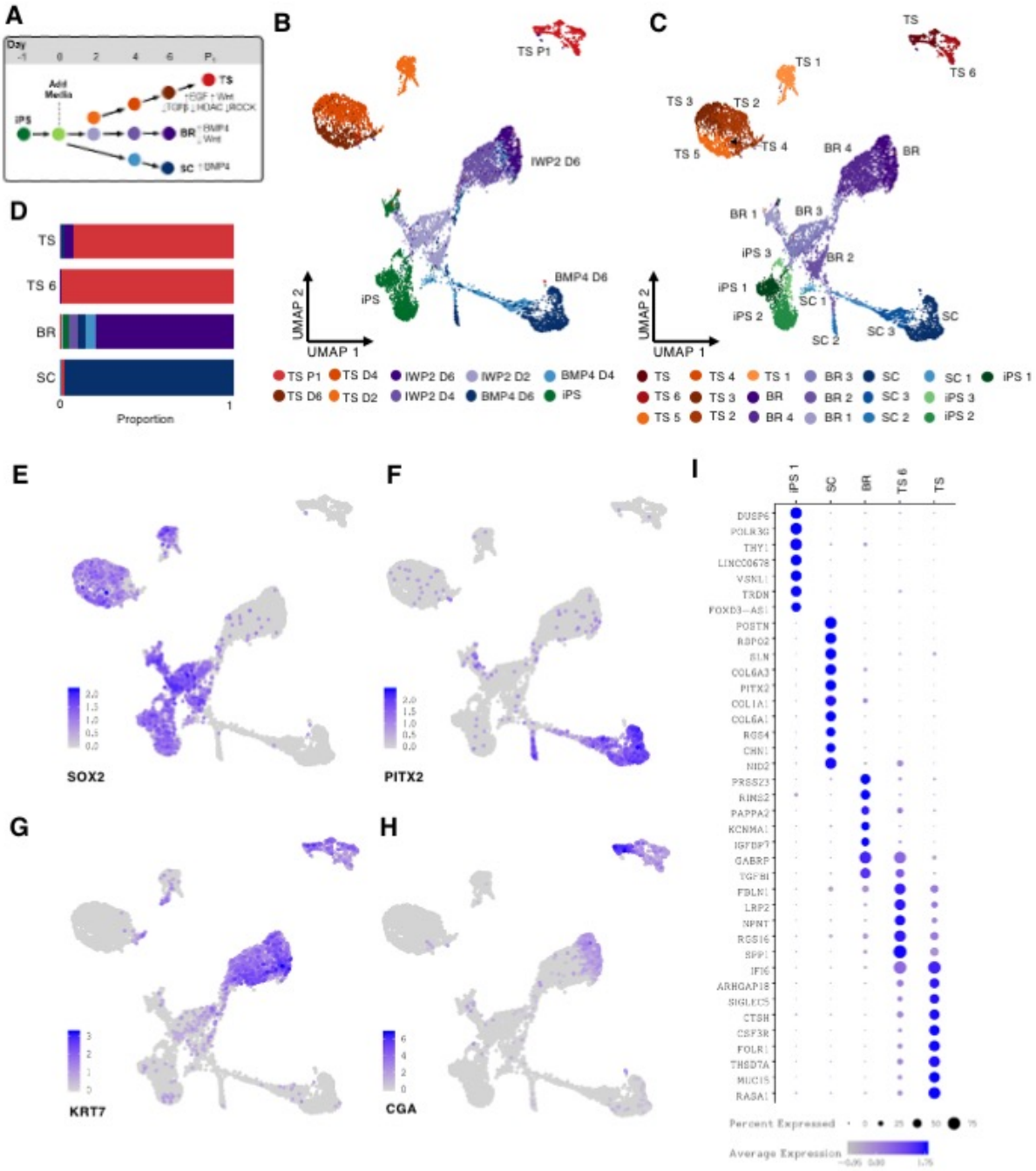


Figure 3: Temporal single Cell RNA Sequencing of iPS to trophoblast differentiation

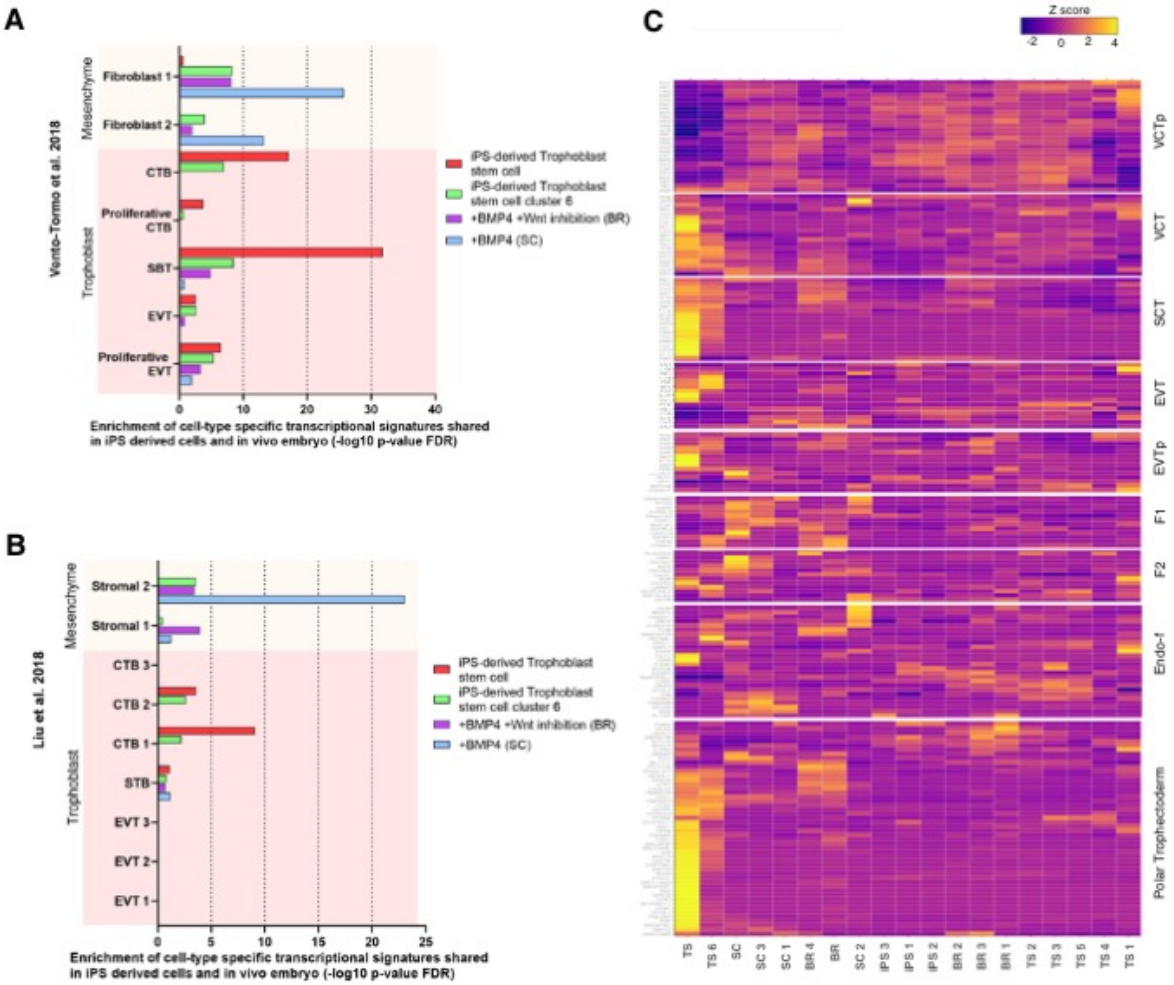


Figure 4: Trophoblast differentiations are enriched for first trimester placental cell types.

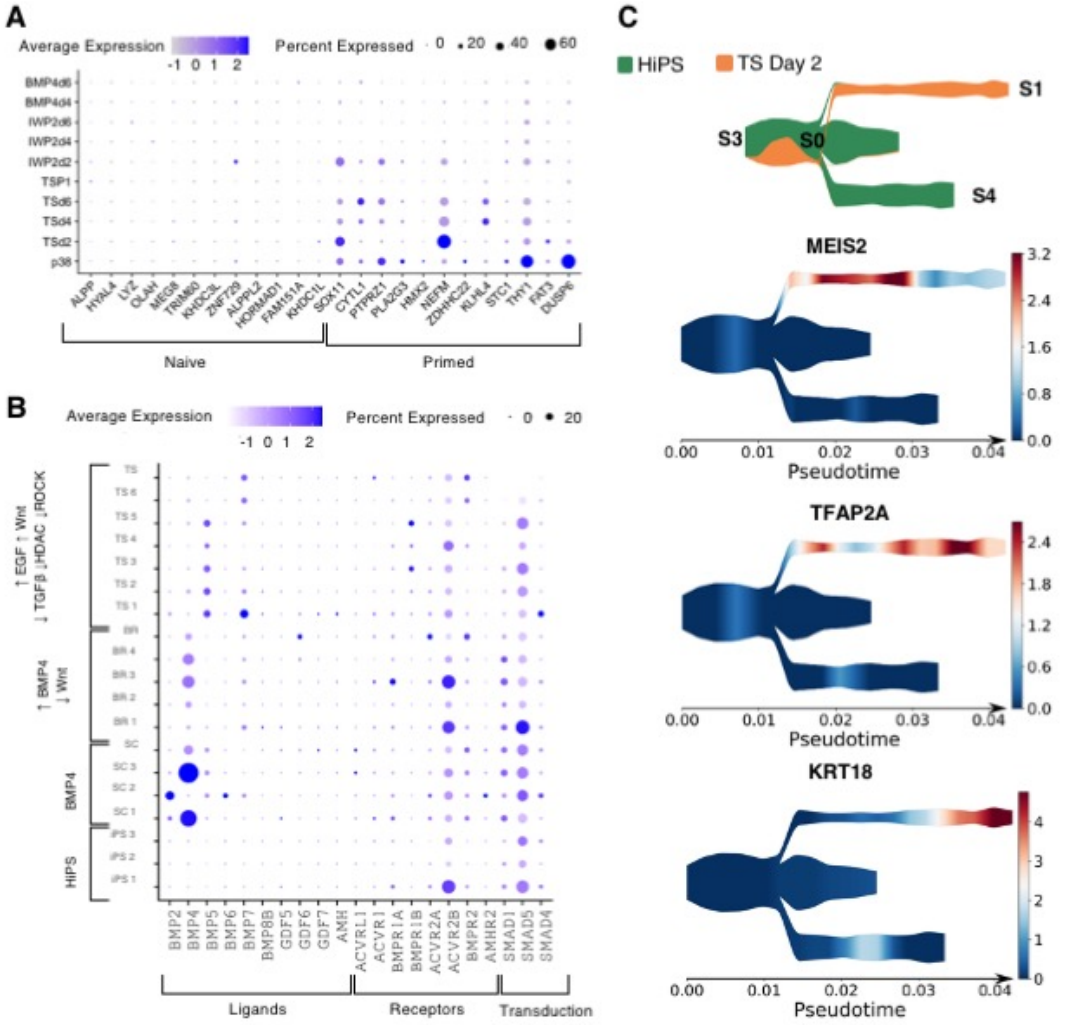


Figure 5: Primed iPS specification to trophoblast stem cells initiates with TFAP2A rewiring without activation of naïve stem cell programs

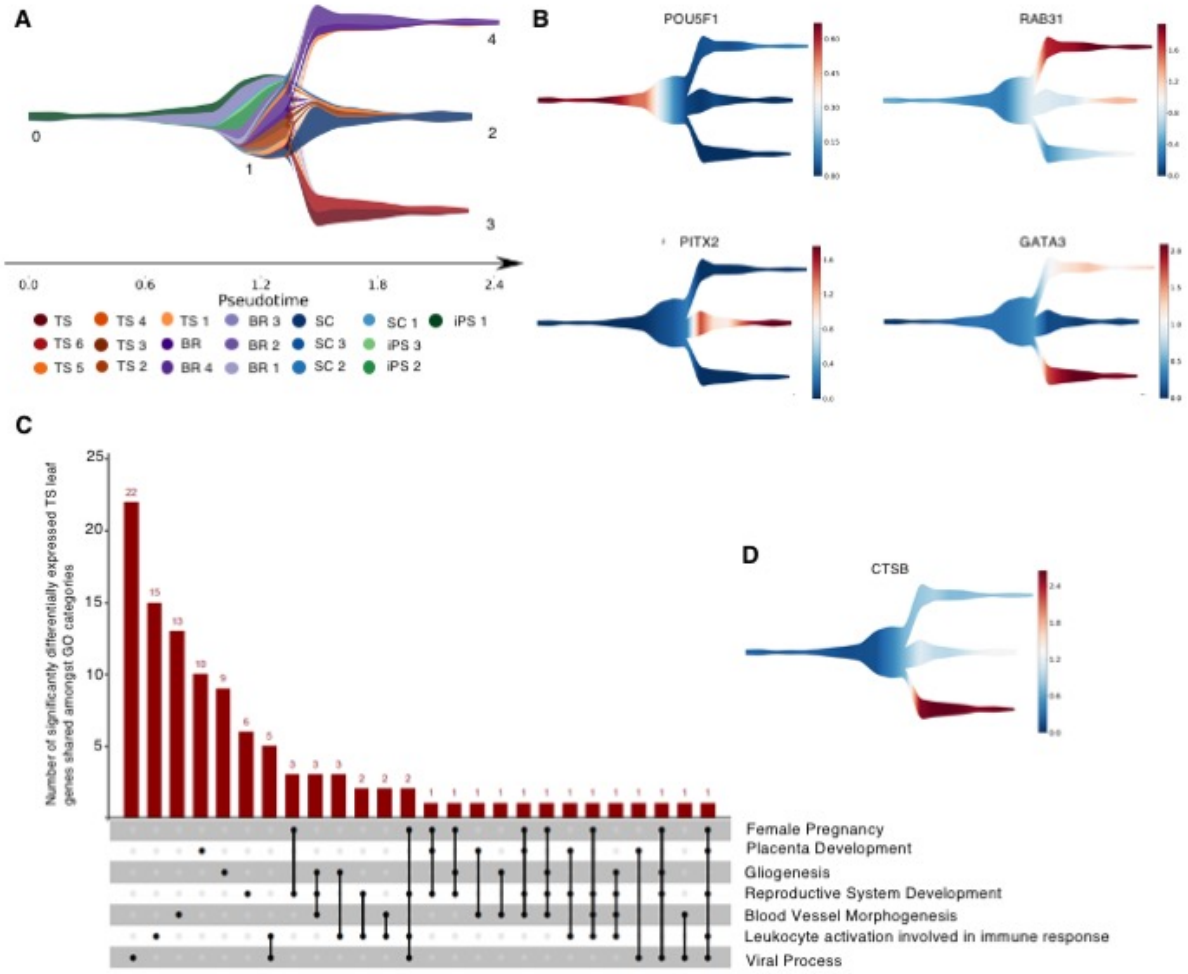


Figure 6: STREAM pseudotime analysis of trophoblast differentiations reveals upregulation of stromal mesenchyme upon BMP4 treatment and trophectoderm transcriptional programming up TS treatment.

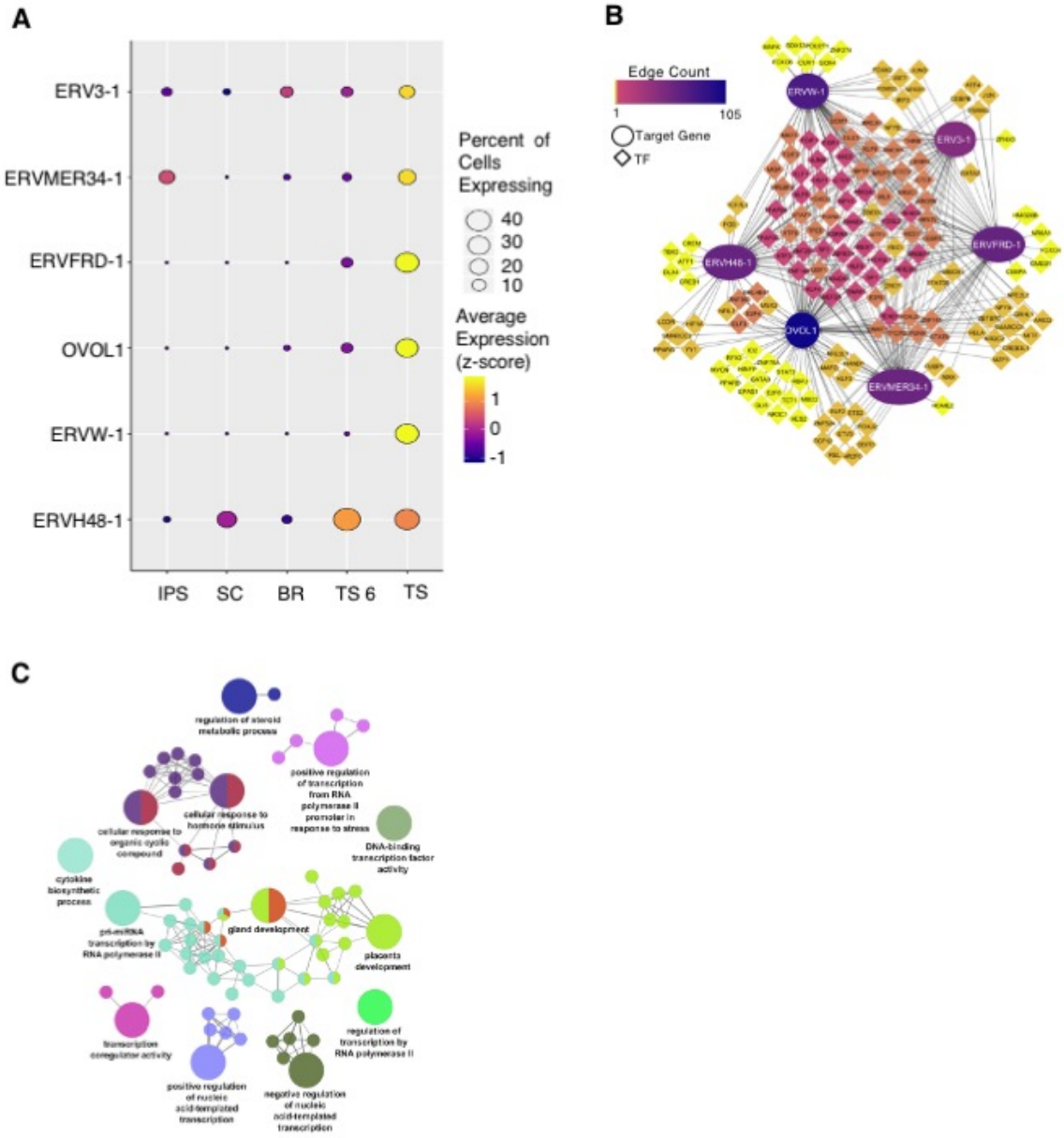


Figure 7: Endogenous retrovirus-derived genes participate in a gene regulatory subnetwork uniquely in TS cells.

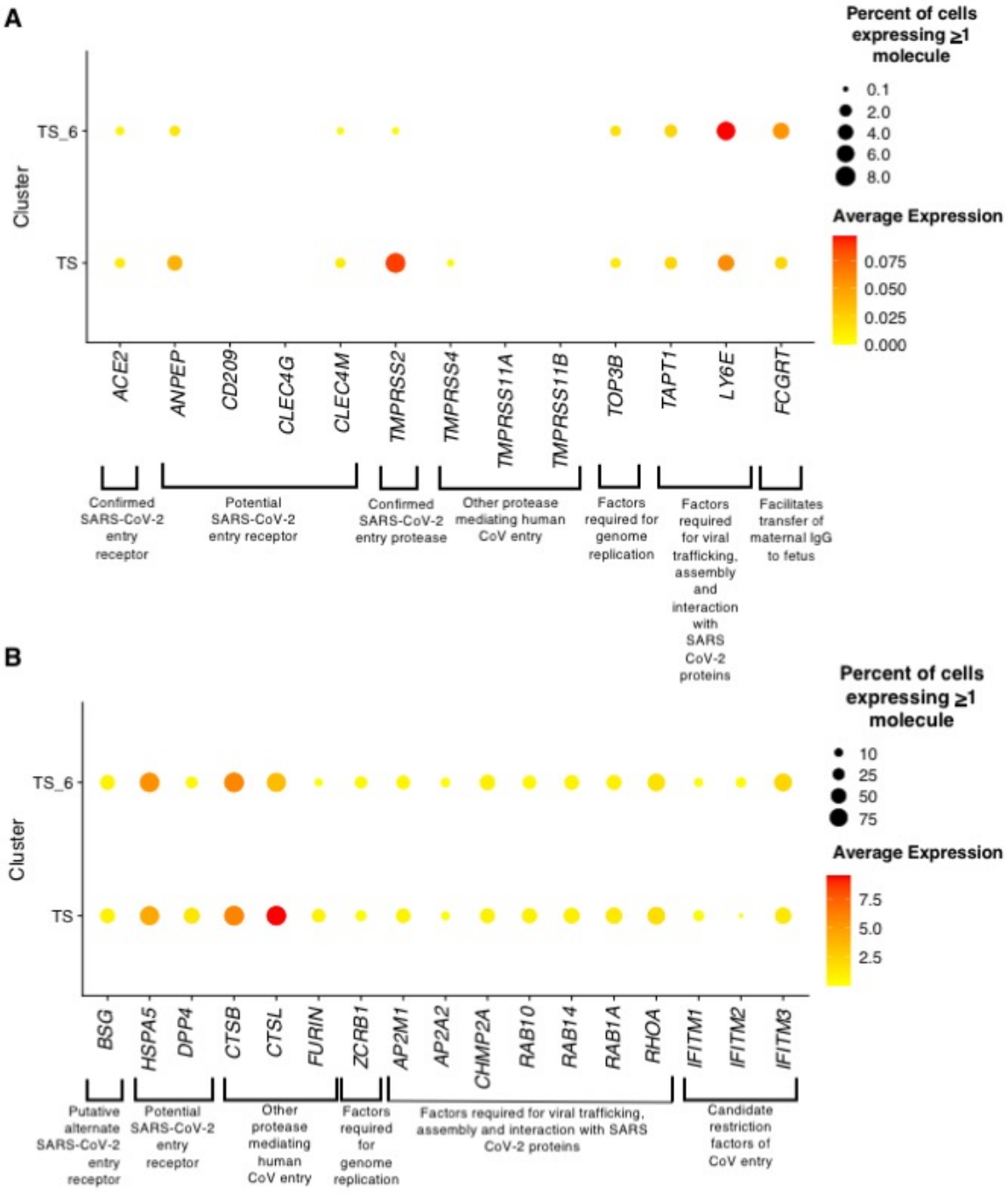


Figure 8: Coronavirus associated host factors are expressed in iPS-derived TS cells

Supplemental Figures

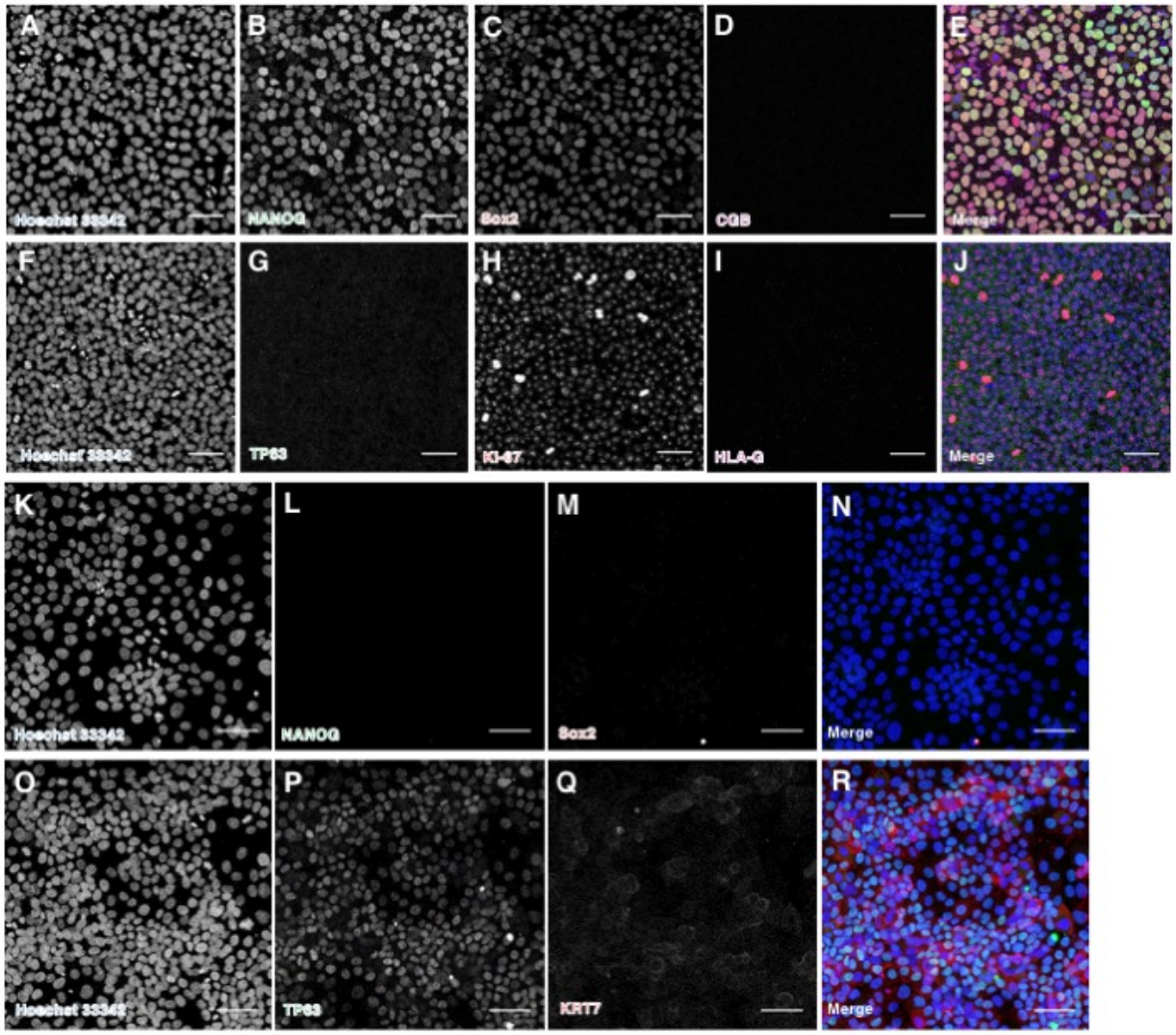


Fig. S1: Differentiated TS cells lose expression of pluripotency markers while gaining expression of trophoblast specific markers.

(A-E) Immunofluorescent images of HiPS line LIBD7C6 stained for Hoechst 33342 (A), NANOG (B), and Sox2 (C), CGB (D), and in composite (E).

(F-J) Immunofluorescent images of HiPS line LIBD7C6 stained for Hoechst 33342 (F), TP63(G), and KI-67 (H), HLA-G (I), and in composite (J).

(K-N) Immunofluorescent images of TS line 2014.06 stained for Hoechst 33342 (K), NANOG(L), and Sox2 (M), and in composite (N).

(O-R) Immunofluorescent images of TS line 2014.06 stained for Hoechst 33342 (O), TP63 (P), KRT7 (Q), and in composite (R).

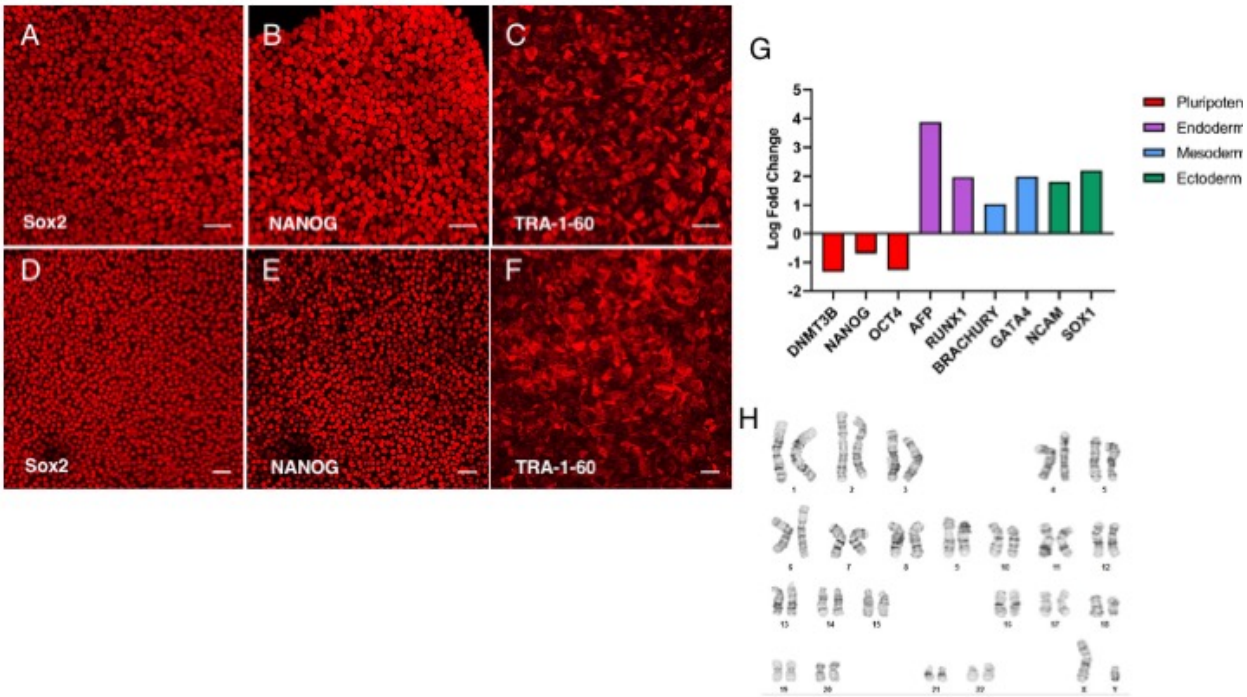


Fig. S2: iPS cell line characterization.

(A-C) Immunofluorescent images of HiPS line LIBD1C8 stained for Sox2 (A), NANOG (B), and TRA-1-60 (C).

(D-F) Immunofluorescent images of HiPS line 2014.06 stained for Sox2 (D), NANOG (E), and TRA-1-60 (F).

(G) Relative expression levels of pluripotency and germ layer specific markers in spontaneously differentiated cells of line 2014.06 normalized to their respective expression levels in undifferentiated HiPS of the same cell line determined by qPCR.

(H) Karyotype analysis of 2014.06

All scale bars 50 um.

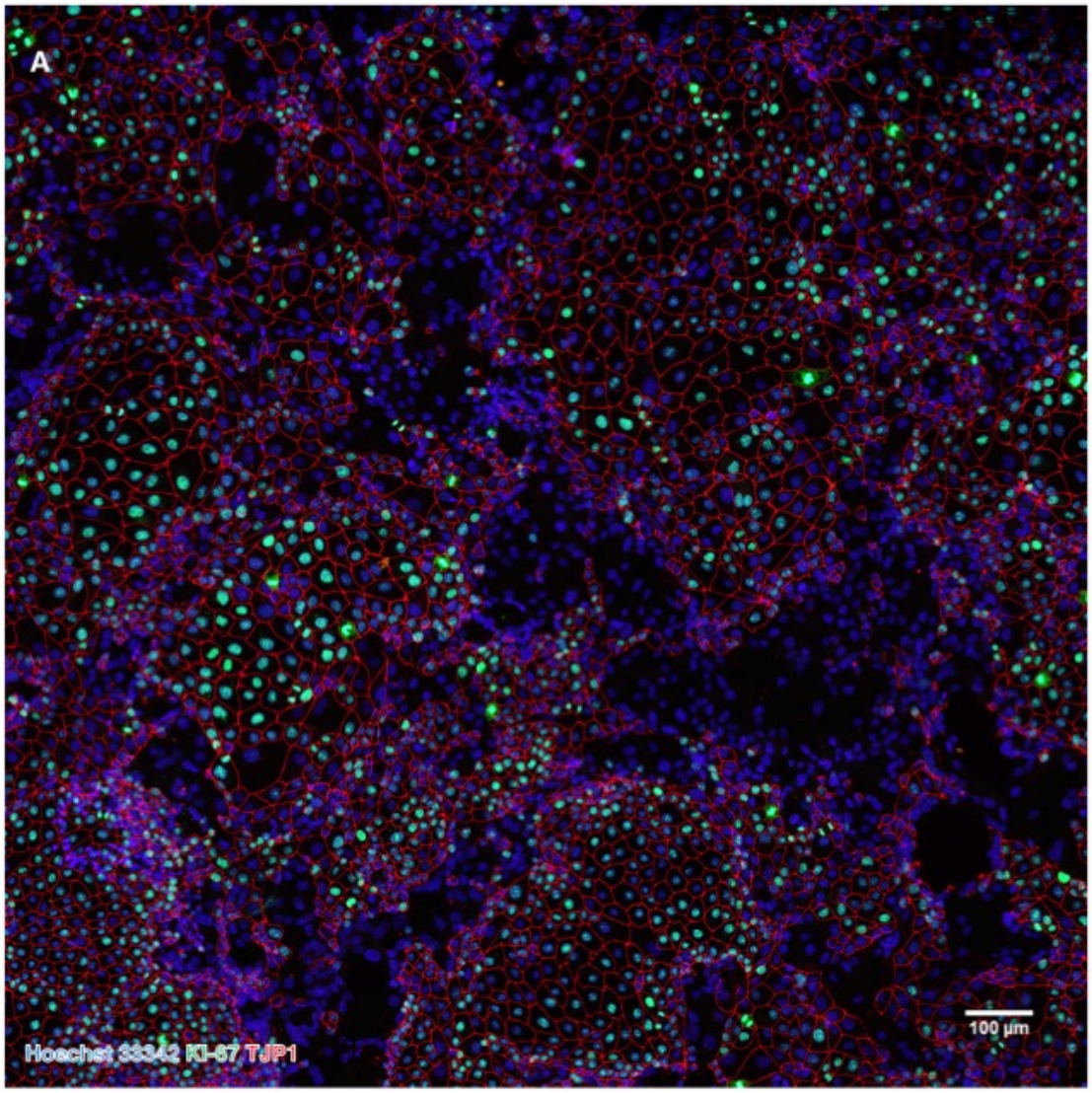
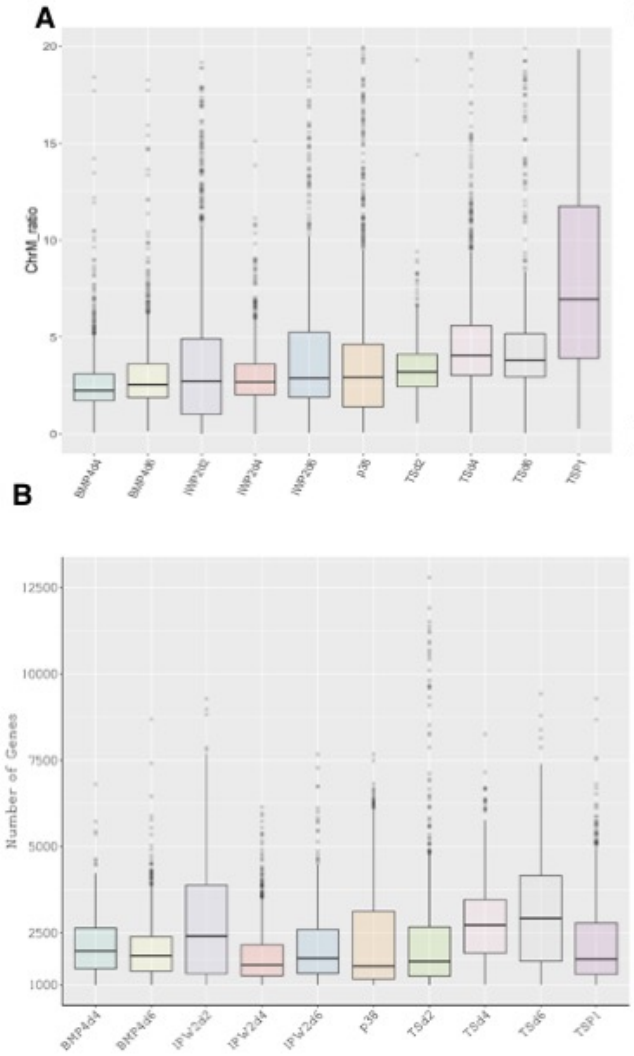


Fig. S3: Tile Scan Confluent TS LIBD1C8 P20 demonstrating extensive syncytium

(A) Tile scan of confluent TS line LIBD1C8 passage 20 grown on Collagen IV and stained for Hoechst 33342 (blue), KI67 (green), and TJP1 (red).



C

Sample	# of Cells	% of Total Cells
iPS	1572	16.01
BMP4 D4	827	8.42
BMP4 D6	1234	12.56
IWP2 D2	1307	13.31
IWP2 D4	1154	11.75
IWP2 D6	598	6.09
TS D2	563	5.73
TS D4	1360	13.85
TS D6	622	6.33
TS P1	584	5.95
Total	9821	100%

D

Cluster	# of Cells	% of Total Cells
BR 4	1374	13.99
TS 2	895	9.11
BR 3	885	9.01
SC	860	8.76
iPS 2	790	8.04
TS 3	651	6.63
BR	647	6.59
BR 2	531	5.41
TS 1	530	5.40
iPS 1	446	4.54
TS 5	436	4.44
SC 3	342	3.48
TS 6	309	3.15
BR 1	277	2.82
TS	272	2.77
SC 1	214	2.18
iPS 3	185	1.88
SC 2	135	1.37
TS 4	42	0.43
Total	9821	100%

Fig. S4: Single Cell RNAseq Quality Control.

(A) Boxplot of Mitochondrial Mapping rate per droplet in the final data set.

(B) Boxplot of number of detected genes per droplet in the final data set.

(C-D) Summary tables of cell number and percent composition by experimental sample (C) and cluster identity (D).

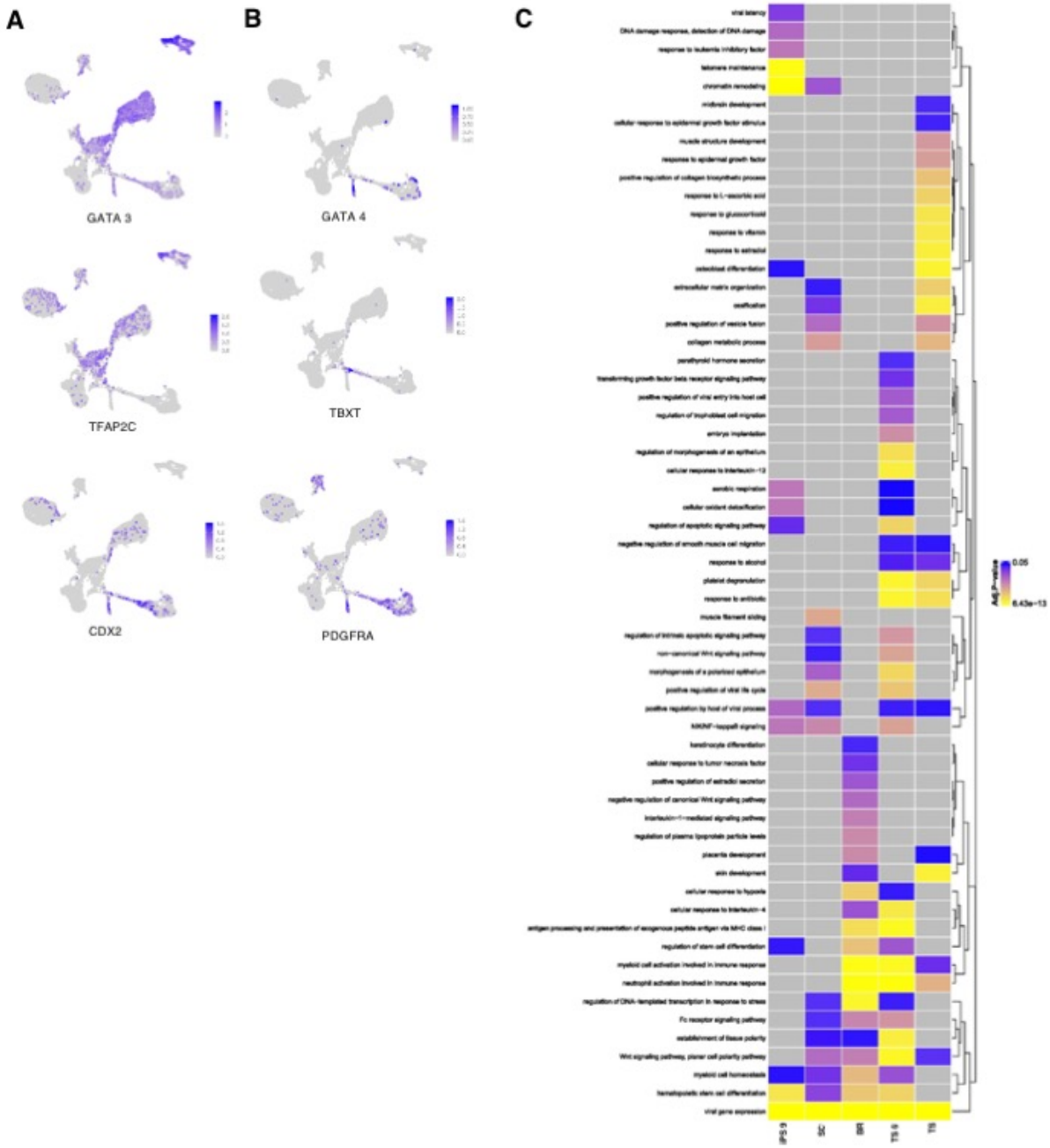


Fig. S5: BMP4 model comparison.

(A-B) Highlighted expression of trophoblast (A) and mesoderm (B) associated transcription factors at single cell resolution.

(C) Heatmap of selected GO terms (p-value <0.05) generated from the 300 most highly expressed genes for each of the indicated mature cell types.

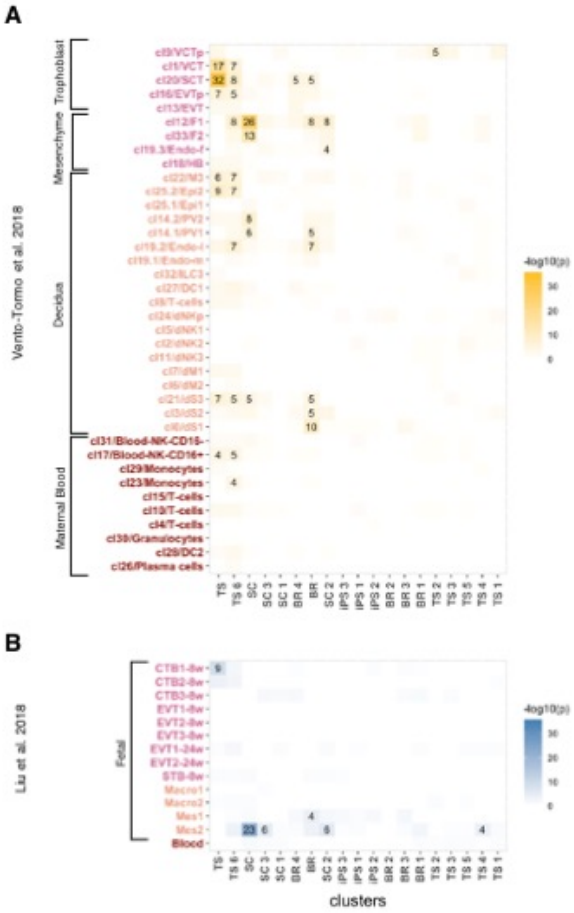


Fig. S6: Trophoblast differentiations are enriched for first trimester placental cell types.

(A,B) Gene set enrichment analysis of HiPS cell differentiations highlighting enrichment of cell clusters to cell types identified in first trimester placentas. Color and number in each square indicates the –log10(p) Bonferroni Corrected hypergeometric test for enrichment of cell-specific genes in IPS-derived clusters and previous placental publications.

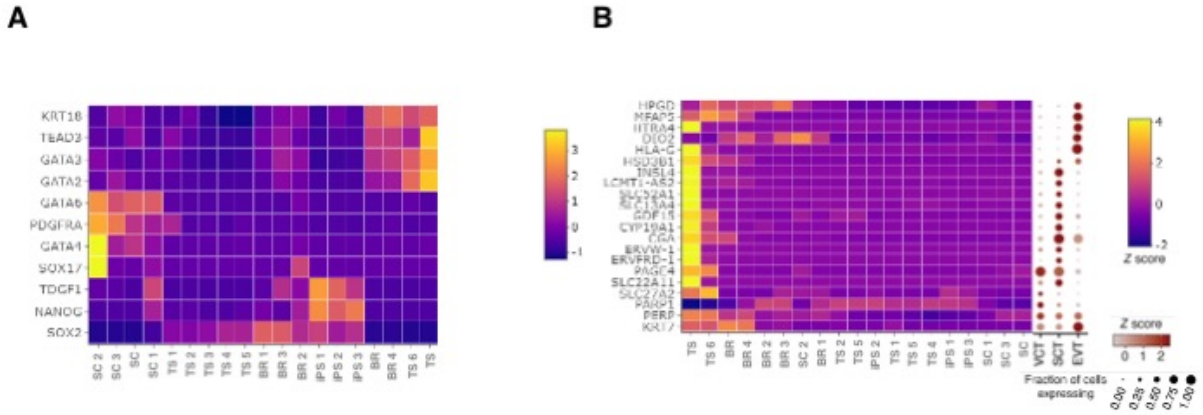


Fig. S7: Lineage specification and cellular cluster identity based on marker gene profiles.

(A) Heatmap of detected high confidence markers for epiblast, primitive endoderm, and trophectoderm from Stirparo et al. 2018.

(B) Heatmap of trophoblast specific markers from Suryawanshi et al. 2018 including modified dotplot from their publication for primary Villous Cytotrophoblasts (VCT), Extravillous Trophoblasts (EVT), and Syncytial Trophoblasts (SCT).

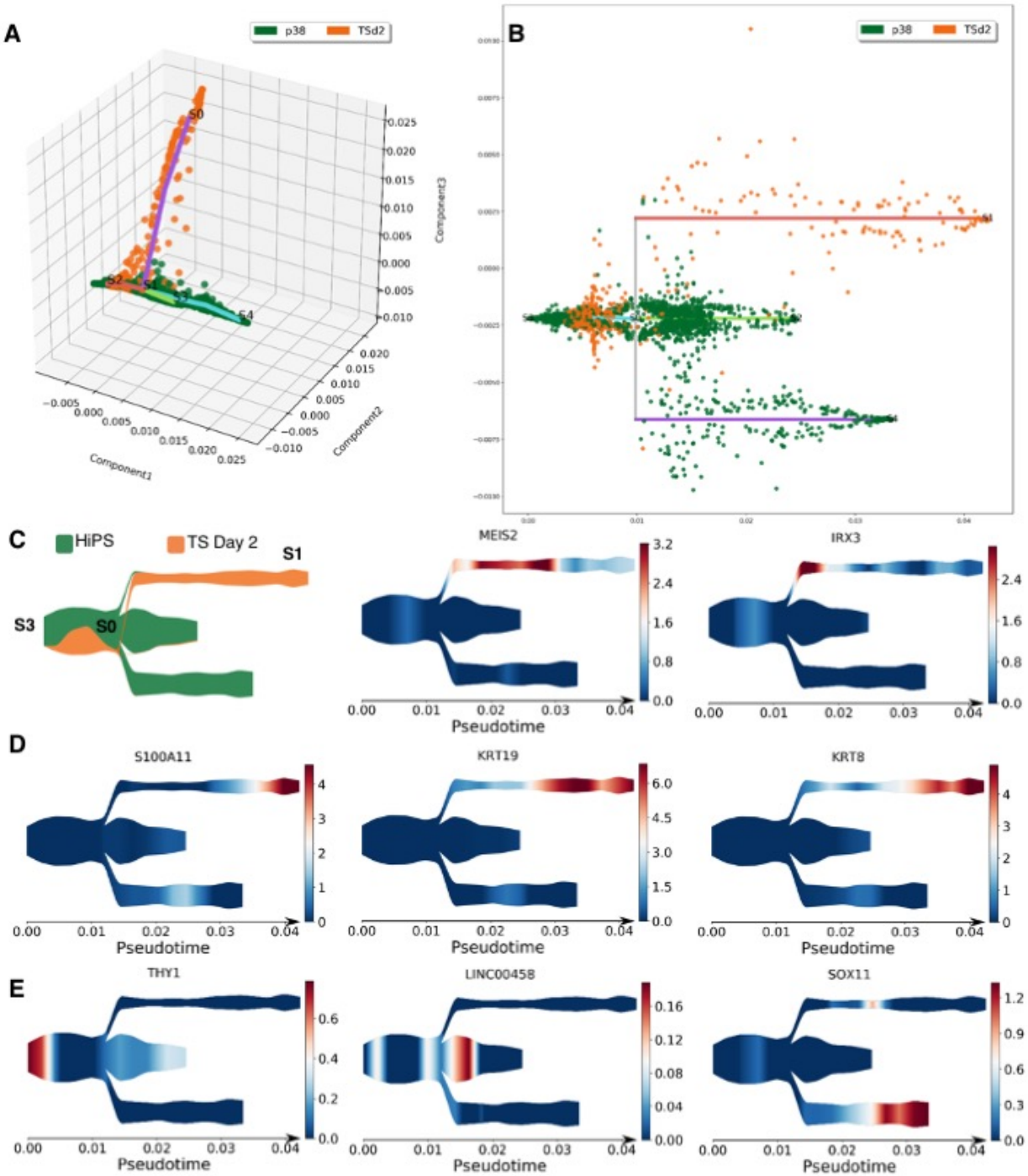


Fig. S8: Developmental trajectory of HiPS and after two days of differentiation in TS media.

(A-B) Visualization of branch assignment using the first 3 SE components (A) and resulting subway map representation of the reconstructed pseudotime trajectory using HiPS and TS D2 cells samples (B).

(C-D) Averaged expression of significant transition genes along the pseudotime trajectory ordered left to right by rank (2-4) for branch S3-S0 (C) and S0-S1 (D).

(E) Averaged expression of selected primed pluripotency markers expressed along psuedotime.

A

Figure S9: PANDA network analysis reveals unique regulation of endogenous Retroviruses in TS cells.

(A) ERV48-1 connected network identified by PANDA in SC cells.

736 **Methods**
737
738 *PSC culture*
739 6-well plates were coated with reduced growth factor Cultrex (1mg/12 ml DMEM/F12) at 37°C
740 for at least one hour. Human ES and iPS were maintained on coated plates in Stemflex media.
741 Media was changed every 48hrs in accordance with manufacturer recommendations. Cells
742 were passaged in small clusters using Versene solution and a split ratio of 1:10-1:12. For
743 cryopreservation, cells were suspended in PSC media and temporarily stored in a deep freezer
744 at −80°C before being transferred to liquid nitrogen for long term storage.
745
746 *Differentiations and Cell Maintenance*
747 6-well plates were coated with reduced growth factor Cultrex (1mg/12 ml DMEM/F12) at 37°C
748 for at least one hour. PSC were harvested in small clusters using Versene solution without
749 centrifugation, seeded at a passage ratio of 1:12, and cultured in 2 mL of Stemflex medium.
750 After 24 hours, cells were washed with DPBS and 2 mL of appropriate media was added to
751 each differentiating well. For TS differentiation, TS medium [DMEM/F12 with Glutamax
752 supplemented with 0.1 mM 2-mercaptoethanol, 0.2% FBS, 0.3% BSA, 1% ITS-X supplement,
753 1.5 µg/ml L-ascorbic acid, 50 ng/ml EGF, 2 µM CHIR99021, 0.5 µM A83-01, 1 µM SB431542,
754 0.8 mM VPA and 5 µM Y27632] was added to each well. For BMP4 based differentiations 10
755 ng/ml BMP4 and/or 2uM IWP2 was added to basal differentiation medium (DMEM/F12 with
756 Glutamax supplemented with ITS and L-ascorbic acid 2-phosphate magnesium) was added to
757 each well. Every 24 hours 2 mL of fresh medium was added to each well. Cells were collected
758 on day 2, 4, and 6 for RNA sequencing. For the TS condition, after 6 days of differentiation cells
759 were passaged using TrypLE express at a split ratio of 1:3 and plated on new Cultrex or col IV
760 (5 ug/mL) coated plates. After the first passage, cells were fed every 48 hours and split at a ratio
761 of 1:3-1:6 every 3-4 days. Mature TS cells between passages 10-20 were used for all
762 experiments unless otherwise noted. Cells were cultured in humidified incubators 5% CO_2 and
763 20% O_2. For hypoxic cultures, O_2 levels were reduced to 5%. For cryopreservation, TS cells
764 were suspended in Cell Banker 1 and temporarily stored in a deep freezer at −80°C before
765 being transferred to liquid nitrogen for long term storage.
766
767 *Immunofluorescence Staining*
768 Cells were washed in Tris Buffered Saline (TBS), fixed with 4% Paraformaldehyde for ten
769 minutes, and washed three times in TBS. Cells were then incubated in blocking solution (3%
770 horse serum, 0.1% Triton-X in TBS) for one hour at room temperature before being incubated
771 with primary antibodies in blocking solution overnight at 4°C. Cells were then rinsed twice and
772 incubated with secondary antibodies in blocking solution for two hours at room temperature.
773 After two washes in TBS cells were counterstained with Hoechst 33342 and rinsed twice in
774 TBS. Cells were imaged using a Zeiss LSM770 or LSM780 confocal microscope. The
775 brightness and/or contrast was globally adjusted for some images.
776
777 *Single Cell sequencing by Dropseq*
778 *Dissociation of Cells for Dropseq*
779 Human iPS cells were grown to 80-90% confluency. After washing cells with DPBS (Thermo
780 Fisher Scientific 14190250) cells were incubated with Versene (Thermo Fisher Scientific

781 15040066) solution for approximately 10 minutes at 37 °C and gently washed up and down with
782 a P1000 pipette to generate a single cell suspension. Cells were then diluted in Stemflex media
783 and centrifuged at 200g for 4 min before being resuspended in 1 mL of 1X PBS, pH 7.4 (Gibco,
784 # 10010023) with 0.01% BSA (Invitrogen, # AM2616) at the concentration of 100 cells/μL.
785
786 *Droplet generation*
787 Barcoded beads (Chemgenes, # Macosko-2011-10) were washed with 30 mL of ethanol (Sigma
788 Aldrich, # E7023-500ML), then removed all ethanol after centrifugation at 1,000g for 1 minute.
789 Washed beads were suspended in 30 mL of TE-TW (10 mM Tris pH8.0, 1 mM EDTA, 0.5%
790 SDS) and spun down at 1,000g for 1 minute. TE-TW was discarded and re-suspended in 20 mL
791 of TE-TW. 100 μm strainer (Corning, Cat# 352360) was used for size selection of beads.
792 Beads were suspended as the concentration of 120 beads/μL in 1 mL of the lysis buffer. To
793 generate droplets, cells, beads, and oil (BioRad, #186-4006) were loaded into the 125-μm
794 microfluidic device after aquapel flush treatment (FlowJEM). The cells and barcoded beads
795 suspension were flown at 2 mL/h each and at 13 mL/h for oil.
796
797 *Droplet breakage, reverse transcription, and exonuclease I treatment*
798 Droplets were collected into the 50 mL tubes and 30 mL of 6X SSC was added, then broken
799 with 1 mL of perfluorooctanol (Sigma, # 370533). Collected beads were washed with 1 mL of 6X
800 SSC twice and with 300 μLof RT buffer. On removing the RT buffer, reverse transcription
801 mixture was added into the beads. To synthesize cDNA, beads in RT mix were incubated with
802 rotation at room temperature for 30 minutes and 42 °C for 90 minutes. The reverse transcription
803 mixture was prepared as in the reference, Macosko, E.Z. et. al., Cell 161, 1202-1214 (2015).
804 Beads were washed once with 1 mL of TE-SDS, twice with 1 mL of TE-TW, and with 1 mL of 10
805 mM Tris pH 8.0. Exonuclease I treatment was performed to remove excess bead primers at 37
806 °C for 45 minutes with rotation followed by washing beads once with 1 mL of TE-SDS, twice
807 with 1 mL of TE-TW, and with 1 mL of molecular grade water.
808
809 *PCR amplification, clean-up with SPRI beads, and library preparation*
810 Unique barcode STAMPs (Single-cell Transcriptomes Attached to MicroParticles) were
811 resuspended in 1 mL H2O, and loaded on a hemocytometer (Incyto, # DHC-N01). Aliquots of
812 3,000 beads in 24.6 μL of H2O were amplified with 0.8 μM SMART PCR primer (IDT, 5'-AAG
813 CAG TGG TAT CAA CGC AGA GT-3') and 2X KAPA HiFi Hotstart Ready Mix (Fisher Scientific,
814 # KK2602) using the following PCR condition: (step 1) 95 °C for 3 min, (step 2) four cycles of:
815 98 °C for 20 sec, 65 °C for 45 sec, 72 °C for 3 min, (step 3) 10 cycles of: 98 °C for 20 sec, 67 °C
816 for 20 s, 72 °C for 3 min, and (step 4) finally, 72 °C for 5 min.
817 Purified cDNA with 0.6X SPRI beads were analyzed with High Sensitive DNA kit (Agilent, #
818 5067-4626), then tagmented and amplified using Nextera XT Library Preparation Kit (Illumina, #
819 FC-131-1096) by manufacturer's manual for sequencing.
820
821 *Sequencing*
822 All libraries were checked its quality with Agilent High Sensitive DNA chips before sequencing.
823 Each library was diluted at 3 nM for an Illumina HiSeq 3000. We used HiSeq SBS 50 cycle kits
824 for sequence (paired-end): 21-bp for read 1 with Custom Read1 primer (IDT, 5'-

825 GCCTGTCCGCGGAAGCAGTGGTATCAACGCAGAGTAC-3'), 21-bp for read 2, and 8-bp for
826 index 1.
827
828 *Preprocessing of Drop-seq data*
829 Raw sequencing data was preprocessed using the pipeline "Drop-seq Alignment Cookbook"
830 v2.0.0 found at https://github.com/broadinstitute/Drop-seq/releases/ and described in (Macosko
831 et al., 2015). Briefly, paired-end reads were filtered in order to remove read pairs with any base
832 with quality of less than 10 in both cellular and molecular barcodes. SMART adapters at 5' end
833 and polyA tails at 3' end with 6 or more bp were removed from the second pair, and then
834 aligned to the reference human (GRCh38) genome using HISAT2 v2.1.0 (Kim et al., 2015) with
835 the default settings. Uniquely mapped reads were retained and grouped by cell barcode. UMI
836 barcodes with a hamming distance = 1 were merged within each cell, and the number of
837 transcripts of each gene was defined based on the total number of each distinct UMI sequence,
838 for that gene, in a given cell.
839
840 *Clustering and Differential Gene Expression analysis for Marker Genes*
841 All downstream analysis for quality control, clustering, and differential expression were
842 performed using Seurat (Stuart et al., 2019) v3.0 R package. Data was initially filtered to remove
843 low quality cells and very low detected genes. We discarded cells with less than 1000 genes
844 detected and with more than 20% of mitochondrial gene content. Genes expressed in less than
845 3 cells were also removed. Variance stabilizing and normalization was performed using
846 sctransform (Hafemeister and Satija, 2019) approach adjusting for library size and the
847 percentage of mitochondrial genes in each cell. To cluster the cells, cell distances were defined
848 with 30 first PCs using a graph-based approach implemented in Seurat, that briefly build a K-
849 nearest neighbor (KNN) graph based on euclidean distance in PCA space, with subsequent
850 refining of edge weights using Jacard similarity. Clusters were defined using Seurat default
851 algorithm – a shared nearest neighbor (SNN) graph based on k-nearest neighbors with
852 modularity optimization using original Louvain algorithm. Gene markers of cell clusters were
853 defined using ROC analysis for only genes with at least 0.25 logFC average difference between
854 the clusters. Graphic visualizations of clusters and gene expression were done using Seurat
855 and ggplot2 in R.
856
857 *Single-cell trajectory reconstruction*
858 Reconstruction of cell differentiation trajectories for cell clusters was performed with STREAM
859 (Single cell Trajectories Reconstruction, Exploration and Mapping) (Chen et al., 2019) Python
860 package. The expression data of high quality cells (9,821 cells) was library size normalized,
861 scaled, natural-log transformed and used as input for STREAM. Data dimension reduction was
862 done with Spectral Embedding (SE) based on most variable genes found with LOESS fitting.
863 Branches representing trajectory structure were learned based on 3 SE dimensions.
864 Marker genes of each branch were defined by Leaf gene detection analysis implemented in the
865 STREAM package. Initially, the average expression of each gene is calculated for all branches.
866 Based on the average expressions, Z-scores are computed and genes are picked as candidates
867 when greater than 1.5. Then a Kruskal–Wallis H-test is calculated for each candidate to test the
868 difference of gene expression among branches. Finally, the Conover test for multiple

comparison is computed for the differentially expressed candidates (Krustal-Wallis p-value < 0.01) between all branches, and genes with Conover p-value < 0.01 are picked as a Leaf gene for that branch. GO enrichment analysis of branch markers were performed for top 100 leaf genes using hypergeometric test with GOstats v2.48.0 R package. Circus plot with GO terms was done using GOplot v1.0.2 R package.
Transition genes were found by performing transition gene analysis implemented in STREAM. The algorithm starts ordering cells based on their inferred pseudotime and splits them in two groups. One group contains 20% of cells closer to the pseudotime start point, and the other is composed of all remaining cells (80%). Gene average expressions are calculated for each group and genes with a log2 fold change greater than 0.25 between groups are picked as candidate genes. Then, Spearman's rank correlation is computed between inferred pseudotime and gene expression of the candidates. Genes with correlation coefficient above 0.4 are reported as transition genes.
Overrepresentation analysis was performed by using the algorithm MSET (Eisinger et al., 2013) implemented in R, using 10000 permutations. The list of transcription factors related to trophoblast differentiation used is described in (Krendl et al., 2017).

Clusters enrichment analysis for placental cell-specific genes.
To characterize the expression profiles of the 19 clusters detected in this work, we analyzed the overlap with cell-specific genes identified by previous single-cell studies. We used two sets of candidate genes. A set of 38 cell-specific gene lists (30 gene each) was drawn from the supplementary table 2 of (Vento-Tormo et al. 2018). These gene lists span both maternal (decidua and blood) and fetal tissues (placenta). Another set of 14 cell-specific gene lists was drawn from the supplementary table 1 of (Liu et al. 2018), which also span analogous tissues.

To profile the expression patterns of the 19 clusters detected in this work, we applied the Cell-Specific Expression Analysis (CSEA) as implemented in the pSI R package (Xu et al., 2014). This approach compares each profile with all the others and identifies genes expressed in one group but not in the others, calculates a score (specificity index, SI) for each gene and attributes a statistical level of significance (pSI). Here, we used the normalized expression matrix to compute cluster-level median gene expression for each gene, which defined 19 cluster profiles. Then, cluster profiles were used to compute SI and pSI for each gene, within each profile. Cluster-specific gene lists were obtained applying the statistical threshold, pSI < 0.05.

Finally, we tested whether cell-specific genes from previous studies are over-represented in our clusters. We used the hypergeometric test and applied the Bonferroni correction for multiple comparisons, considering all the tested gene lists [$\alpha = 0.05/ (19\times (38+14)) = 5.1\times10^{-5}$].

PANDA regulatory networks
Gene regulatory networks were constructed using the pandaR v1.14.0 R package which integrates multiple types of data to infer direct interactions (edges) between TFs and target genes. PANDA initiates with a prior regulatory network that can be built by mapping TF binding sites to the genome, and refines this initial network integrating gene expression data of target genes and TF PPI data. The main idea behind the algorithm is that target genes from a given

TF are likely co-expressed, and TFs that interact with each other are more likely to regulate a similar group of target genes. These two assumptions are used to infer the edge weight of each TF-target interaction, and basically reflects the congruence between the regulatory profile of a TF with target gene co-expression. Iteratively the algorithm refine the initial network structure and infer a final consensus regulatory network.

A regulatory network was built for each cluster using as initial network a TF-motif binding map described in (Sonawane et al., 2017) and downloaded in (https://sites.google.com/a/channing.harvard.edu/kimberlyglass/tools/resources). Gene expression data was library size normalized, scaled, and natural-log transformed. Genes expressed in less than 10% of the cells of each cluster were filtered out. PPI network of TFs was built using STRING database v11.0 (downloaded from https://string-db.org/) and score interactions were divided by 1000 to initiate the PANDA.

Edges with negative weights estimated by PANDA were discarded in order to explore only interactions with greater evidence. The relative importance of individual genes in the network topology was addressed with the betweenness centrality measure calculated using the Igraph v1.2.4 R package. For each cluster, the significance of betweenness values were calculated based on the betweenness probability distribution of 100 random networks constructed with the same number of edges of that cluster. Graphic visualization, gene clusterization, and GO enrichment analysis of gene clusters within networks were performed using Cytoscape v3.7.2 (Shannon et al., 2003) and ClueGO v2.5.4.ENRICHMENT/RANKING ANALYSIS (Bindea et al., 2009).

Software and Analysis:

Analysis pipelines are publicly deposited on Github:
https://github.com/apuapaquola?tab=repositories

Datasets:

Single cell RNA sequencing data is deposited at:

943 **Key Reagents and Resources**

944

Reagent or Resource Type	Name	Supplier	Product Number	Additional information
Antibody	anti-hCGb (mouse monoclonal)	Abcam	ab9582	1:100
Antibody	anti-HLAG (mouse monoclonal)	Abcam	ab52455	1:100
Antibody	anti-KI67 (rabbit polyclonal)	Abcam	ab15580	1:250
Antibody	anti-KRT7 (mouse monoclonal)	ThermoFisher Scientific	MA5-11986	1:1000
Antibody	anti-TJP1 (mouse monoclonal)	ThermoFisher Scientific	33-9100	1:100
Antibody	anti-TP63 (goat polyclonal)	Novus Biologicals	AF1916	1:100
Antibody	anti-VIM (rabbit monoclonal)	Abcam	ab92547	1:500
Antibody	Donkey anti-mouse IgG AF-488	Jackson Immunoresearch Laboratories	715-545-151	1:500
Antibody	Donkey anti-mouse IgG Cy3	Jackson Immunoresearch Laboratories	715-165-151	1:500
Antibody	Donkey anti-rabbit IgG AF-488	Jackson Immunoresearch Laboratories	711-545-152	1:500
Antibody	Donkey anti-rabbit IgG Cy3	Jackson Immunoresearch Laboratories	711-005-152	1:500
Antibody	Donkey anti-rabbit IgG AF-647	Jackson Immunoresearch Laboratories	711-605-152	1:500
Antibody	Donkey anti-goat IgG Cy3	Jackson Immunoresearch Laboratories	705-165-147	1:500
Chemical, Peptide, or Recombinant Protein	2-mercaptoethanol	ThermoFisher Scientific	21985023	
Chemical, Peptide, or Recombinant Protein	Bmp4	Peprotech	120-05-5ug	

Chemical, Peptide, or Recombinant Protein	Bovine Serum Albumin	Cell Signaling Technology	9998S	
Chemical, Peptide, or Recombinant Protein	Cell Banker 1 Cryopreservation Media	Amsbio	11888	
Chemical, Peptide, or Recombinant Protein	Collagen IV	Corning	354233	
Chemical, Peptide, or Recombinant Protein	Cultrex Reduced Growth Factor Basement Membrane Extract, PathClear	R&D Systems	3433-005-01	
Chemical, Peptide, or Recombinant Protein	DMEM/F12, Glutamax supplement	ThermoFisher Scientific	10565018	
Chemical, Peptide, or Recombinant Protein	DPBS, no calcium, no magnesium	ThermoFisher Scientific	14190250	
Chemical, Peptide, or Recombinant Protein	Epidermal Growth Factor	Peprotech	AF-100-15	
Chemical, Peptide, or Recombinant Protein	Fetal Bovine Serum	ThermoFisher Scientific	16140071	
Chemical, Peptide, or Recombinant Protein	Gelatin	Sigma-Aldrich	G1393-100mL	
Chemical, Peptide, or Recombinant Protein	Hoechst 33342	ThermoFisher Scientific	H3570	
Chemical, Peptide, or Recombinant Protein	Insulin-transferrin-selenium	Corning	25-800-CR	
Chemical, Peptide, or Recombinant Protein	ITS-X	ThermoFisher Scientific	51500056	
Chemical, Peptide, or Recombinant Protein	IWP2	BioGems	6866167-1mg	
Chemical, Peptide, or Recombinant Protein	L-Ascorbic Acid	BioGems	2520691	

Chemical, Peptide, or Recombinant Protein	L-ascorbic acid 2-phosphate magnesium	Sigma-Aldrich	A8960-5G	
Chemical, Peptide, or Recombinant Protein	SB431542	BioGems	3014193	
Chemical, Peptide, or Recombinant Protein	StemFlex Media	ThermoFisher Scientific	A3349401	
Chemical, Peptide, or Recombinant Protein	TrypLE express	ThermoFisher Scientific	12604013	
Chemical, Peptide, or Recombinant Protein	Valproic Acid Sodium salt	BioGems	1066656	
Chemical, Peptide, or Recombinant Protein	Y-27632	BioGems	1293823	
Cell Line	2014.06 Human Induced Pluripotent Stem Cells	Lieber Institute	N/A	
Cell Line	CF-1 Mouse Embryonic Fibroblasts	ThermoFisher Scientific	A34181	
Cell Line	LIBD1C8 Human Induced Pluripotent Stem Cells	Sawada et al., 2020	N/A	Generated in lab
Cell Line	LIBD6C2 Human Induced Pluripotent Stem Cells	Sawada et al., 2020	N/A	Generated in lab
Cell Line	LIBD7C6 Human Induced Pluripotent Stem Cells	Sawada et al., 2020	N/A	Generated in lab
Cell Line	WA-01 (H1) Human Embryonic Stem Cells	Wicell	WB0197	

948

949 Bibliography

950 Abd El-Aleem, S.A., and Dekker, L.V. (2018). Assessment of the cellular localization of the
951 annexin A2/S100A10 complex in human placenta. J. Mol. Histol. *49*, 531–543.

952 Aguiar, J.A., Tremblay, B.J.-M., Mansfield, M.J., Woody, O., Lobb, B., Banerjee, A.,
953 Chandiramohan, A., Tiessen, N., Cao, Q., Dvorkin-Gheva, A., et al. (2020). Gene expression
954 and in situ protein profiling of candidate SARS-CoV-2 receptors in human airway epithelial cells
955 and lung tissue. Eur. Respir. J.

956 Al-Haddad, B.J.S., Oler, E., Armistead, B., Elsayed, N.A., Weinberger, D.R., Bernier, R., Burd,
957 I., Kapur, R., Jacobsson, B., Wang, C., et al. (2019). The fetal origins of mental illness. Am. J.
958 Obstet. Gynecol. *221*, 549–562.

959 Arai, C., Yoshizaki, K., Miyazaki, K., Saito, K., Yamada, A., Han, X., Funada, K., Fukumoto, E.,
960 Haruyama, N., Iwamoto, T., et al. (2017). Nephronectin plays critical roles in Sox2 expression
961 and proliferation in dental epithelial stem cells via EGF-like repeat domains. Sci. Rep. *7*, 45181.

962 Bernardo, A.S., Faial, T., Gardner, L., Niakan, K.K., Ortmann, D., Senner, C.E., Callery, E.M.,
963 Trotter, M.W., Hemberger, M., Smith, J.C., et al. (2011). BRACHYURY and CDX2 mediate
964 BMP-induced differentiation of human and mouse pluripotent stem cells into embryonic and
965 extraembryonic lineages. Cell Stem Cell *9*, 144–155.

966 Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.-H.,
967 Pagès, F., Trajanoski, Z., and Galon, J. (2009). ClueGO: a Cytoscape plug-in to decipher
968 functionally grouped gene ontology and pathway annotation networks. Bioinformatics *25*, 1091–
969 1093.

970 Blaise, S., de Parseval, N., Bénit, L., and Heidmann, T. (2003). Genomewide screening for
971 fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on
972 primate evolution. Proc Natl Acad Sci USA *100*, 13013–13018.

973 Blakeley, P., Fogarty, N.M.E., del Valle, I., Wamaitha, S.E., Hu, T.X., Elder, K., Snell, P.,
974 Christie, L., Robson, P., and Niakan, K.K. (2015). Defining the three cell lineages of the human
975 blastocyst by single-cell RNA-seq. Development *142*, 3151–3165.

976 Blond, J.L., Besème, F., Duret, L., Bouton, O., Bedin, F., Perron, H., Mandrand, B., and Mallet,
977 F. (1999). Molecular characterization and placental expression of HERV-W, a new human
978 endogenous retrovirus family. J. Virol. *73*, 1175–1185.

979 Blond, J.L., Lavillette, D., Cheynet, V., Bouton, O., Oriol, G., Chapel-Fernandes, S., Mandrand,
980 B., Mallet, F., and Cosset, F.L. (2000). An envelope glycoprotein of the human endogenous
981 retrovirus HERV-W is expressed in the human placenta and fuses cells expressing the type D
982 mammalian retrovirus receptor. J. Virol. *74*, 3321–3329.

983 Chen, B., Dodge, M.E., Tang, W., Lu, J., Ma, Z., Fan, C.-W., Wei, S., Hao, W., Kilgore, J.,
984 Williams, N.S., et al. (2009). Small molecule-mediated disruption of Wnt-dependent signaling in
985 tissue regeneration and cancer. Nat. Chem. Biol. *5*, 100–107.

986 Chen, H., Albergante, L., Hsu, J.Y., Lareau, C.A., Lo Bosco, G., Guan, J., Zhou, S., Gorban,

987 A.N., Bauer, D.E., Aryee, M.J., et al. (2019). Single-cell trajectories reconstruction, exploration
988 and mapping of omics data with STREAM. Nat. Commun. *10*, 1903.

989 Chen, I.-H., Wang, H.-H., Hsieh, Y.-S., Huang, W.-C., Yeh, H.-I., and Chuang, Y.-J. (2013).
990 PRSS23 is essential for the Snail-dependent endothelial-to-mesenchymal transition during
991 valvulogenesis in zebrafish. Cardiovasc. Res. *97*, 443–453.

992 Chua, C.E.L., and Tang, B.L. (2014). Engagement of the small GTPase Rab31 protein and its
993 effector, early endosome antigen 1, is important for trafficking of the ligand-bound epidermal
994 growth factor receptor from the early to the late endosome. J. Biol. Chem. *289*, 12375–12389.

995 Cole, L.A. (2010). Biological functions of hCG and hCG-related molecules. Reprod. Biol.
996 Endocrinol. *8*, 102.

997 Cornelis, G., Heidmann, O., Degrelle, S.A., Vernochet, C., Lavialle, C., Letzelter, C., Bernard-
998 Stoecklin, S., Hassanin, A., Mulot, B., Guillomot, M., et al. (2013). Captured retroviral envelope
999 syncytin gene associated with the unique placental structure of higher ruminants. Proc Natl
1000 Acad Sci USA *110*, E828-37.

1001 Dong, C., Beltcheva, M., Gontarz, P., Zhang, B., Popli, P., Fischer, L.A., Khan, S.A., Park, K.-
1002 M., Yoon, E.-J., Xing, X., et al. (2020). Derivation of trophoblast stem cells from naïve human
1003 pluripotent stem cells. Elife *9*.

1004 Drukker, M., Tang, C., Ardehali, R., Rinkevich, Y., Seita, J., Lee, A.S., Mosley, A.R., Weissman,
1005 I.L., and Soen, Y. (2012). Isolation of primitive endoderm, mesoderm, vascular endothelial and
1006 trophoblast progenitors from human pluripotent stem cells. Nat. Biotechnol. *30*, 531–542.

1007 Dunn-Fletcher, C.E., Muglia, L.M., Pavlicev, M., Wolf, G., Sun, M.-A., Hu, Y.-C., Huffman, E.,
1008 Tumukuntala, S., Thiele, K., Mukherjee, A., et al. (2018). Anthropoid primate-specific retroviral
1009 element THE1B controls expression of CRH in placenta and alters gestation length. PLoS Biol.
1010 *16*, e2006337.

1011 Eisinger, B.E., Saul, M.C., Driessen, T.M., and Gammie, S.C. (2013). Development of a
1012 versatile enrichment analysis tool reveals associations between the maternal brain and mental
1013 health disorders, including autism. BMC Neurosci. *14*, 147.

1014 Erb, T.M., Schneider, C., Mucko, S.E., Sanfilippo, J.S., Lowry, N.C., Desai, M.N., Mangoubi,
1015 R.S., Leuba, S.H., and Sammak, P.J. (2011). Paracrine and epigenetic control of trophectoderm
1016 differentiation from human embryonic stem cells: the role of bone morphogenic protein 4 and
1017 histone deacetylases. Stem Cells Dev. *20*, 1601–1614.

1018 Erwin, J.A., del Rosario, B., Payer, B., and Lee, J.T. (2012). An ex vivo model for imprinting:
1019 mutually exclusive binding of Cdx2 and Oct4 as a switch for imprinted and random X-
1020 inactivation. Genetics *192*, 857–868.

1021 Fukamachi, K., Matsuoka, Y., Ohno, H., Hamaguchi, T., and Tsuda, H. (2002). Neuronal
1022 leucine-rich repeat protein-3 amplifies MAPK activation by epidermal growth factor through a
1023 carboxyl-terminal region containing endocytosis motifs. J. Biol. Chem. *277*, 43549–43552.

1024 Gao, X., Nowak-Imialek, M., Chen, X., Chen, D., Herrmann, D., Ruan, D., Chen, A.C.H.,
1025 Eckersley-Maslin, M.A., Ahmad, S., Lee, Y.L., et al. (2019). Establishment of porcine and
1026 human expanded potential stem cells. Nat. Cell Biol. *21*, 687–699.

1027 Glass, K., Huttenhower, C., Quackenbush, J., and Yuan, G.-C. (2013). Passing messages
1028 between biological networks to refine predicted interactions. PLoS ONE *8*, e64832.

1029 Goldman, D.C., Bailey, A.S., Pfaffle, D.L., Al Masri, A., Christian, J.L., and Fleming, W.H.
1030 (2009). BMP4 regulates the hematopoietic stem cell niche. Blood *114*, 4393–4401.

1031 Guan, L., Li, T., Ai, N., Wang, W., He, B., Bai, Y., Yu, Z., Li, M., Dong, S., Zhu, Q., et al. (2019).
1032 MEIS2C and MEIS2D promote tumor progression via Wnt/β-catenin and hippo/YAP signaling in
1033 hepatocellular carcinoma. J. Exp. Clin. Cancer Res. *38*, 417.

1034 Guleria, I., and Pollard, J.W. (2000). The trophoblast is a component of the innate immune
1035 system during pregnancy. Nat. Med. *6*, 589–593.

1036 Hafemeister, C., and Satija, R. (2019). Normalization and variance stabilization of single-cell
1037 RNA-seq data using regularized negative binomial regression. Genome Biol. *20*, 296.

1038 Hendee, K.E., Sorokina, E.A., Muheisen, S.S., Reis, L.M., Tyler, R.C., Markovic, V., Cuturilo,
1039 G., Link, B.A., and Semina, E.V. (2018). PITX2 deficiency and associated human disease:
1040 insights from the zebrafish model. Hum. Mol. Genet. *27*, 1675–1695.

1041 Horii, M., Bui, T., Touma, O., Cho, H.Y., and Parast, M.M. (2019). An Improved Two-Step
1042 Protocol for Trophoblast Differentiation of Human Pluripotent Stem Cells. Curr. Protoc. Stem
1043 Cell Biol. *50*, e96.

1044 Huang, D., Wang, S., Wang, A., Chen, X., and Zhang, H. (2016). Thymosin beta 4 silencing
1045 suppresses proliferation and invasion of non-small cell lung cancer cells by repressing Notch1
1046 activation. Acta Biochim Biophys Sin (Shanghai) *48*, 788–794.

1047 Johnson, W.E. (2019). Origins and evolutionary consequences of ancient endogenous
1048 retroviruses. Nat. Rev. Microbiol. *17*, 355–370.

1049 Kazanskaya, O., Glinka, A., del Barco Barrantes, I., Stannek, P., Niehrs, C., and Wu, W. (2004).
1050 R-Spondin2 is a secreted activator of Wnt/beta-catenin signaling and is required for Xenopus
1051 myogenesis. Dev. Cell *7*, 525–534.

1052 Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low
1053 memory requirements. Nat. Methods *12*, 357–360.

1054 Krendl, C., Shaposhnikov, D., Rishko, V., Ori, C., Ziegenhain, C., Sass, S., Simon, L., Müller,
1055 N.S., Straub, T., Brooks, K.E., et al. (2017). GATA2/3-TFAP2A/C transcription factor network
1056 couples human pluripotent stem cell differentiation to trophectoderm with repression of
1057 pluripotency. Proc Natl Acad Sci USA *114*, E9579–E9588.

1058 Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar,
1059 K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome.
1060 Nature *409*, 860–921.

1061 Leach, J.L., Sedmak, D.D., Osborne, J.M., Rahill, B., Lairmore, M.D., and Anderson, C.L.
1062 (1996). Isolation from human placenta of the IgG transporter, FcRn, and localization to the
1063 syncytiotrophoblast: implications for maternal-fetal antibody transport. J. Immunol. *157*, 3317–
1064 3322.

1065 Lee, C.Q.E., Turco, M.Y., Gardner, L., Simons, B.D., Hemberger, M., and Moffett, A. (2018).

1066 Integrin α2 marks a niche of trophoblast progenitor cells in first trimester human placenta.
1067 Development *145*.

1068 Liu, A.-X., Jin, F., Zhang, W.-W., Zhou, T.-H., Zhou, C.-Y., Yao, W.-M., Qian, Y.-L., and Huang,
1069 H.-F. (2006). Proteomic analysis on the alteration of protein expression in the placental villous
1070 tissue of early pregnancy loss. Biol. Reprod. *75*, 414–420.

1071 Liu, Y., Fan, X., Wang, R., Lu, X., Dang, Y.-L., Wang, H., Lin, H.-Y., Zhu, C., Ge, H., Cross,
1072 J.C., et al. (2018). Single-cell RNA-seq reveals the diversity of trophoblast subtypes and
1073 patterns of differentiation in the human placenta. Cell Res. *28*, 819–832.

1074 Li, Z., Kurosawa, O., and Iwata, H. (2019). Establishment of human trophoblast stem cells from
1075 human induced pluripotent stem cell-derived cystic cells under micromesh culture. Stem Cell
1076 Res. Ther. *10*, 245.

1077 Lv, S., Cheng, G., Zhou, Y., and Xu, G. (2013). Thymosin beta4 induces angiogenesis through
1078 Notch signaling in endothelial cells. Mol. Cell. Biochem. *381*, 283–290.

1079 Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas,
1080 A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly Parallel Genome-wide Expression
1081 Profiling of Individual Cells Using Nanoliter Droplets. Cell *161*, 1202–1214.

1082 Mansour, A.A., Schafer, S.T., and Gage, F.H. (2020). Cellular complexity in brain organoids:
1083 Current progress and unsolved issues. Semin. Cell Dev. Biol.

1084 Messmer, T., von Meyenn, F., Savino, A., Santos, F., Mohammed, H., Lun, A.T.L., Marioni, J.C.,
1085 and Reik, W. (2019). Transcriptional Heterogeneity in Naive and Primed Human Pluripotent
1086 Stem Cells at Single-Cell Resolution. Cell Rep. *26*, 815-824.e4.

1087 Meyer, K., Kwon, Y.-C., Liu, S., Hagedorn, C.H., Ray, R.B., and Ray, R. (2015). Interferon-α
1088 inducible protein 6 impairs EGFR activation by CD81 and inhibits hepatitis C virus infection. Sci.
1089 Rep. *5*, 9012.

1090 Mischler, A., Karakis, V., Mahinthakumar, J., Carberry, C., San Miguel, A., Rager, J., Fry, R.,
1091 and Rao, B.M. (2019). Two distinct trophectoderm lineage stem cells from human pluripotent
1092 stem cells. BioRxiv.

1093 Mouse Genome Sequencing Consortium, Waterston, R.H., Lindblad-Toh, K., Birney, E.,
1094 Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., et al.
1095 (2002). Initial sequencing and comparative analysis of the mouse genome. Nature *420*, 520–
1096 562.

1097 Myatt, L. (2006). Placental adaptive responses and fetal programming. J Physiol (Lond) *572*,
1098 25–30.

1099 Niakan, K.K., and Eggan, K. (2013). Analysis of human embryos from zygote to blastocyst
1100 reveals distinct gene expression patterns relative to the mouse. Dev. Biol. *375*, 54–64.

1101 Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., and Rossant, J.
1102 (2005). Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. Cell
1103 *123*, 917–929.

1104 Oh, Y., Nagalla, S.R., Yamanaka, Y., Kim, H.S., Wilson, E., and Rosenfeld, R.G. (1996).

1105 Synthesis and characterization of insulin-like growth factor-binding protein (IGFBP)-7.
1106 Recombinant human mac25 protein specifically binds IGF-I and -II. J. Biol. Chem. *271*, 30322–
1107 30325.

1108 Okae, H., Toh, H., Sato, T., Hiura, H., Takahashi, S., Shirane, K., Kabayama, Y., Suyama, M.,
1109 Sasaki, H., and Arima, T. (2018). Derivation of human trophoblast stem cells. Cell Stem Cell *22*,
1110 50-63.e6.

1111 Osafune, K., Caron, L., Borowiak, M., Martinez, R.J., Fitz-Gerald, C.S., Sato, Y., Cowan, C.A.,
1112 Chien, K.R., and Melton, D.A. (2008). Marked differences in differentiation propensity among
1113 human embryonic stem cell lines. Nat. Biotechnol. *26*, 313–315.

1114 Overgaard, M.T., Boldt, H.B., Laursen, L.S., Sottrup-Jensen, L., Conover, C.A., and Oxvig, C.
1115 (2001). Pregnancy-associated plasma protein-A2 (PAPP-A2), a novel insulin-like growth factor-
1116 binding protein-5 proteinase. J. Biol. Chem. *276*, 21849–21853.

1117 Parry, S., Holder, J., and Strauss, J.F. (1997). Mechanisms of trophoblast-virus interaction. J.
1118 Reprod. Immunol. *37*, 25–34.

1119 Petropoulos, S., Edsgärd, D., Reinius, B., Deng, Q., Panula, S.P., Codeluppi, S., Plaza Reyes,
1120 A., Linnarsson, S., Sandberg, R., and Lanner, F. (2016). Single-Cell RNA-Seq Reveals Lineage
1121 and X Chromosome Dynamics in Human Preimplantation Embryos. Cell *165*, 1012–1026.

1122 Pique-Regi, R., Romero, R., Tarca, A.L., Luca, F., Xu, Y., Alazizi, A., Leng, Y., Hsu, C.-D., and
1123 Gomez-Lopez, N. (2020). Does the human placenta express the canonical cell entry mediators
1124 for SARS-CoV-2? Elife *9*.

1125 Porazinski, S., Wang, H., Asaoka, Y., Behrndt, M., Miyamoto, T., Morita, H., Hata, S., Sasaki,
1126 T., Krens, S.F.G., Osada, Y., et al. (2015). YAP is essential for tissue tension to ensure
1127 vertebrate 3D body shape. Nature *521*, 217–221.

1128 Roberts, R.M., Ezashi, T., Sheridan, M.A., and Yang, Y. (2018). Specification of trophoblast
1129 from embryonic stem cells exposed to BMP4. Biol. Reprod. *99*, 212–224.

1130 Rossant, J. (2018). Genetic control of early cell lineages in the mammalian embryo. Annu. Rev.
1131 Genet. *52*, 185–201.

1132 Sawada, T., Benjamin, K.J.M., Brandtjen, A.C., Tietze, E., Allen, S.J., Paquola, A.C.M.,
1133 Kleinman, J.E., Hyde, T.M., and Erwin, J.A. (2020). Generation of four postmortem dura-derived
1134 iPS cell lines from four control individuals with genotypic and brain-region-specific transcriptomic
1135 data available through the BrainSEQ consortium. Stem Cell Res. *46*, 101806.

1136 Schneider, M.D., Gaussin, V., and Lyons, K.M. (2003). Tempting fate: BMP signals for cardiac
1137 morphogenesis. Cytokine Growth Factor Rev. *14*, 1–4.

1138 Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N.,
1139 Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated
1140 models of biomolecular interaction networks. Genome Res. *13*, 2498–2504.

1141 Singh, M., Bansal, V., and Feschotte, C. (2020). A single-cell RNA expression map of human
1142 coronavirus entry factors. BioRxiv.

1143 Sonawane, A.R., Platig, J., Fagny, M., Chen, C.-Y., Paulson, J.N., Lopes-Ramos, C.M., DeMeo,

1144 D.L., Quackenbush, J., Glass, K., and Kuijjer, M.L. (2017). Understanding Tissue-Specific Gene
1145 Regulation. Cell Rep. *21*, 1077–1088.

1146 Stirparo, G.G., Boroviak, T., Guo, G., Nichols, J., Smith, A., and Bertone, P. (2018). Integrated
1147 analysis of single-cell embryo data yields a unified transcriptome signature for the human pre-
1148 implantation epiblast. Development *145*.

1149 Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y.,
1150 Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell
1151 Data. Cell *177*, 1888-1902.e21.

1152 Suryawanshi, H., Morozov, P., Straus, A., Sahasrabudhe, N., Max, K.E.A., Garzia, A., Kustagi,
1153 M., Tuschl, T., and Williams, Z. (2018). A single-cell survey of the human first-trimester placenta
1154 and decidua. Sci. Adv. *4*, eaau4788.

1155 Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S.,
1156 and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science
1157 *282*, 1145–1147.

1158 Vallier, L., Touboul, T., Chng, Z., Brimpari, M., Hannan, N., Millan, E., Smithers, L.E., Trotter,
1159 M., Rugg-Gunn, P., Weber, A., et al. (2009). Early cell fate decisions of human embryonic stem
1160 cells and mouse epiblast stem cells are controlled by the same signalling pathways. PLoS ONE
1161 *4*, e6082.

1162 Vento-Tormo, R., Efremova, M., Botting, R.A., Turco, M.Y., Vento-Tormo, M., Meyer, K.B., Park,
1163 J.-E., Stephenson, E., Polański, K., Goncalves, A., et al. (2018). Single-cell reconstruction of the
1164 early maternal-fetal interface in humans. Nature *563*, 347–353.

1165 Vivanti, A.J., Vauloup-Fellous, C., Prevot, S., Zupan, V., Suffee, C., Do Cao, J., Benachi, A.,
1166 and De Luca, D. (2020). Transplacental transmission of SARS-CoV-2 infection. Nat. Commun.
1167 *11*, 3572.

1168 Wakeland, A.K., Soncin, F., Moretto-Zita, M., Chang, C.-W., Horii, M., Pizzo, D., Nelson, K.K.,
1169 Laurent, L.C., and Parast, M.M. (2017). Hypoxia Directs Human Extravillous Trophoblast
1170 Differentiation in a Hypoxia-Inducible Factor-Dependent Manner. Am. J. Pathol. *187*, 767–780.

1171 Wang, R.-Y., Chen, L., Chen, H.-Y., Hu, L., Li, L., Sun, H.-Y., Jiang, F., Zhao, J., Liu, G.-M.-Y.,
1172 Tang, J., et al. (2013). MUC15 inhibits dimerization of EGFR and PI3K-AKT signaling and is
1173 associated with aggressive hepatocellular carcinomas in patients. Gastroenterology *145*, 1436-
1174 48.e1.

1175 Weinberger, L., Ayyash, M., Novershtern, N., and Hanna, J.H. (2016). Dynamic stem cell states:
1176 naive to primed pluripotency in rodents and humans. Nat. Rev. Mol. Cell Biol. *17*, 155–169.

1177 Xu, R.-H., Chen, X., Li, D.S., Li, R., Addicks, G.C., Glennon, C., Zwaka, T.P., and Thomson,
1178 J.A. (2002). BMP4 initiates human embryonic stem cell differentiation to trophoblast. Nat.
1179 Biotechnol. *20*, 1261–1264.

1180 Xu, X., Wells, A.B., O'Brien, D.R., Nehorai, A., and Dougherty, J.D. (2014). Cell type-specific
1181 expression analysis to identify putative cellular mechanisms for neurogenetic disorders. J.
1182 Neurosci. *34*, 1420–1431.

1183 Yabe, S., Alexenko, A.P., Amita, M., Yang, Y., Schust, D.J., Sadovsky, Y., Ezashi, T., and
1184 Roberts, R.M. (2016). Comparison of syncytiotrophoblast generated from human embryonic
1185 stem cells and from term placentas. Proc Natl Acad Sci USA *113*, E2598-607.

1186 Yang, L., Soonpaa, M.H., Adler, E.D., Roepke, T.K., Kattman, S.J., Kennedy, M., Henckaerts,
1187 E., Bonham, K., Abbott, G.W., Linden, R.M., et al. (2008). Human cardiovascular progenitor
1188 cells develop from a KDR+ embryonic-stem-cell-derived population. Nature *453*, 524–528.

1189 Zhang, P., Li, J., Tan, Z., Wang, C., Liu, T., Chen, L., Yong, J., Jiang, W., Sun, X., Du, L., et al.
1190 (2008). Short-term BMP-4 treatment initiates mesoderm induction in human embryonic stem
1191 cells. Blood *111*, 1933–1941.