Senescent syncytiotrophoblast secretion during early-onset preeclampsia

Author names

Olivia Nonn[1,2,3,4,5,21], Olivia Debnath[6,21], Daniela S. Valdes[1,2,3,4,21], Katja Sallinger[5,7], Ali Kerim Secener[8,9], Cornelius Fischer[8], Sebastian Tiesmeyer[6], Jose Nimo[3], Thomas Kuenzer[10], Juliane Ulrich[1,2,3], Theresa Maxian[11], Martin Knöfler[11], Philipp Karau[6], Hendrik Bartolomaeus[1,2,3,4], Thomas Kroneis[5,7], Alina Frolova[1,2,3,12], Lena Neuper[5], Nadine Haase[1,2,3,4], Alexander Malt[6], Niklas Müller-Bötticher[6], Kristin Kräker[1,2,3,4], Sarah Kedziora[1,2,3,4], Désirée Forstner[5], Roland Eils[6], Ruth Schmidt-Ullrich[2,3], Sandra Haider[11], Stefan Verlohren[13], Christina Stern[14], Meryam Sugulle[15,16], Stuart Jones[17], Basky Thilaganathan[18], Tu'uhevaha J Kaitu'u-Lino[19], Stephen Tong[19], Berthold Huppertz[5], Amin El-Heliebi[5,7], Anne Cathrine Staff[15,16], Fabian Coscia[3], Dominik N. Müller[1,2,3,4], Ralf Dechend[1,2,3,4,20], Martin Gauster[5], Naveed Ishaque[6,22], Florian Herse[1,2,3,22]

Supplemental Methods

Patient samples

Tissue sampling was done in a multicentre-design. Patients were recruited in Berlin (German), Graz and Vienna (Austria), Oslo (Norway), London (UK) and Melbourne (Australia). The studies were approved by each regional ethics committee and described individually and headlined by the analysing method.

First trimester tissue used for snRNA-sequencing

Placental and matching decidual tissue were collected from electively terminated pregnancies with informed consent of healthy individuals (gestational age 5 – 11 weeks). Exclusion criteria were maternal age under 18, maternal BMI >25 and self-reported maternal pathologies. Ethical approval was obtained from the Medical University of Graz Ethics Committee (31-019 ex 18/19; 26-132 ex 13/14). Immediately after surgical extraction, tissue was stored at 4°C in culture medium DMEM/F12 1:1, 1 g/dL glucose, Gibco®, Life Technologies (TM), Thermo Fisher Scientific, Vienna, Austria) and processed in no more than 4 hours. Villous and decidual tissues were separately rinsed twice in cold (4°C) 0.9% NaCl to remove blood, afterwards snap frozen in liquid nitrogen and stored at -80°C until processing. Patient characteristics can be found in Supplemental Table S1.

Healthy term tissue used for snRNA-sequencing

Healthy term samples were collected immediately after delivery at the inpatient clinic of the Department of Obstetrics and Gynecology, University Hospital Graz, Austria. The study was approved by the local Ethics committee at the Medical University of Graz (31-019 ex 18/19; 26-132 ex 13/14) and informed consent was obtained from each participating woman. Representative tissue samples (1x1x1 cm) of the medial third of the placenta were cut from vital cotyledons that were macroscopically free of infarct areas or other obvious pathologies that are assumed to have happened during delivery. This should avoid sampling degraded RNA and ensure a high-quality yield for further analysis, well knowing that it might skew towards possibly inaccurate phenotypes on either side of disease and healthy samples. Amnion was dissected and remaining tissue was rinsed twice in cold (4°C) 0.9% NaCl to remove blood, afterwards

snap frozen in liquid nitrogen and stored at -80°C until processing. Patient characteristics can be found in Supplemental Table S1.

Early-onset preeclampsia and healthy term tissue used for snRNA-sequencing

Pregnant women were recruited in Oslo University Hospital prior to elective caesarean section after informed written consent, as previously described[39] , from women with either early-onset preeclampsia (eoPE) or normotensive pregnancies. eoPE was defined as new onset hypertension (blood pressure ≥140/90 mmHg) and new onset proteinuria (≥1+ on dipstick, or ≥30 protein/creatinine ratio) at ≥20 weeks gestation and with delivery prior to gestational week 34. Placental villous tissue biopsies were cut from the centre of central normal appearing cotyledons, were snap frozen in liquid nitrogen and stored at −80°C until use. The study was approved by the regional committee for Medical and Health Research Ethics in South-Eastern Norway and performed according to the Helsinki Declaration. Patient characteristics can be found in Supplemental Table S1.

Validation cohorts

Validation cohorts were recruited individually as described below. Tissue was proceeded as described in “RNA isolation and RT-qPCR”. Data was z-transformed and merged for the following cohorts: Graz, Oslo, Berlin. The cohort from Melbourne also included subjects with pure FGR and was handled separately.

<u>Validation cohort Graz</u>

Study samples were recruited retrospectively immediately after delivery at the inpatient clinic of the Department of Obstetrics and Gynecology, University Hospital Graz, Austria between 2018 and 2019. Preeclampsia (PE) was defined according to the ISSHP guidelines. Women receiving low dose aspirin were excluded. The study was approved by the local Ethics committee at the Medical University of Graz (26-132 ex 13/14 and 31-019 ex 18/19) and informed consent was obtained from each participating woman. Patient characteristics can be found in Supplemental Table S1.

<u>Validation cohort Oslo</u>

Pregnant women were recruited prior to elective caesarean section after informed written consent, as previously described, from women with either preeclampsia or

normotensive pregnancies. PE was defined as new onset hypertension (blood pressure ≥140/90 mmHg) and new onset proteinuria (≥1+ on dipstick, or ≥30 protein/creatinine ratio) at ≥20 weeks gestation. In addition, eoPE was defined as delivery prior to gestational week 34. Placental villous tissue biopsies were cut from the centre of central normal appearing cotyledons, and were snap frozen in liquid nitrogen and stored at −80°C until use. The study was approved by the Regional committee for Medical and Health Research Ethics in South-Eastern Norway and performed according to the Helsinki Declaration. Patient characteristics can be found in Supplemental Table S1.

Validation cohort Berlin

Samples from 19 placentas <34 weeks were collected from March 2013 to July 2014 at the Department of Obstetrics at Charité University Medicine, Campus Virchow Clinic, Berlin, Germany. The trial protocol was approved by the local ethics committee and written and informed consent was obtained from all participants. Women were recruited at the time of clinical admission. PE was defined according to the International Society for the Study of Hypertension in Pregnancy (ISSHP) 2000, as new onset hypertension of >140/90 mmHg at two occasions six hours apart, in combination with proteinuria of >300 mg/24 h or >2+ dip stick. Patient characteristics can be found Supplemental Table S1.

Validation cohort Melbourne

Placental tissue was collected from patients at Mercy Hospital for Women (Melbourne, Australia) with early-onset preeclampsia (<34 weeks' gestation) and gestation-matched controls (<34 weeks' gestation,) who delivered via caesarean section. Early-onset preeclampsia was diagnosed in accordance with American College of Obstetricians and Gynecologists (ACOG) guidelines. Preterm control placentas were obtained from normotensive patients with no evidence of placental insufficiency who delivered preterm due to complications such as placenta previa. Tissue was sampled from four separate sites of the placenta, washed in ice-cold Dulbecco's phosphate-buffered saline (dPBS) and preserved with RNAlater Stabilization Solution (InvitrogenTM, Waltham, USA). Samples were stored at –80°C until RNA extraction. Ethics approval was granted by Mercy Health Human Research Ethics Committee (R11/34) and participants presenting to the Mercy Hospital for Women gave informed, written consent for sample collection. Patient characteristics can be found in Supplemental Table S1.

Single-nucleus sequencing (snRNA-Seq)

Nuclei capture, library generation, sequencing

Approximately 100-200 mg frozen placental and corresponding, separately sampled, decidual tissue was processed according to an optimised nuclei isolation protocol by Krishnaswami et al.[40] Briefly, frozen tissue was disrupted with a pre-cooled glass Dounce in homogenisation buffer (1X NIM2 [1X protease inhibitor, 1 µM DDT, 250 mM sucrose, 25 mM KCl, 5 mM MgCl2, 10 mM pH8.0 Tris], 0.4 U/µL RNAseIn, 0.2 U/µL Superasin, 0.1% v/v Triton X-100) and filtered through a flow-cytometry (BD Falcon) tube with a 35 µm cell sieve cap. Homogenate was incubated in the dark, on ice, for two minutes with DAPI (5 µg/µL) and centrifuged for eight minutes (1,000xg, 4°C). Pellet was resuspended with staining buffer, transferred to a FACS-tube (BD Falcon) with a 35 µm cell-sieve cap and analysed using the BD FACS ARIA III flow cytometer using the BD FACSDiva software (BD Bioscience). After FACS sorting with a cut-off at 90% viable single nuclei, nuclei from the landing buffer (1% BSA, 0.2 U/µL RNAseIn) were counted using a digital counting chamber (Elvira) to achieve the concentration of 400-500 nuclei/µl and were loaded onto 10x Genomics Chromium chips. 10x Genomics single-index v2 and v3 libraries were prepared according to manufacturer's instructions (Chromium Single Cell 3' Kits v2 User Guide – CG00052, Chromium Single Cell 3' Kits v3.1 Dual Index User Guide – CG000315). Libraries were sequenced on an Illumina HiSeq-4000 (pair-ended) aiming for a minimum coverage of 50,000 raw reads per nucleus.

Data pre-processing and quality control

The processing and identification of Unique Molecular Identifiers (UMI) and nucleus barcode extraction of raw 3' snRNA-Seq data (demultiplexed FASTQs) was performed using Cell Ranger software (versions 3.0.2, 6.0.1 & 6.1.2) from 10x Genomics. Specifically, the SP014 (10X V2 library chemistry), SP082 and SP136 batches (10X V3 library chemistry) were processed with versions 3.0.2, 6.0.1, and 6.1.2 respectively. The transcripts were aligned against the pre-built human reference genome GRCh38 premRNA version 3.0.0, which was built from the GRCh38 precompiled reference (https://cf.10xgenomics.com/supp/cell-exp/refdata-cellranger-GRCh38-3.0.0.tar.gz), and modified for use with snRNA-Seq data by extracting "transcripts" features from the gene model GTF and instead annotating these as "exon", as described in the protocol defined by 10x Genomics (https://support.10xgenomics.com/single-cell-gene-

expression/software/release-notes/build#grch38_3.0.0). The unfiltered feature-barcode matrix per sample was considered for further removing ambient RNA and technical artifacts.

Subsequently, systematic biases and empty droplets were modelled and removed using the remove-background function implemented in CellBender v0.2.0 and setting model= “full.” Target false positive rate (FPR) was kept at default value of 0.01. The parameter --expected-cells was set to the estimated number of cells from CellRanger. After evaluating the CellRanger rank-ordered UMI plot, 15,000 was determined to be a suitable threshold for --total-droplets-included for modelling ambient RNA. For three villous samples (18-033-v, 18-098-v and 20-027-v), whose expected cell numbers from Cellranger were 16,811, 13,310 and 13,057 respectively, --total-droplets-included were set to default 25,000 and highly consistent number of cells were found after running CellBender (16,442, 12,922 and 12,839 respectively). The number of cells per sample achieved after CellBender filtering and their corresponding QC metrics (total UMI counts and number of expressed genes with at least one positive count per nuclei) are reported in Supplemental Tables S4 and S5 for placenta and decidua, respectively.

Filtered expression matrices were loaded into python v3.7.9 and further processed using scanpy v1.8.2. The post-quantification quality control was computed with the calculate_qc_metrics function implemented in scanpy. Nuclei having fewer than 200 expressed genes or for which the total mitochondrial transcript expression was higher than 5% were excluded. Only those genes expressed in more than three nuclei were included. Data quality was assessed by plotting the number of unique molecular identifiers (UMIs) and total number of genes per sample. After quality control filtering, the samples were log-normalised to 10,000 reads using scanpy. The resulting number of cells and corresponding QC per sample (total UMI counts and number of genes expressing one positive count per nuclei) were reported for both decidua and placenta in Supplemental Tables S4 and S5, respectively. On an average, the scanpy filtering removed 5.62% of nuclei after quality check for both tissue samples.

Doublet detection:

Initially, the neotypic doublets were predicted using Scrublet scr.scrub_doublets() function. The overall distribution of doublet scores was not bimodal, implicating that

our dataset was not systematically affected by potential doublets. Nonetheless, the identified doublets were flagged (using a threshold of 0.35 based on the doublet score histogram) and were cautiously carried forward to the downstream analysis instead of immediately filtering them out. Since droplets that appear to harbor doublets might arise from several effects they were cell-typed and evaluated in[41] (see Evaluation of biological doublets section).

Data harmonization, clustering, and cell annotations for placenta

For the data harmonization of placenta samples, firstly, 6000 highly variable genes were computed using scanpy's highly_variable_genes function, using the dispersion-based method (flavor='seurat_v3') and otherwise default parameters. The donor identifier was used as the key batch to minimize selection of batch-specific genes. Subsequently, the samples were integrated using scVI v0.14.5. Using, scvi.model.SCVI and get_latent_representation functions in scVI, a shared latent space of 15 dimensions for all placental single nuclei was inferred. Precisely, 128 nodes per hidden layer, 2 hidden layers used for encoder and decoder for the variational inference, and 0.1 drop-out rate was used. Zero-inflated negative binomial distribution (ZINB) was used to model gene expression. Apart from using donor_id (each sample) as batch key, further categorical covariates (10X library chemistry used, procurement centre of samples, gestational week) and continuous covariates (total counts, total number of genes with at least one positive count, percentage of mitochondrial expression, percentage of XIST expression per nucleus) were used to minimize the influence of technical variation in the cell typing.

The K-nearest neighbour graph was computed on the scVI inferred latent space using pp.neighbors function in scanpy with $k = 15$ and otherwise default parameters. To further reduce the high dimensional latent spaces to 2D, visualization was generated using Uniform Manifold Approximation and Projection (UMAP). Particularly, the umap-learn v0.5.2 implementation in python was used and the maximum number of iterations was set to 500 (for better convergence) and random state to 0 (for reproducibility).

Cell-typing (annotations) was initially performed on the control placental samples (both early and late gestation) based on robust and specific expression of marker genes (Supplemental Table S8). At first, clusters were identified in an unsupervised manner using Leiden community algorithm implemented in scanpy (with an initial resolution limit of 2) and initially annotated using marker genes extracted from literature plus top

signatures obtained from Seurat's FindAllMarkers Logistic Regression (LR) method and Empirical Bayes method by model.differential_expression function in scVI. Leiden clusters lacking robust/specific biological markers were merged into the closest cluster. Thereafter, a LR classifier model (optimized by the stochastic gradient descent algorithm) implemented in Celltypist v0.2.0 was trained based on our control cluster labels and was used to predict the cell annotations in diseased (eoPE) samples. A confusion-matrix was used to evaluate the performance of the classifier (predicted labels) given the known ground-truth (from Leiden clusters annotation). Spurious Leiden clusters mapping to a specific sample and lacking appropriate markers were removed. Particularly, a fibroblast (n=669) and erythroblast (n=930) subpopulation were excluded because they mapped solely to two specific early donors and hence, do not contribute to comparative cell typing. Additionally, 547 misclassified nuclei firmly clustering with vCTBp but also expressing high STB/EVT markers were excluded (with further help from the pseudotime analysis where these nuclei could not be modelled in a specific differentiation path). For further internal validation of cluster phenotype, we computed module scores using known markers list using the score_genes function in scanpy. Finally, all the clusters assigned to a phenotype (label) were evaluated using robust and specific marker genes (described in the Differential expression analysis section; also refer to Supplemental Table S5 and Supplemental Table S8).

Data harmonization, clustering, and cell annotations for decidua

Similar to placenta, the top 6000 highly variable genes were computed using scanpy's highly_variable_genes function using donor_id as the batch key. Here, the cell typing was initially performed on the 10X V2 samples (because they were sequenced earlier) by annotating unsupervised Leiden clusters based on robust and specific markers expression. Using the get_latent_representation function in scVI, a shared latent space of 10 dimensions was inferred keeping the other parameters same as used for the placenta. Like placenta, markers were extracted from literature as well as top signatures obtained from Bayes-method scVI model.differential_expression function and Seurat's FindAllMarkers LR method. Leiden clusters mapping uniquely to a donor were excluded for the purpose of comparative cell typing. Subsequently, the cell labels were transferred to the 10X V3 samples using scANVI.

In parallel to scANVI, a LR classifier model from Celltypist was trained using the annotated cluster labels and was used to predict the cell annotations in 10X V3

samples. A confusion-matrix was used to evaluate the performance of CellTypist classifier and scANVI (predicted labels). Ultimately, each cluster was inspected using biological markers knowledge and final decisions were made.

A cluster (initially annotated as stromal given its proximity to the DSC1/2 and consisting of 2911 nuclei) was later excluded because of the expression of conflicting markers such as *NOTUM, HPGD* and *HLA-G* (denoting EVT lineage). It also expressed certain macrophage genes and was difficult to classify. The CellTypist LR classifier assigned them a very low confidence score (~0) indicating the cluster was very likely contaminated. Another cluster (initially thought of as NKT cells; 1119 nuclei) were removed because of high macrophage gene expression. Finally, the cell type/state annotations of decidua were shown using a set of robust and specific markers genes as in Supplemental Figure S5 and Supplemental Table S8).

Cell-type/state composition per gestational time, condition, and biological replicate is reported in Supplemental Table S3. The resulting number of cells and corresponding QC metrics (total UMI counts and number of genes expressing one positive count per nuclei) for each placenta and decidua sample were tabulated in Supplemental Tables S4 and S5, respectively. 5.31% of additional nuclei were removed after filtering out donor-specific and ambiguous clusters in both tissues. The final UMAP embedding stratifying the cellular hierarchies for decidua and villi are shown in Figure 2. Importantly, the cell type/state level QC were also evaluated as displayed in Supplemental Figure S4 and Supplemental Table S6.

Evaluation of clustering robustness

To ensure the effective annotation of cell types/states, amortized Latent Dirichlet Allocation (LDA) implemented within scVI was used to find topic profiles for both tissues. Conceptually, a distinct cell types/state should map to a unique topic. Subclusters share the topics of the mother cluster and in addition, usually harbor a unique topic. For example, dNK1 and dNK2 have shared as well as unique topics. This modelling approach can also be used to identify potential doublets when cells exhibit multiple conflicting topics (mainly due to opposing lineage markers), similar to marker-based approaches used in other single cell studies of placenta and decidua[42].

LDA was performed at several stages (initially using the number of Leiden clusters equal to the number of topics and ultimately, to the number of final labels) to see if the

learned topics were mainly dominant in cells close together in the UMAP space. Problematic clusters were confirmed to not map to unique/known topics and hence removed from all downstream analysis.

Evaluation of biological doublets

Scrublet predictions of neotypic doublets in placenta were polarized towards vSTB1. Out of 2443 total doublets, scrublet predicted 1656, 136 and 50 doublets for vSTB1, vSTB2 and vSTBim cell states (accounting for 2.8%, 1.02% and 1.90% of their populations) respectively, which were investigated further. Precisely, for each cluster, the marker gene expression was investigated, separated by those predicted doublets and singlets states which did not reveal cross contamination of marker genes between predicted doublets and non-doublets for the vSTB states (Supplemental Figure S2). Overall, no doublet centric clusters were found and no other cell types/states were confounded with conflicting marker expression.

Syncytiotrophoblast sub-clusters

Immature STB were marked by the expression of paternally imprinted *DLK1* (regulator of cell growth and differentiation), *SPARC, TMSB10,* and *ACTB* which indicate an association with extracellular matrix remodelling and promotion of changes to cell shape. The immature nuclei robustly express the secretory phenotype (characterized by several PSGs and maternally imprinted *TFPI2*) and a classical STB like profile (expression of *CGA, CYP19A1, KISS1, ADAM12, SDC1* and others) for which it was classified under vSTB group.

Other placenta cell types/states

vCTBp was considered as the trophoblast progenitor as they are actively cycling given the expression of genes like *MKI67, TOP2A, STMN1* and *CENPK/CENPE*. They exhibit robust expression of *YAP1, TEAD1, TP63, CCNA2, ITGA6*- all known for their roles as progenitor. vCTBpf is primarily fusogenic and is characterized by very specific markers such as *GREM2, ERVFRD-1, ERVV-1/2, OTUB2,* and *DYSF*.

Placental *F13A1+/FGF13+* resident macrophages (Hofbauer cells, vHBC) uniquely express hyaluronan receptor *LYVE1* in the immune cell subset. We additionally identified antigen presenting *HLA-DRA+* placenta associated maternal monocytes/macrophages (vPAMM. A cycling population of HBC was identified

(vHBCp) characterized by traditional HBC genes as well as proliferative genes like *MKI67* and *TOP2A*.

Evaluation of integration performance

The performance of data harmonization/integration was evaluated using the Adjusted Rand Index (ARI), Adjusted Mutual Information (AMI) & cell-type adjusted silhouette width (ASW) per batch as used in prior integration benchmarking studies.[43]

Our strategy investigated batch effect correction across 4 analyses: (1) technical replicates to establish an upper bound of integration quality; (2) different library preparation methods; (3) sampling sites; (4) sampling site and library preparation (Supplemental Figure S1, Supplemental Table S7). First, we evaluated batch-effects between technical replicates. Batch effects in our experimental design were carefully modelled by including technical replicates for a sample (577_1 and 577_2), to establish an upper bound for quality of integration. Both samples had comparable numbers of cells (6,185 and 6,081), median genes expressed (788 and 756) and similar cell type compositions. The relative proportion of cell types/states in two technical replicates was depicted using kernel density plot (KDE) that reflected the density of cells in an embedded space and concordance in cell composition. As expected, both the ARI (0.0037) and AMI (0.003) score were close to 0, indicating near perfect integration. Similarly, the average ASW per batch (0.92) was close to 1, also indicating very decent batch effect removal. Furthermore, ASW was evaluated at the cell-type level which indicated excellent scores for a larger number of cell types, however, intermediate scores for villous myocytes and slightly lower for villous T-cells (Supplemental Table S7). This is why we did not focus on these cell types in our study.

Second, we investigated the role of library preparation chemistry. Similar to establishing an upper bound to similarity of samples through analysing technical replicates, a more extreme effect of using different library preparation (10X V2 and 10X V3 chemistry) was investigated. In case of early (first-trimester pregnancy), samples were processed using 10X V3 (n=7) and 10X V2 (n=3) chemistries. The groups showed 67,037 and 12,848 cells, and median genes expressed 3,088 and 1,465 using 10X V3 and 10X V2 chemistry, respectively. The relative cell type composition was also similar with relatively more vHBCp, vCTBp, vFB, vMC, vHBC, and vPAMM in 10x V3 samples. ARI and AMI (0.002, and 0.009) values close to zero suggested negligible

effect of library chemistry. The average ASW per batch was 0.88 indicating a decent removal of library effects. Furthermore, ASW was evaluated at the cell-type/state level which indicated excellent scores (>0.90) for a larger number of clusters relevant for the first-trimester pregnancy such as vCTB, vCTBp, vCTBpf, vCCT and vSTB1 (Supplemental Table S7).

All late term controls were processed using 10X V2 chemistry. Hence, the library was not a confounder. We present differences in eoPE samples in Case 4.

Third, we investigated the role of sampling site (procurement) in late samples. Late samples were split equally between both Oslo and Graz sites (n=3 and n=3, respectively). Both samples had comparable numbers of cells (15,733 and 13,969), median genes expressed (827 and 894), and similar cell type compositions with some exceptions. The ARI (0.016) and AMI (0.027) were close to zero indicating that our cell type labels were not influenced by sampling site. The average cell type ASW per batch was 0.88 indicating appropriate removal of sampling effects within a cell type identity. Cell-type ASW per batch was decent (> 0.90) for the major cell types such as vCTB, vSTB1/2 & vSTBim indicating optimal batch-mixing while conserving biological information. There were increased proportions of PAMM, vT-cells, vFB and vMC in Graz samples; hence, no key conclusions were drawn from these cell populations in the manuscript. Despite observing slightly more vCTB in Graz samples, they mixed well between site as reflected by ASW (Supplemental Table S7; Sampling site late controls).

Fourth we investigated the role of sampling site (procurement) and chemistry in eoPE samples. Further, eoPE samples were split between both Oslo and Graz sites (n=3 and n=2, respectively). The numbers of cells for each site were 17,604 and 12,365, median genes expressed 790 and 1892 genes respectively, and similar cell type compositions. A higher number of profiled genes is because the Graz samples were processed using 10X V3 chemistry compared to 10X V2 used for Oslo samples. The ARI (0.004), and AMI (0.028) scores were close to zero indicating that our cell type labels were not majorly influenced by sampling sites and library chemistry. The average ASW was 0.80 indicating good integration. ASW for major cell-types such as vSTB1/2/im were very high, which rules out influence of sampling site (Supplemental

Table S7; Sampling site eoPE). However, we observed relatively more vFB in Graz eoPE samples that might have lowered its ASW. Immune cells such as PAMM & vTcells were slightly depleted in Graz samples relative to Oslo, which apparently also have lowered their ASW score. However, no major conclusions were drawn from these cell types in the manuscript.

Our results suggested that cell types/states were not majorly confounded by individual donors. Cell-type ASW per batch is high for most of the individual clusters (and split across conditions). Cell types central for eoPE and term comparative analysis (like, vSTB subgroups) showed an exceptionally high ASW score which indicates optimal batch-mixing while conserving biological information (Supplemental Table S7). Our metrics were in par with the ASW per batch reported by Luecken et al. 2021.[43] Importantly, our average global ASW per batch for our technical replicates (0.92) revealed near perfect integration, which is also roughly comparable to our overall integration using all samples (0.863).

Differential expression analysis

Cell-type marker analyses for both decidua and placenta were performed using multi-variate LR generalized linear model implemented in Seurat's FindAllMarkers() and were further internally validated using the empirical Bayes method in scVI model.differential_expression function.

In the case of LR, the number of UMI, number of genes, and percentage of mitochondrial transcripts per nuclei were used as continuous covariates. Additionally, ~disease (if a nucleus is from a control or eoPE sample) and library (10X V2 or V3 chemistry) were used as categorical covariates to minimize the effects of eoPE and libraries. Only genes having a log fold-change cut-off of 0.25 and expressed in at least 25% of cells within each cluster were considered significant cell markers. An adjusted p-value cut-off was kept at 0.05 (after Bonferroni correction for multiple testing; Supplemental Table S9).

Integration of publicly available placenta scRNA-seq data with our snRNA-seq data

The 10X scRNA-seq data from Pique-Regi et al[11] were harmonized with our own 10X snRNA-seq data to identify differentially expressed preterm gene-sets per cell

type/states that can be further applied for correcting gestational age or preterm effects during eoPE vs term control analysis.

The data integration was performed on only the pertinent samples: eoPE and late term control samples from our study and preterm and term samples from the Pique-Regi study separately for villous and decidual samples. Given the lack of pre-processed count matrices from the Pique-Regi study, the raw FASTQ data from dbGAP accession phs001886.v4.p1 were downloaded and feature-barcode count matrices were generated using CellRanger 6.1.2. The results were highly concordant with a mean difference of only 43.9 cells compared to the original study, all with highly similar UMI and gene counts. Thereafter, data harmonization and label-transfer were performed using scVI/scANVI (based on Bayesian variational inference). Precisely, donor_id (each sample) was used as batch key, further categorical covariates (dataset: scRNA-seq vs snRNA-seq, 10X library chemistry: 10X V2 vs 10X V3, procurement centre of samples: Oslo, Graz & Detroit, condition: preterm vs term, fetal sex) and continuous covariates (total counts, total number of genes with at least one positive count, percentage of mitochondrial expression, XIST counts per nucleus) were used to minimize the influence of batches and technical variation in the cell typing.
Projecting the cells into UMAP embeddings revealed good batch mixing of the datasets for the conserved cell types/states such as vCTB as shown in Supplemental Figure S3. Integration was evaluated using adjusted rand index (ARI), adjusted mutual information (AMI) and cell-type specific absolute silhouette width (ASW) per batch. To be precise, ARI and AMI were 0.051 and 0.118 respectively, that indicated good integration for the placental clusters (Supplemental Table S7). Furthermore, mean ASW (scaled between 0-1; where 1 means perfect integration) per donor was 0.803 with fairly high scores per cell type (Supplemental Table S7). For the decidua, the ARI and AMI were 0.03 and 0.12 respectively; mean ASW per donor was 0.826 indicating appropriate integration.

After integrating the data, differential gene expression analysis was performed while accounting for batch-effects. For two cell-state/class populations of interest (preterm vs term controls), differentially expressed genes were identified using scvi.model.SCVI.differential_expression() using the ‘change’ DE mode[44] and adjusted for covariates specified during integration. We filtered the list using FDR < 0.05 and

Bayes factor > 3. The term controls for each cluster were downsampled to match the cell numbers in preterm controls to avoid composition specific effects. The preterm genes were not identified in vFB, vMC and vVEC cell groups given extreme sparsity in preterm control groups. For the decidua, the dNK1/2 and dMAC1/2 subclusters were merged into dNK and dMAC classes for preterm gene calculation. Preterm genes were not identified for dVEC, dLEC and dLECp owing to sparsity of these populations in the preterm control group. The preterm genes for both placenta and decidua cell types/states are reported in Supplemental Table S2.

After identifying preterm genes, a preterm score per cell types/states was computed and ultimately, eoPE associated genes were identified as described below.

Identification of preterm genes in vSTB

Given the lack of publicly available single cell/nucleus data describing preterm STB(s), we instead leveraged the fact that STB accounts for the majority (~90% of nuclei) population in the villi. This rationalized the use of bulk gene expression data to represent the transcriptome of STB nuclei. On this note, dataset published by Leavey et. al 2015[16] which profiled a large set of villous samples using bulk gene expression profiling deposited under GEO accession GSE75010 was used. The data was downloaded using GEOquery R package and background subaction, quantile normalization, and summarization of raw microarray probe intensity values were performed with RMA RMA16 function from the oligo R package. Custom CDF from Brainarray project, version 25.0.0, was used to summarize and annotate the probes to Entrez Gene ID. Quality control was done with ArrayQualityMetrics and factoextra (http://www.sthda.com/english/rpkgs/factoextra) R packages.

Samples were identified from that study that were most suitable for gestational age correction based on matching the characteristics of our cohort: age: 18 – 40, BMI: 18.5 – 35 (WHO normal to obese), delivery mode: caesarean section, no FGR and no chronic hypertension. This selection resulted n=10 preterm (<34 weeks) samples and n=16 term (>37 weeks) samples on which differential expression analysis was performed with the limma R package.[45] All the computations were done using Bioconductor version 3.16 (BiocManager 1.30) and R version 4.2.2. A set of 13 genes was found to be differentially expressed between preterm and term (BH-corrected adjusted p-value < 0.01 and additionally tested for sex-bias) that were also expressed

to varying extents in our STB populations. Thereafter, a STB specific preterm score was computed for each STB state using this set of 13 genes and utilized as a confounding variable in our Logistic Regression model when identifying eoPE specific genes (described below).

Differential analysis of eoPE disease markers and gestational age correction

To determine the differentially expressed genes for disease (eoPE) against late controls, the LR framework (implemented Seurat's FindMarkers function) was applied to respective cell types/states. The number of UMI, gene counts, percentage of mitochondrial transcripts and percentage of sex-specific transcripts per nuclei were used as continuous covariates.

Importantly, a cell type/state specific preterm score was calculated using the preterm vs term significant genes (using microarray approach for vSTB groups and scVI integrated Pique-Regi et al[11] data for other cell types/states as described in the above section) and used as a continuous covariate in the LR model. This was explicitly performed to prevent strong preterm specific effects in the analysis since eoPE arises 6-8 weeks prior to healthy term.

Only those genes having a log2 fold-change cut-off of 0.25 and expressed in at least 10% of cells within each group were reported as significant given adjusted p-value < 0.05 (Bonferonni corrected). Both up- and down-regulated genes were computed. No cell type/state exhibited significant composition shift in eoPE relative to term controls (except for vHBC); hence, down sampling was only performed for the vHBC. For cell types such as vCCT, dEVT, vCTBp, vCTBpf, dDSTB, dPC, dBcells no analysis was performed owing to extreme sparsity in eoPE group. None of our samples were confounded with a major co-occurring disease. The eoPE markers are reported in Supplemental Table S9.

Reconstruction of differentiation trajectories, lineage relationships and computation of pseudotime genes

To infer the cluster and lineage relationships between the different trophoblast cell types/states, STREAM v1.1. (https://github.com/pinellolab/STREAM) and diffusion pseudotime were used. Specifically, the trajectory inference was restricted to the early controls of the trophoblast cell types including vCTBp (progenitor), vCTB, vCTBpf,

vSTBim, vSTB1/2 and vCCT. In the late term controls, there is a striking discrepancy in the cell-type composition given a massive increase of vSTB sub-populations which signifies degradation rather than differentiation.

At first, the scVI harmonized control data as subsetted for the relevant cell types and learned the trajectory principal graph using STREAM 1.1. Using previously computed latent variables, cells were clustered in the reduced UMAP space for recovering the main and possibly finer structures of trophoblast differentiation. Thereafter, the principal graph was inferred on the manifold learnt from the dimension_reduction function using the first six components. K-means clustering was used for the initial graph seeding using seed_elastic_principal_graph(). The elastic principal graphs are structured data approximators, consisting of each cell as a vertex interconnected by edges. The inference of this graph relied on a greedy optimization procedure based on which a minimum spanning tree (MST) was constructed using the Kruskal's algorithm. No branch pruning or shifting of nodes were performed to obtain the optimal principal graph (Supplemental Figure S6).

Ultimately, the transition and leaf markers were computed for all lineage paths (vSTB, vCTB and vCCT) by considering MKI67 positive vCTBp as a root node (start of the pseudotime) respectively.

The transition genes were calculated by considering fold change in average gene expression of the first 20% and the last 80% of the cells for an individual branch based on the inferred pseudotime. For the genes exhibiting log2 fold change cut-off of 0.20, further Spearman's rank correlation was calculated between pseudotime and gene expression of all the cells along the individual branches (as implemented in STREAM's detect_transition_markers function). Ultimately, genes above a specified correlation threshold (=0.35) were reported as transition genes. For leaf gene detection, the z-scores of all leaf branches were calculated based on the average normalized gene expressions. Particularly, Kruskal–Wallis H-test followed by a post-hoc pairwise Conover's test (as implemented in STREAM's detect_leaf_markers function) was used for multiple comparisons of mean rank sums test among all leaf branches. A z-score cut-off of 1, and p-value cut-off of 0.01 were used to identify the candidate leaf genes. The expression of highly robust cell fate markers along the pseudotime provided a strong validation for our trajectories.

To further evaluate the lineage relationships and global transcriptomic similarity between different cell types (for trophoblast), Diffusion pseudotime analysis (Haghverdi et. al 2016) was performed that orders cells based on their transcriptomic similarity in a Markovian space. This method considers each cell to be represented by a Gaussian wave function and diffusion distances are based on a robust connectivity measure between cells, which is estimated over all possible paths of a certain length between the cells. The Eigen functions of the Markovian transition probability matrix (diffusion components; DC1 and DC2) were used for low-dimensional representation and visualization of trophoblast data. Additionally, a force-directed graph based on Fruchterman-Reingold algorithm was shown (Supplemental Figure S6).

Receptor-Ligand interaction databases

An extremely important factor deciding the results of the R-L interaction study is the underlying database used. Two popular databases, CellChatDB and FANTOM5, were used that allowed identification of well-established interactions such as MIF-ACKR3/CXCR7 and INHBA-ENG/END, which were unique to CellChatDB and FANTOM5 respectively.

Receptor-Ligand interaction differential analysis of eoPE vs term controls

The differentially expressed ligand-receptor interactions were inferred using Connectome v1.0.1 (specifically, the differential connectomics pipeline). For the maternal-fetal interface, the strategy was to use only secretory ligands for vSTB groups that can practically cross the maternal-fetal barrier and can be in contact with the maternal blood (decidua) where it can influence the vessels. Only significantly differentially upregulated ligands (p-value 0.01 after adjusting for multiple testing) in eoPE relative to term controls exhibiting a log2FC cut-off of 0.25 and detected in a minimum of 10% of diseased cells were considered. It was assumed that once a ligand is activated (upregulated), it would bind the receptor (irrespective of the latter being differential or not). Biologically, we can describe such instance as ligand pressure (where ligand is high, but receptor is either non-differential or low; Figure 6A). Multivariate Logistic Regression was used for differential calculation and the statistics are consistent with our former described DEG test (including covariate corrections).
For the within tissue interaction map (decidua and villi interaction) using immune and endothelial cell types, LR and Connectome were used. The log2FC cut-off of 0.25 and

ligands/receptors detected in a minimum of 10% of diseased cells were considered for both up and downregulated differential candidates. The p-value was adjusted for covariates (as described for eoPE vs late term DEG). Visualization was performed using circos plots.

Receptor-Ligand interaction analysis in eoPE

The analysis for decidual STB and EVT ligands (with maternal VEC and dSMC) were restricted to eoPE samples only given their extreme sparsity in late term. Interactions were derived using Connectome using both FANTOM5 and CellChatDB databases. The min.cells.per.ident was kept at 75 and Diagnostic Odds Ratio (DOR) was calculated for each interaction pair. High DOR is an indicative of high specificity and sensitivity with a low rate of false positives and false negatives. For dDSTB, the interaction list was filtered using pct.source (senders) >= 25% and pct.target (receivers) >= 20% (pct= percentage of cells expressing a ligand/receptor) and further, filtered by DOR.source > 3 and ligand expression > 1.5. For the interactions with dEVT, relatively robust criteria were used for narrowing down the important interaction partners (from an initial list of > 2000 pairs). Particularly, DOR.source of 5, edge strength (product of the receptor and ligand expression) of 3 and minimum percentage of ligand expressing source of 50% was used to ensure cell specific communication. The R-L figure was shown in Supplemental Figure S8.

Computational validation of major R-L interactions

All Connectome results were cross checked SingleCellSignalR for the vSTB, dDSTB and dEVT based interactions. All the R-L interactions were recapitulated (when not limited by database).

Subsequently, we applied additional tools (NATMI, logFC Mean (inspired by iTalk), CellphoneDB, CellChat) implemented within LIANA framework[46] and we were able to recapitulate all R-L interactions across numerous databases. The results are summarized in Supplemental Table S11.

Pathway and network analyses of marker genes

The list of DEG based on cut-off values (logFC +/-0.25 and a significance level of 0.05) were used as background for networks. Variable genes were excluded using the webtool diVenn (divenn.noble.org). Genes were then used as input in the stringDB for PPI networks (confidence level = 0.15, no added proteins in shells). Networks were

then further analysed in Cytoscape (version 3.8.2). Hub genes, defined as genes with high correlation in candidate modules, were identified from the PPI network calculating top 5 genes of all topological analysis methods of CytoHubba in Cytoscape plug-in (DMNC; MNC, MCC, ecCentricity, Bottleneck, Degree, EPC, Closeness). The candidate hub genes were merged into one network, decomposed into communities using clustermaker and GLay Cytoscape plug-in based on Newman and Girvan's edge-betweenness algorithm. The hub network was analysed to visualise the network degree of nodes by size of nodes. The original background logFC was used for continuous mapping colours. The hub gene network was used to calculate transcription factors via iRegulon cytoscape plug-in (standard threshold: enrichment score threshold 3.0, ROC threshold for AUC calc 0.03, Rank threshold 5000, minimum identity between orthologous genes: 0.0, max FDR on motif similarity: 0.001). Predicted transcription factors were visualised as PPI (confidence level 0.15) via stringDB and validated by adding expression values from the DEG list.

Pathway analyses are based on these DEG lists, hub genes, and transcriptions factors and were carried out via web-tools Metascape and Enrichr.

In Situ **Sequencing**

<u>High sensitivity library preparation</u>

Fresh tissue samples of early villi were FFPE processed and stored at +4°C. A custom gene panel was used to detect specific cell-type and cell pathway genes of interest. The *in situ* sequencing method was processed according to manufacturer instructions (Cartana, part of 10x Genomics). 5µm tissue sections were baked at 60°C for one hour, deparaffinised in xylene, rehydrated in 100% and 70% ethanol, and permeabilised using citrate buffer (pH 6) for 45min at >95°C in a steamer. Sections were dehydrated in an ethanol series from 70 to 100% and air-dried (Secure Seal, Grace Biolabs, Bend, United States). Gene specific chimeric padlock probes were added, directly hybridised to the RNA at 37°C in an RNAse free humid chamber overnight and ligated at 37°C for 2 hours. Ligation derived circular oligonucleotide structures (padlocks) amplified overnight at 30°C. RNA-degradation during tissue processing was minimised by adding 0.1% v/v diethyl pyrocarbonate (DEPC) to all buffers and reagents not provided by the manufacturer.

Imaging

Imaging was performed using a digital slide scanner (Olympus SLIDEVIEW VS200) connected to external LED source (Excelitas Technologies, X-Cite Xylis). Fluorescence filters cubes and wheels were equipped with a pentafilter (AHF, excitations: 352-404 nm, 460-488 nm, 542-566 nm, 626-644 nm, 721-749 nm; emissions 416-452 nm, 500-530 nm, 579-611 nm, 665-705 nm, 767-849 nm) and single cube filters (Kromnigon; SpectraSplit 440, SpectraSplit 488, SpectraSplit Cy3, and SpectraSplit 594). Images were obtained with a sCMOS camera (2304 × 2304, ORCA-Fusion C14440-20UP, 16 bit, Hamamatsu), and Olympus Universal-Plansuperapochromat 40× (0.95 NA/air, UPLXAPO40X). To avoid signal cross-talk, the pentafilter was used to image DAPI, Cy5 and AF750 signals, and the single cubes to image AF488 and Cy3 were used. Imaged regions were recorded to perform repetitive cycle imaging. After imaging, labelling mix was stripped from each slide by adding three times 100% formamide for 1min, followed by a washing step.

Hybridizing and Sequencing

In situ sequencing steps were repeated six times with six different adapter probe pools, each imaged in five channels (DAPI, FITC, Cy3, Cy5, AF750). After stripping, adapter probes were hybridised at 37°C for 1 h in a RNAse free humid chamber, washed and sequencing probes hybridised at 37°C for 30 min in a RNAse free humid chamber. Sections were washed, dehydrated in an ethanol series, air-dried, and mounted with SlowFade Gold Antifade Mountant (Thermo Fisher Scientific). Library preparation protocols were optimised for placental tissue using high (*MALAT1*) and low (*RPLP0*) control probes before using the final probe panels. Background without any adapter probe pool was imaged in 6 channels for autofluorescence subtraction.

Image analysis and spot calling

Imaging data was analysed with the custom pipeline provided by CARTANA that handles image processing and gene calling. All code was written in MATLAB and additionally a CellProfiler pipeline (v.2.1.1) was used, that includes the ImageJ plugins MultiStackReg, StackReg and TurboReg as previously described. In short, TIFF images from all sequencing cycles were aligned to the general stain of library preparation and split into multiple smaller images. The median intensity of all RCP signals of each channel was calculated with an additional CellProfiler pipeline

(v.4.0.7), this value was used to normalise RCP signal intensities of each channel to a pixel intensity of 10,000. The received multiplication factor value for each channel was integrated in the CellProfiler pipeline and the background of each channel subtracted from each sequencing cycle, to reduce the autofluorescence of the tissue. A pseudo-anchor was created for each cycle by making a composition of the four readout detection probe channels into one merged image. The pseudo-anchor was used to perform a second, more exact alignment. RCPs of the labelling mix were detected, x and y coordinates saved and fluorescence intensities measured. The highest intensity value in each sequencing cycle was assigned as positive event and used for decoding in Matlab. For signal visualisation, the selected transcripts were plotted on a DAPI-stained image.

In situ **sequencing data analysis**

In situ sequencing data handling

In situ sequencing data handling was performed using the plankton.py v0.1.0 package (https://github.com/HiDiHlabs/planktonpy) in python 3.10.4. For the conclusive analysis routine, the *in situ* sequencing data was displayed as decoded spots of x and y coordinates of all detected mRNA molecules, each with an associated gene label. In total, three *in situ* sequencing slide scans were analysed (106KS, 107KS, 156KS). 156KS (early control) contained genes from the customized placenta/cell typing pane l that was designed for retrieving cell and tissue types, and both 106KS (late control) and 107KS (eoPE) contained genes from the custom/pathway panel which was targeted at analysing cell state and metabolic activity (code book for panels available via zenodo doi: 10.5281/zenodo.5243240). The cell typing sample was taken during the early stage of pregnancy.

For visualization of the detected mRNA molecules in their histological context, matching DAPI stains of each sample slide were pre-processed by transforming it to grayscale, normalizing the colour values between 0 and 1 and pushing the low-exposure areas by raising all values to the power of 0.4.

Identification of cell type specific markers in the placenta panel

Analysis of the placenta cell typing data had the aim of contextualizing major cell types determined by snRNA-Seq analysis. Genes from the cell typing panel were conceptualized as cell type markers, with CTB, STB and HBC cell types considered

for further spatial analysis and plotting since they had good marker coverage and constituted important spatial landmarks in the villous anatomy (with the walls being layered with STB and CTBs, and HBCs forming distinct, compact cells in the intra-villous matrix).

Accordingly, a gene-cell-type affinity measure was derived through the gene molecule counts in the snRNA-Seq data set for CTB, STB and HBC. This was done per gene by contrasting molecule counts in the cells belonging to a cell type of interest in the snRNA-Seq data against the molecule counts of an opposing set of cells using plankton.py's score_affinity function. Hence, for each analysis, two contrastive sets of cell types were defined: (i) CTB vs STB for CTBs; (ii) STB vs CTBs for STB; and (iii) HBC vs all other cells for HBCs. The transcriptome of CTB and STB could be expected to be more similar since both are trophoblasts. Therefore, to determine definitive cell type markers, CTB and STB were contrasted against each other, which would cancel out potential common trophoblast marker genes. Each genes' mean molecule counts in all cells assigned to the two contrastive cell type sets was determined. The logarithm of the ratio of these mean count indicators was used as score for a gene's affinity to a certain cell type. To improve visual clarity during plotting, a threshold of 0.5 was used to assign colour labels to each gene in both analyses. Genes exceeding this affinity score threshold determined markers for CTBs (*ASPM, ATAD2, BRIP1, CD24, CDH1, CENPE, DIAPH3, FBN2, KANK1, SEMA6D, TIMP3*), STBs (*ADAMTS20, CGA, CYP19A1, ENTPD1, KISS1, KLRD1, LEP, LINC00474, MYCNUT, PAPPA2, PLAC4, PLXDC2*), and HBCs (*CD163L1, CD36, F13A1, FGF13, LYVE1, MEF2C, SPP1*), with the remaining genes assigned as 'other'.

<u>mRNA molecule spatial context analysis in the placenta panel</u>

Using the plankton.py's run_umap function, a weighted neighbourhood graph was built using the 800 nearest neighbours of each molecule, with neighbours weighted according to their Euclidean distance using a Gaussian probability density function (PDF) at a bandwidth of 9 µm, which would roughly cover the area of a single cell and its immediate environment. Then, a model of local mRNA distribution was created over all genes by summing over each gene's molecules' weights. Furthermore, a regularization mechanism was introduced by increasing each distribution's value of the gene of its molecule of origin by 1.15.

A 2D embedding of recurring spatial context was determined by applying the python umap-learn v0.5.3 algorithm to the local distributions. The number of neighbours in the UMAP algorithm was set to 24 and the minimal distance of points in the embedded representation was set to 0.2. UMAP used a Euclidean distance metric and was initiated at a random state of 42. The determined gene-cell type associations were used in a cell type visualization plot of the early placenta sample 156KS. All molecules were plotted as a scatter plot on top of the greyscale renderings of the DAPI stain. Molecules with a cell type-association affinity score above 0.5 were coloured accordingly, while the remaining molecules were rendered as grey.

Villous wall detection

Having demonstrated the principal plausibility of the spatial information in our *in situ* sequencing data during the cell-typing analysis, our experimental design required a follow up comparative analysis of pathways between a late control and an eoPE sample. The pathway categories chosen for the analysis of this second experiment were ‘vascularization’ (genes *IDO2, ZEB1, TEK, CDH5, KDR*), senescence (genes *MMP11, INHBA*) and trophoblasts (genes *LGR5, FGFR2, MET*).

Spatial analysis was restricted to the densely populated and well-structured villous walls in both samples, as this is the most structured part of the tissue. Villous walls were determined using a basic edge detection algorithm applied to matching DAPI signal, where villous walls were clearly remarked by dense nucleation. A greyscale rendering of the DAPI stain was smoothed using an optical Gaussian filter at a 2 µm bandwidth. Scikit-image's (v0.19.2) feature.canny() implementation of the canny edge detection algorithm (using a sigma value of 3.7) was used to extract the villous walls in the smoothed image. Every molecule within a radius of 5 µm to any point of the detected edges was defined as being part of the villous walls, and all other molecules were discarded from further analysis. The wall filter algorithm was visualized by plotting the underlying DAPI stain in matplotlib’s violet-blue ‘magma’ colour scheme. The detected edges from the second step of the wall filter algorithm were plotted on top of the stain as orange lines. In the bottom-right half of the plot, the present mRNA molecules were plotted, coloured green or blue according to their assignment as wall/not-wall members (Supplemental Figure S10).

Spatial relationship of vascular and senescence markers

For visualization of the spatial senescence-vascularization relation, all ‘wall’ molecules were plotted on top of a black-and-white rendering of the DAPI stain, coloured according to their gene assignments of ‘vascularization’ (red) and ‘senescence’ (yellow), with the other molecules plotted in white. Optical inspection of the scattered senescence and vascularization markers hinted that senescence marker topography was more structured in the control sample as compared to the eoPE sample, with senescence marker expression being reduced around vascularization clusters in the tissue.

To statistically model this observation, the villous wall molecules were subdivided into two categories depending on their location of expression: (i) a vessel proximal category that contained all molecules within regions of 5 µm of another vessel marker; and (ii) a vessel-distal category that contained the other molecules. A null hypothesis was formed, according to which the distributions of genes should be equal within the two categories. For each gene, we reported the p-value of the violation of this null hypothesis in a binomial test. Scipy’s (v1.8.1) stats.binom.cdf implementation was used, with parameters ‘p’ defined as the total percentage of ‘proximal’ molecules, ‘k’ the gene-specific number of proximal molecules and ‘n’ the total number of molecules of the respective gene in the sample. The sorted p-values for all genes present in the pathway sample were displayed in a vertical bar graph, with the bars coloured according to their membership to the categories ‘senescence’, ‘vascularization’, ‘trophoblast’ or ‘other’ (Supplemental Figure S10). The p-values of senescence, vascularization and a control category of ‘trophoblasts’ were extracted and plotted per sample as scatters on a vertical line. The scatters of both samples were displayed next to each other for visual comparison (Figure 5E, Supplemental Figure S10).

RNA isolation and RT-qPCR

Cell pellets or pulverized tissue were lysed in QIAzol lysis reagent (Qiagen, Austin, Texas). RNA was isolated according to the manufacturer’s instructions (AllPrep DNA/RNA/Protein Mini, Qiagen, Austin, Texas). In samples from the cohort from Melbourne, RNA was isolated according to the manufacturer’s instructions (GenEluteTM Mammalian Total RNA Miniprep Kit (Sigma-Aldrich, St. Louis, USA). RNA quality was determined using an Agilent 2100 Bioanalyzer (Agilent Technologies,

Santa Clara, CA, USA). Quality check was followed by reverse transcription of 1 µg total RNA per reaction using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster City, CA, USA), according to the manufacturer's manual. For Graz cohort, qPCR was performed with Blue S'Green qPCR Kit (Biozym, Vienna, Austria) using a Bio-Rad CFX96 cycler. For Melbourne cohort, a total of 1 µg placental RNA was reverse transcribed to cDNA using the High-Capacity cDNA Reverse Transcription Kit (Life Technologies, Carlsbad, USA), as per the manufacturer's instructions. RNA was converted to cDNA using the iCycler iQ5 (BioRad, Hercules, USA) and qPCR was performed on the CFX384 (BioRad, Hercules, USA). For all other qPCRs, the QuantStudio 3 Real-Time PCR System (Applied Biosystems) with either TaqMan Fast Universal PCR Master Mix or Fast SYBR Green Master Mix (both Thermo Fisher Scientific) were used. Primer and probes (as reported in table below) were designed using Real-time PCR (TaqMan) Primer and Probes Design Tool (online tool) from GenScript and synthesized by BioTez (Berlin Germany). Primers were diluted to a final concentration of 10 mM, probes to 5 mM. The target mRNA expression was quantitatively analysed with standard curve method. All expression values were normalized to the housekeeping gene *18S or TBP and for Melbourne cohort, to TOP1 and CYC1.* Validation cohorts were analysed individually and for combined presentation merged by z-transformation.

Target gene (human)	Forward sequence	Reverse sequence	Probe sequence
18S	5´ ACA TCC AAG GAA GGC AGC AG 3´	5´ TTT TCG TCA CTA CCT CCC CG 3´	5' FAM-CGC GCA AAT TAC CCA CTC CCG ACA-TAMRA 3'
GDF15	5' TGGGAAGATTCGA ACACCGA 3'	5' CCCGAGAGATACGC AGGTG 3'	5' FAM-CTGGGATCCGGCGG CCACCT-TAMRA 3'
INHBA	5' GCAGACCTCGGAG ATCATCA 3'	5' GAAATCTCGAAGTGC AGCGT 3'	NA
SERPIN E1	5' GTTCTGCCCAAGTT CTCCCT 3'	5' ACATGTCGGTCATTCC CAGG 3'	NA

NA, not applicable.

First trimester Serum ELISA Measurement

Women were recruited in the first trimester of pregnancy and a serum sample was taken before risk assessment via the FMF algorithm[47]. We excluded women with comorbidities such as chronic hypertension or diabetes mellitus and proceeded to match women that developed early onset preeclampsia to controls 1:2 (n=28 vs n=56, Supplemental Table S1). The matching was based on the variables maternal age, gestational age at first scan, and BMI. This was done using the R package *Matching*, which finds for each case two matching controls that minimise the weighted distance of their matching variables. We excluded patients that were prescribed prophylactic Aspirin from being part of the control group to reduce confounding. The serum samples from the 84 selected case and control patients (matched on maternal age, GA at first scan, and BMI) were then analysed for leptin, perlecan, GDF15 and activin A according manufactures protocol: human HSPG (Perlecan) ELISA Kit (ab274393; Abcam), human GDF-15 Quantikine ELISA Kit (DGD150; R&D Systems), human/mouse/rat Activin A Quantikine ELISA Kit (DAC00; R&D Systems) and human Leptin Quantikine ELISA Kit (DLP00; R&D Systems). Due to missing samples and one sample that was removed after unreliable measurements, the group sizes were eoPE: n=27 and healthy term controls: n=49. A conditional logistic regression model was fit to the new data with predictor variables being included using forward selection. A conditional logistic regression model was calculated as absolute model without prior risk assessment based on the cohort published earlier[48], a second model included the risk assessment by the FMF algorithm as offset. ROC curves and AUC are calculated from the method described in[49], while confidence intervals stem from the DeLong method. R-scripts, data-tables and detailed results are available via https://github.com/HiDiHlabs/preeclamspsia_Nonn_etal/.

BUMPS Cohort – plasma samples at 36 weeks‘ gestation – ELISA Measurement

The biomarker and ultrasound measures for preventable stillbirth (BUMPS) study is a large prospective study conducted at the Mercy Hospital for Women, Melbourne, Australia, which involved the collection of blood at 36 weeks' gestation (35^{+0} – 37^{+0}) preceding diagnosis of term preeclampsia.[50] Preeclampsia was diagnosed according to the ACOG guidelines. Ethical approval was obtained from the Mercy Health and Human Research Ethics Committee (Approval number: 2019–012) and participants gave informed, written consent. A case cohort was selected from the first 1000 patients

enrolled in the BUMPS study. This included all patients who later developed preeclampsia (n=23), and a random selection of 199 patients from the rest of the population who did not develop preeclampsia. Clinical characteristics are provided in Supplemental Table S1. Whole blood was collected in 9mL EDTA tubes. Tubes were centrifuged and plasma supernatant was obtained and stored at -80°C until sample analysis.

Maternal plasma GDF15 levels were measured with a commercial electrochemiluminescence immunoassay platform (Roche Diagnostics). PAI1 was measured in maternal plasma using the human SerpinE1 ELISA from RnD Systems, with plasma diluted 1:300. The AUROC was calculated for each variable and for a multivariable logistic regression model obtained by forward selection of variables. Adjusted odds ratios were calculated using multivariable logistic regression models including BMI and nulliparity as potential confounders.

Immunofluorescence staining (Tissue)

Formalin fixed paraffin embedded (FFPE) placenta tissue sections (5 µm) were mounted on Superfrost Plus slides. Standard deparaffinisation was followed by antigen retrieval (AGR) in the multifunctional microwave tissue processor KOS in Tris-EDTA buffer pH 9.0 or citrate buffer pH 6.0 for 40 min at 93°C. Thereafter, sections were washed with PBS/T and incubated with Ultra V Block for 7 min at RT. For double staining, primary antibodies were mixed and diluted in antibody diluent and incubated on sections for 30 min at RT. Subsequently, slides were washed with PBS/T and incubated with secondary anti-mouse or anti-rabbit antibodies for 30 min at RT. Finally, slides were washed and nuclei stained with DAPI (1: 2,000; Invitrogen). Rabbit immunoglobulin fraction and negative control mouse IgG1 were used as described above and revealed no staining. Tissue sections were mounted with ProLong Gold antifade reagent (Invitrogen) and fluorescence micrographs were acquired with an Olympus microscope (BX3-CBH).

Target	Clone	Species	Dilution	Company
GDF15		rabbit	1:1000 (1st Trim) 1:250 (Term)	Sigma-Aldrich, #HPA011191
anti-rabbit IgG (Alexa Fluor 555)		goat	1:200	Invitrogen
anti-mouse IgG (Alexa Fluor 488)		goat	1:200	Invitrogen

914
915

916 **Immunohistochemistry**

917 Formalin fixed paraffin embedded (FFPE) placenta tissue sections were deparaffinised
918 according to standard procedures. Antigen retrieval (AGR) was performed in a
919 microwave oven in citrate buffer pH6 for 40 min. After a washing step with TBS/T
920 sections were incubated with Hydrogen Peroxide Block (Epredia, Netherlands) to
921 quench endogenous peroxidase followed by a further blocking step with UltraVision
922 Protein Block (Epredia). Primary antibodies were diluted in antibody diluent and
923 incubated on the sections for 45 min at RT. Slides were washed with TBS/T and
924 thereafter the UltraVision LP HRP Polymer Detection System (Epredia) was used
925 according to the manufacturer's instructions. The polymer complex was visualized with
926 AEC (AEC substrate kit, Abcam, UK), sections were counterstained with hemalaun
927 and mounted with Kaisers glycerin gelatine (Merck, Germany). An Olympus VS200
928 slide scanner was used to scan the slides.
929

Target	Clone	Species	Dilution	Company (cat.#)
GDF15	-	rabbit	1:1000 (1st Trim) 1:250 (Term)	Sigma-Aldrich, #HPA011191
HLA-G	4H84	mouse	1:6000	BD Biosciences, #557577
βhCG	-	rabbit	1:1000	Thermo Scientific, #RB-059-A
PAI-1	PR17272-21	rabbit	1:1000	Abcam, ab182973

930
931

Data and code availability

The snRNA-Seq raw data of the 33 villous and decidual samples generated in this study have been deposited in the European Genome-Phenome Archive under the accession number EGAS00001005681. The data are available under controlled access due to the sensitive nature of sequencing data, and access can be obtained by contacting the appropriate Data Access Committee listed for each dataset in the study. Access will be granted to commercial and non-commercial parties according to patient consent forms and data transfer agreements. Images of the ISS data are available via Zenodo (doi: 10.5281/zenodo.5243240). The Visium data are available via Zenodo (doi: 10.5281/zenodo.5336504). The remaining data are available within the article and Supplemental Information. The R script for statistical models of the cohorts is available. Scripts used to analyse the data and generate figures are available via https://github.com/HiDiHlabs/preeclamspsia_Nonn_etal/.

Data collection

No software was used for data collection.

Data analysis

Single-nucleus RNA sequencing analysis. The alignment and pre-processing of the snRNA-Seq data were performed using Cellranger version 3.0.2, 6.0.1 & 6.1.2. Ambient RNA and background noise correction were performed using CellBender 0.2.0. The data were processed using scanpy 1.8.2 in python 3.7.9. scvi-tools 0.14.5 was used for data harmonization. UMAP was computed using umap-learn 0.5.2. Trajectory analysis was performed using stream 1.1 and scanpy 1.8.2. Seurat 4.0 was used for marker analysis. Cell-cell interaction analyses were performed using Connectome 1.0.1 and LIANA 0.1.4. Gene/transcription factor regulatory network analyses and visualization were performed using STRING, iRegulon, and Cytoscape 3.8.2. For visualisation, igraph 1.3.2, circlize 0.4.15, dplyr 1.0.9, ComplexHeatmap 2.10.0; seaborn 0.10.0, and python-igraph 0.7.1 were used. Generally, scikit-learn 1.0.2, statsmodel 0.12.1, scipy 1.5.3, pandas 1.1.4, and numpy 1.19.4 were used.
ISS analysis was performed using python 3.10.4 and jupyter 1.0.0. Data handling was done using plankton 0.1.0, which uses pandas 1.4.3. All plots were generated using matplotlib 3.5.2. SnRNA-Seq data was integrated using scanpy 1.9.1. Image analysis for villi wall detection was performed with Scikit-image 0.19.2. Scikit-learn 1.1.1 was

used to assign wall pixels and for spatial model building and nearest neighbor analysis. Numpy 1.22.4 was used for all algebraic operations on matrix representations of the data. Scipy 1.8.1 was used for statistical model building during pathway analysis. For conditional regression model analyses and visualisations, R 4.1.2, magrittr 2.0.2, Matching 4.9-11, tidyr 1.2.0, survival 3.2-13 and pROC 1.18.0 were used.

Two-group baseline characteristic and experimental comparisons (trm.ctrl vs eoPE; ptrm vs eoPE; FGR vs noFGR) were statistically analyzed and visualized using the GraphPad Prism software (v9.4.1). Values were evaluated for normality using the Shapiro-Wilk, Kolmogorov-Smirnov and Anderson-Darling tests. Homoscedasticity plots were visually inspected per group to assess variance.Significance was tested between groups using either a two-tailed unpaired t-test with Welch's correction or a two- tailed Mann-Whitney U test. which do not assume equal variances.

References Methods

39. Harsem NK, Staff AC, He L, Roald B. The decidual suction method: a new way of collecting decidual tissue for functional and morphological studies. *Acta Obstet Gynecol Scand*. 2004;83:724-730. doi: 10.1111/j.0001-6349.2004.00395.x

40. Krishnaswami SR, Grindberg RV, Novotny M, Venepally P, Lacar B, Bhutani K, Linker SB, Pham S, Erwin JA, Miller JA, et al. Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nat Protoc*. 2016;11:499-524. doi: 10.1038/nprot.2016.015

41. Slyper M, Porter CBM, Ashenberg O, Waldman J, Drokhlyansky E, Wakiro I, Smillie C, Smith-Rosario G, Wu J, Dionne D, et al. A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors. *Nat Med*. 2020;26:792-802. doi: 10.1038/s41591-020-0844-1

42. Suryawanshi H, Morozov P, Straus A, Sahasrabudhe N, Max KEA, Garzia A, Kustagi M, Tuschl T, Williams Z. A single-cell survey of the human first-trimester placenta and decidua. *Sci Adv*. 2018;4:eaau4788. doi: 10.1126/sciadv.aau4788

43. Luecken MD, Buttner M, Chaichoompu K, Danese A, Interlandi M, Mueller MF, Strobl DC, Zappia L, Dugas M, Colome-Tatche M, Theis FJ. Benchmarking atlas-level data integration in single-cell genomics. *Nat Methods*. 2022;19:41-50. doi: 10.1038/s41592-021-01336-8

44. Boyeau P, Lopez R, Regier J, Gayoso A, Jordan MI, Yosef N. Deep Generative Models for Detecting Differential Expression in Single Cells. *bioRxiv*. 2019:794289. doi: 10.1101/794289

45. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43:e47. doi: 10.1093/nar/gkv007

46. Dimitrov D, Turei D, Garrido-Rodriguez M, Burmedi PL, Nagai JS, Boys C, Ramirez Flores RO, Kim H, Szalai B, Costa IG, et al. Comparison of methods and resources for cell-cell communication inference from single-cell RNA-Seq data. *Nat Commun*. 2022;13:3224. doi: 10.1038/s41467-022-30755-0

47. Guy GP, Leslie K, Diaz Gomez D, Forenc K, Buck E, Khalil A, Thilaganathan B. Implementation of routine first trimester combined screening for pre-eclampsia: a clinical effectiveness study. *BJOG*. 2021;128:149-156. doi: 10.1111/1471-0528.16361

48. O'Gorman N, Wright D, Poon LC, Rolnik DL, Syngelaki A, de Alvarado M, Carbone IF, Dutemeyer V, Fiolna M, Frick A, et al. Multicenter screening for pre-eclampsia by maternal factors and biomarkers at 11-13 weeks' gestation: comparison with NICE guidelines and ACOG recommendations. *Ultrasound Obstet Gynecol*. 2017;49:756-760. doi: 10.1002/uog.17455

49. Xu H, Qian J, Paynter NP, Zhang X, Whitcomb BW, Tworoger SS, Rexrode KM, Hankinson SE, Balasubramanian R. Estimating the receiver operating characteristic curve in matched case control studies. *Stat Med*. 2019;38:437-451. doi: 10.1002/sim.7986

50. Bartho LA, Kandel M, Walker SP, Cluver CA, Hastie R, Bergman L, Pritchard N, Cannon P, Nguyen TV, Wong GP, et al. Circulating Chemerin Is Elevated in Women With Preeclampsia. *Endocrinology*. 2023;164. doi: 10.1210/endocr/bqad041

Supplemental tables

Supplemental Table S1: Detailed snRNAseq and validation pregnancy cohorts characteristics.

Extensive description of maternal characteristics from the early and late pregnancy tissue samples of control and early onset pre-eclampsia used for snRNAseq separated by sampleID. Additional tabs contain the summary of baseline characteristics between groups for our multi-centre and longitudinal cohorts.

Sample ID	gestational age (days)	crown-rump-length (mm)	maternal age	maternal height (cm)	maternal weight (kg)	BMI	Rh status
17/021	63	29	40	163	58	21.8	pos
17/022	49	8	23	162	48	18.3	pos
17/025	54	8	31	167	67	24.0	neg
18/082	74	48	28	165	62	22.8	pos
18/108	76	39	25	169	60	21.0	pos
18/032	60	18	20	161	45	17.4	pos
18/017	58	19	35	176	78	25.2	neg
18/033	64	26	32	169	60	21.0	pos
18/096	35	3	28	175	56	18.3	pos
20/027	43	5	32	167	55	19.7	pos
mean	*57.6*	*20.3*	*29.4*	*167.4*	*58.9*	*20.95*	

HYPE_2024_23362_Supplementary Table 01_all cohorts

Sample ID	Tissue	Gestational time	Diagnosis	Fetal sex	l age (days)	Maternal BMI pre-pregnancy	Maternal age (y)	SBP delivery (mmg/Hg)	DBP delivery	protein (stick)	Placental weight (g)	Fetal weight (g)	Smoking status	Maternal comorbidities	FGR/SGA*	Tissue of origin
TRM1-v	Villi	late	C	male	272	25.7	38	122	74	0	840	4300	Non-smoker	Factor 2 deficiency	no FGR according to consensus definition by DRT, 20190114	fetal
TRM2-v	Villi	late	C	female	267	28.9	32	109	87	0	580	3430	Non-smoker	None	no FGR according to consensus definition by DRT, 20190114	fetal
TRM3-v	Villi	late	C	male	270	20.2	33	110	68	0	980	4044	Non-smoker	None	no FGR according to consensus definition by DRT, 20190114	fetal
327-v	Villi	late	C	male	275	22.7	35	137	69	0	928	3630	Non-smoker	None	no FGR according to consensus definition by DRT, 20190114	fetal
328-v	Villi	late	C	male	278	25.8	38	109	65	0	595	3332	Non-smoker	Known hyperlipidemia	no FGR according to consensus definition by DRT, 20190114	fetal
372-v	Villi	late	C	male	283	23.1	37	136	89	0	580	3540	Non-smoker	Asthma	no FGR according to consensus definition by DRT, 20190114	fetal
327-d	Decidua	late	C	male	275	22.7	35	137	69	0	928	3630	Non-smoker	None	no FGR according to consensus definition by DRT, 20190114	maternal
328-d	Decidua	late	C	male	278	25.8	38	109	65	0	595	3332	Non-smoker	Known hyperlipidemia	no FGR according to consensus definition by DRT, 20190114	maternal
372-d	Decidua	late	C	male	283	23.1	37	136	89	0	580	3540	Non-smoker	Asthma	no FGR according to consensus definition by DRT, 20190114	maternal
073-d	Decidua	late	C	female	277	25.0	25	118	85	0	540	3414	NA	NA	NA	maternal
389-v	Villi	late	eoPE	male	192	23.7	26	152	87	3	220	753	Non-smoker	None	fetal doppler and/or baby<3percentile	fetal
419-v	Villi	late	eoPE	male	236	21.9	33	147	87	3	280	1254	Non-smoker	None	fetal doppler and/or baby<3percentile	fetal
181-v	Villi	late	eoPE	male	206	31.9	33	206	110	3	500	1059	NA	NA	NA	fetal
120-v	Villi	late	eoPE	male	196	29.0	31	150	90	3	240	897	NA	NA	NA	fetal
577-1-v	Villi	late	eoPE	male	222	26.1	35	149	94	3	380	1420	Non-smoker	Allergies	fetal doppler and/or baby<3percentile	fetal
577-2-v	Villi	late	eoPE	male	222	26.1	35	149	94	3	380	1420	Non-smoker	Allergies	fetal doppler and/or baby<3percentile	fetal
389-d	Decidua	late	eoPE	male	192	23.7	26	152	87	3	220	753	Non-smoker	None	fetal doppler and/or baby<3percentile	maternal
419-d	Decidua	late	eoPE	male	236	21.9	33	147	87	3	280	1254	Non-smoker	None	fetal doppler and/or baby<3percentile	maternal
274-d	Decidua	late	eoPE	female	227	23.2	43	180	97	2	400	1590	Non-smoker	None	fetal doppler and/or baby<3percentile	maternal
100-d	Decidua	late	eoPE	male	184	20.9	32	150	90	1	125	648	NA	None	fetal doppler and/or baby<3percentile	maternal
102-d	Decidua	late	eoPE	male	192	27.8	30	180	110	3	160	1215	NA	None	fetal doppler and/or baby<10percentile	maternal

v, villi; d, decidua; C, control; eoPE, early onset preeclampsia; FGR, fetal growth restriction; SGA, small for gestational age; NA, not available.

HYPE_2024_23362_Supplementary Table 01_all cohorts

	term no labour	term in labour	preterm in labour
Clinical parameters			
Maternal age (years; median [IQR])	32 (28-35)	25 (22-31.5)	27 (23.5-29)
Body mass index (kg/m^2; median [IQR])	31.2 (30.3-37.7)	43.3 (36.9-44.1)	32.5 (28.9-39.9)
Primiparity	0% (0/3)	66.7% (2/3)	33.3% (1/3)
Cesarean section	100% (3/3)	66.7% (2/3)	33.3% (1/3)
Gestational age at delivery (weeks; median [IQR])	39.6 (39.3-39.6)	39.1 (38.8-39.9)	35.1 (33.4-35.2)
Birthweight (g)	4030 (3900-4160)[a]	3310 (3160-3405)	2145 (1680-2220.5)
Ethnicity			
African-American	66.7% (2/3)	66.7% (2/3)	100% (3/3)
Caucasian	33.3% (1/3)	0% (0/3)	0% (0/3)
Other	0% (0/3)	33.3% (1/3)	0% (0/3)

Adapted from
doi.org/10.7554/eLife.52004

HYPE_2024_23362_Supplementary Table 01_all cohorts

Characteristic	pre-term control (n=11)	eoPE (n=23)	*p* between groups
Foetal sex	45.5% female	43.5% female	-
Gestational age (days)	214.7 ± 22.3	209.8 ± 15.8	0.5224
Maternal BMI pre-pregnancy	25.6 ± 7.9	25.3 ± 5.3	0.914
Maternal age	31.8 ± 5.8	30.4 ± 5.2	0.497
SBP delivery	113.3 ± 17.7	163.7 ± 21.9	**<0.0001**
DBP delivery	71.5 ± 10.5	98.4 ± 13.7	**<0.0001**
Urinary protein (stick) *	0.13 ± 0.4	2.47 ± 1.1	**<0.0001**
Gravidity	2.64 ± 1.6	2.22 ± 1.3	0.453
Parity	1.18 ± 0.6	1.00 ± 0.9	0.481
Foetal weight (g)	1686 ± 524.5	1244 ± 375.4	**0.024**

Mean ± standard deviation shown, significance tested via unpaired t-test with Welch's
* n = 8 early control, n = 19 eoPE

HYPE_2024_23362_Supplementary Table 01_all cohorts

preterm correction

	preterm (<34 weeks) no FGR		term (>37 weeks) no FGR		
	n = 10		n = 16		
	mean	sd	mean	sd	p-value
Maternal age (y)	33.8	5.3	33.9	4.1	0.945
BMI (kg/m2)	23.5	3.9	25.5	5.2	0.270
Gestational age (days)	30.5	2.4	38.8	0.6	**0.000001**
Maximum systolic BP	120.9	9.6	121.1	6.7	0.963
Maximum diastolic BP	78.5	9.1	77.2	7.5	0.706
Fetal sex	F: 5, M: 5	-	M: 9, F: 7	-	-
Previous nulliparity	No: 6, Yes: 4	-	No: 12, Yes: 4	-	-

FGR correction

	preterm (<34 weeks) FGR		preterm (<34 weeks) no FGR		
	n = 8		n = 15		
	mean	sd	mean	sd	p-value
age	34.0	3.0	30.7	6.4	0.111
bmi	29.0	5.7	26.6	4.3	0.321
ga	30.5	2.3	29.7	1.6	0.430
max_sbp	166.6	21.1	174.0	14.5	0.396
max_dbp	107.9	14.0	106.8	8.3	0.846
gender	F: 5, M: 3	-	F: 9, M: 6	-	-
prev_nulliparity	Yes: 6, No: 2	-	Yes: 11, No: 4	-	-

adapted from
DOI: 10.1161/HYPERTENSIONAHA.116.07293

HYPE_2024_23362_Supplementary Table 01_all cohorts

Characteristic	**healthy controls**, N = 56 [1]	**eoPE cases**, N = 28 [1]	**p-value** [2]
GA at first scan	88.0 (86.0, 92.0)	88.0 (85.0, 92.0)	>0.9
Maternal age	34.0 (29.0, 37.3)	34.0 (29.0, 38.0)	0.9
BMI	27 (23, 31)	27 (23, 33)	0.9
Aspirin	0 (0%)	15 (54%)	<0.001
GA Birth (d)	279 (274, 284)	250 (237, 253)	<0.001
Birth weight (g)	3,525 (3,063, 3,853)	1,883 (1,510, 2,298)	<0.001
UtA Mean PI (MoM)	0.94 (0.78, 1.07)	1.10 (0.82, 1.29)	0.045

[1]Median (IQR); n (%)

[2]Wilcoxon rank sum test; Pearson's Chi-squared test

GA, gestational age; UtA-PI, uterine artery pulsatility index; eoPE, early onset preeclampsia.

HYPE_2024_23362_Supplementary Table 01_all cohorts

	Controls (n=201)	**Preeclampsia/FGR** (n=21)	**P value**
Maternal Age Median (IQR)	32 (30-35)	34 (30.5-37)	0.14
Gestation at Sampling (days) Median (IQR)	253 (250.5-256)	254 (251-256.5)	0.3
BMI (kg/m^2) † median (IQR)	24.8 (22.4-28.2)	27.9 (24.1-30.9)	0.03
Highest Systolic BP (mmHg) Median (IQR)	135 (125-140)	152 (150-172)	<0.0001
Highest Diastolic BP (mmHg) Median (IQR)	80 (75-100)	100 (96.5-105)	<0.0001
Parity no. (%)			
0	97 (48.3)	19 (90.5)	
1	80 (39.8)	2 (9.5)	0.0011
≥2	24 (11.9)	0 (0)	
Gravidity no. (%)			
Primiparous	72 (35.8)	11 (52.4)	ns
Multiparous	129 (64.2)	10 (47.6)	
Birth weight (g) Mean (SEM)	3458 (31.2)	3078 (110.7)	0.0003
Fetal sex			
Male	99 (49.3)	10 (47.6)	ns
Female	102 (50.7)	11 (52.4)	

BP, blood pressure; FGR, fetal growth restriction

Patient clinical characteristics. BUMPS cohort (Melbourne, Australia) from whom plasma samples were collected around 36 weeks' gestation preceding diagnosis of preeclampsia. All comparisons were made versus controls that were representative of the population (randomly sampled cohort).

† Data on BMI (body mass index at the first pregnancy visit) missing for 2/199 controls

HYPE_2024_23362_Supplementary Table 01_all cohorts

Supplemental Table S2: Preterm genes per cell annotated cell type or state.
Preterm gene sets identified by comparing preterm vs term controls for every conserved cell type/state after snRNA/scRNA-seq data harmonization as described in Methods. scVI (change mode) algorithm was used for the analysis. Metrics such as proba_de (probability of a gene to be differential), bayes Factor, LFC (log-fold change), cell-type proportions (non-zeros proportion), and raw normalized mean counts are reported for both group1 (preterm) and group2 (term) are reported. As per scVI convention, mean, median and standard deviation (std) is reported for the LFC effect size variable. Both placenta (first tab) and decidua (second tab) are tabulated. Differentially expressed preterm genes identified by comparing preterm and term controls in the bulk microarray data from Leavey et. al 2015[21]. Limma-R package was used for the analysis, and only significant genes with BH-corrected adjusted p-value < 0.01 and also tested for sex-bias are reported (third tab). Limma test performed to identify genes differentially expressed between eoPE with FGR vs. no-FGR condition using bulk microarray data. Insignificant adjusted p-values (> 0.05) reflect that no genes were differentially regulated (fourth tab).

Supplemental Table S3: Cell type or state composition per biological sample sequenced.
Nuclei number contribution per condition, i.e., early, term late controls and eoPE pregnancies (first tab) and per biological replicate annotated to a specific cell type or state in the maternal-fetal interface. Contributions are tabulated as absolute and relative values and separated by gestational age timepoint and condition. Total number of cells per donor, condition, and cell type/state are reported. Second and third tabs separate tissue of origin.

Supplemental Table S4: Quality Control metrics for Placenta villi.
Comprehensive table tabulating placenta QC at the level of CellRanger (mean reads per cell, reads mapped confidently to genome and fraction reads in cells); Cellbender filtering (after removing ambient RNA and random barcode swapping); Scanpy filtering and final downstream filtering. Estimated number of cells, median number of UMI, and median number of genes per sample during QC from each step are reported at the donor level (first tab) and cell level (second tab), respectively. Total number of UMI counts (total_counts), number of genes with at least one count (n_genes_by_counts), %MT-transcripts per nuclei (pct_counts_MT_genes).

Supplemental Table S5: Quality Control metrics for Decidua.
Comprehensive table tabulating decidua QC at the level of CellRanger (mean reads per cell, reads mapped confidently to genome and fraction reads in cells); Cellbender filtering (after removing ambient RNA and random barcode swapping); Scanpy filtering and final downstream filtering. Estimated number of cells, median number of UMI, and median number of genes per sample during QC from each step are reported at the donor level (first tab) and cell level (second tab), respectively. Total number of UMI counts (total_counts), number of genes with at least one count (n_genes_by_counts), %MT-transcripts per nuclei (pct_counts_MT_genes).

Supplemental Table S6: Quality Control metrics for each cell type/state of placenta and decidua.
Total number of UMI counts (total_counts), number of genes with at least one count per nuclei (n_genes_by_counts), %MT-transcripts per nuclei (pct_counts_MT_genes) are reported for each cell type/state.

Supplemental Table S7: Investigation of batch-effects using technical parameters.
Tabulated cell-type/state absolute silhouette width (ASW), adjusted rand index (ARI) and adjusted mutual information (AMI) score per batch in case of technical replicates (557_1 and 557_2 samples); late term controls and eoPE pregnancies (first tab); early pregnancy (second tab) and comparing our study metrics to published datasets (third tab). ASW are scaled from 0-1; where 1 means perfect integration or batch removal. ARI/AMI scores close to zero indicate that the batch and cluster labels are independent of each other.

Supplemental Table S8: Markers identifying annotated cell types and states.
Summary of significant differential genes derived from logistic regression analysis (after covariate correction) that characterise each annotated cell type or state in the manuscript. Lists are tabulated per tissue of origin (placenta, decidua). The list is restricted to Bonferroni adjusted $p < 0.05$ and log2 fold-change ± 0.25. Metascape analysis showing comparative pathway enrichment analysis of STB1 and STB2 cell states (third tab) based on top-100 Logistic Regression markers. Significant vSTBjuv markers relative to vSTB1/2 were identified using Logistic Regression. List is restricted to Bonferroni adjusted $p < 0.01$ and log2 fold-change ± 0.25; top-100 markers are reported. The number of vSTB1/2 was downsampled to that of vSTBjuv to avoid composition bias (fourth tab). Significant STB1 vs STB2 markers using Logistic Regression; Bonferroni adjusted $p < 0.01$ and log2 fold-change ± 0.25. The number of vSTB1 was downsampled to match with that of vSTB2 to avoid composition bias (fifth tab).

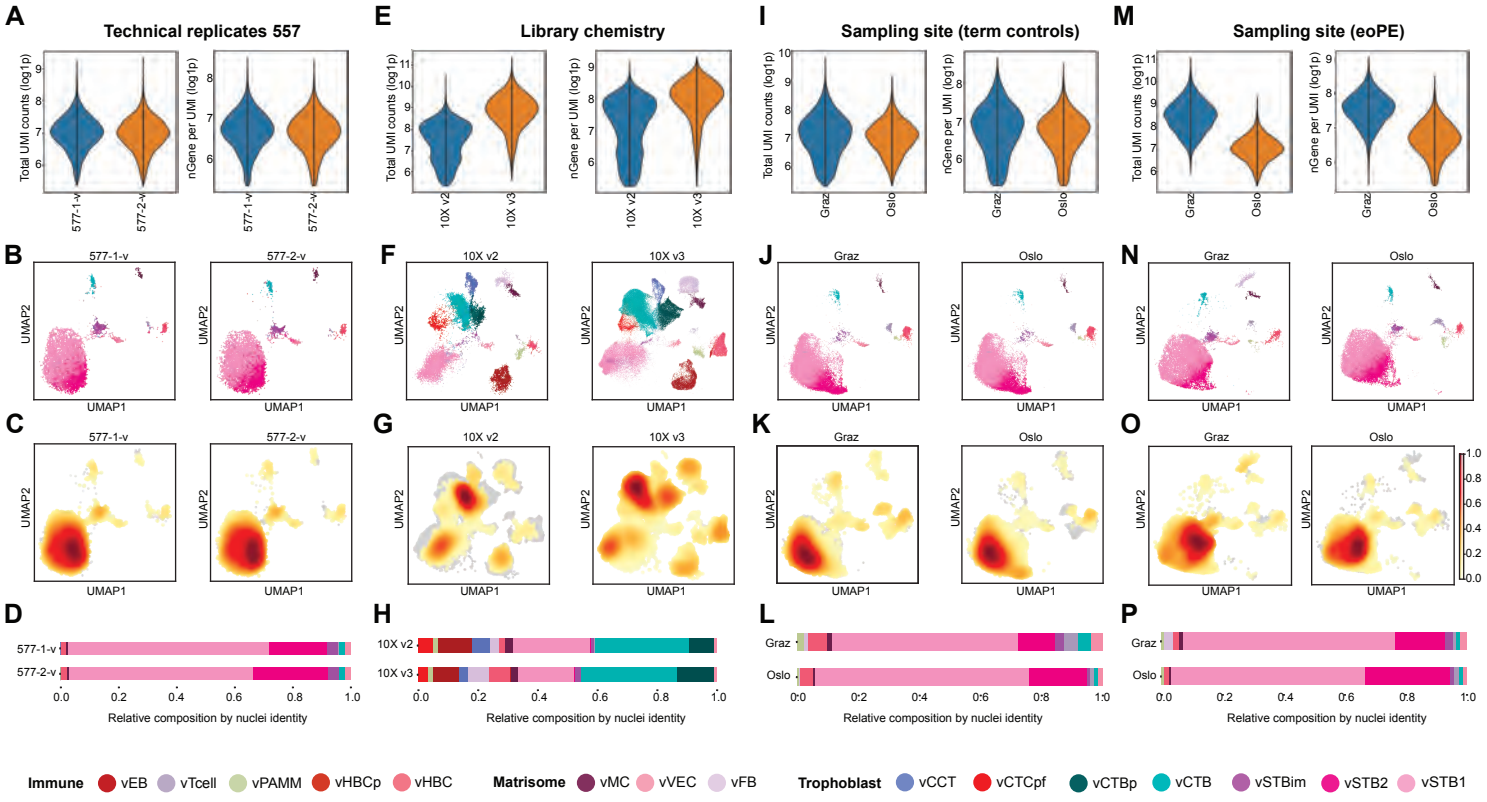
Supplemental Table S9: eoPE markers per cell type/state.
Differentially expressed genes between pre-eclampsia and late pregnancy controls using multivariate logistic regression. p-values were Bonferroni adjusted for: number of UMIs, gene counts, %mitochondrial transcripts, %sex transcripts and preterm gene scores per nucleus. List is restricted to Bonferroni adjusted p-value < 0.05 and log2 fold-change ± 0.25. Tabs separate the cell type/state.

Supplemental Table S10: Differentially expressed genes overlapping in villi compartments.
Significantly dysregulated genes in early onset pre-eclampsia that overlap between Immune, Meso/endothelial and Trophoblast villi compartments as defined by Figure 2E and described in Supplemental Figure S4.

Supplemental Table S11: Receptor ligand interaction analyses.
Receptor ligand differentially upregulated interactions of secreted vSTB and highly expressed dDSTB ligands in eoPE binding to maternal dVEC and dSMC. dEVT ligands interacting with decidua; vCCT ligands interacting with decidua matrisome; up/downregulated interactions within villi and within decidua. Lists informed visualisations in Figure 3C, Figure 6A; Supplemental Figure S8.

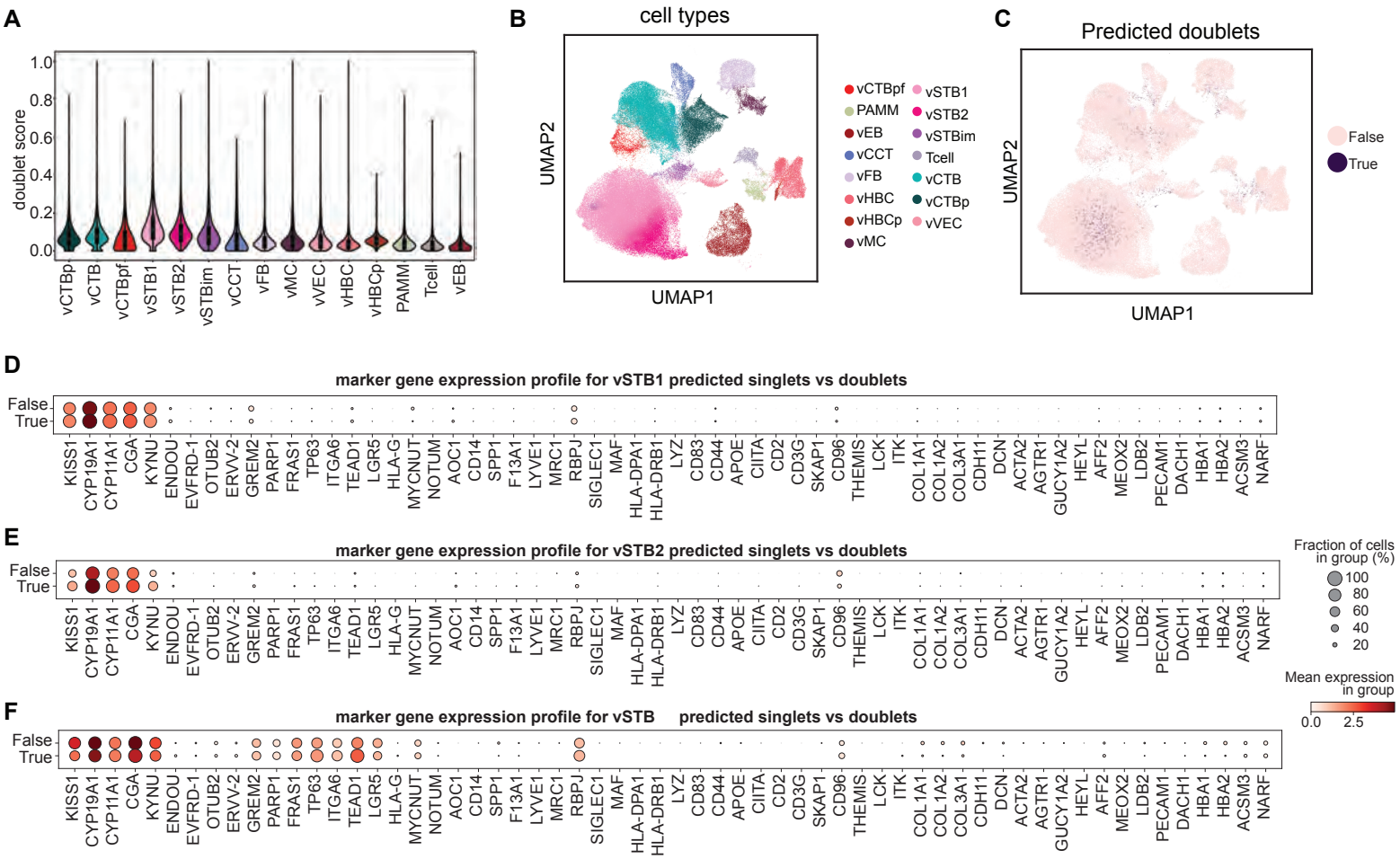
Supplemental Table S12: Pathway enrichment analysis of syncytiotrophoblast functional states- STB1 and STB2.
Metascape enrichment analysis of eoPE genes uniquely dysregulated in vSTB1 (first tab); vSTB2 (second tab); vSTBjuv (third tab) and iRegulon identified transcription factors for enriched unique STB targets (fourth tab). eoPE genes are as per Supplemental Table 9.

Supplemental Table S13: Syncytiotrophoblast dysregulated genes associated with a senescence secretory phenotype.
Overlap of dysregulated genes in the syncytiotrophoblast nuclei states. Genes in one or more of these states associated to a senescence phenotype compartment based on comparison to the human senescence associated proteins (SASP) atlas PMID: 31945054.



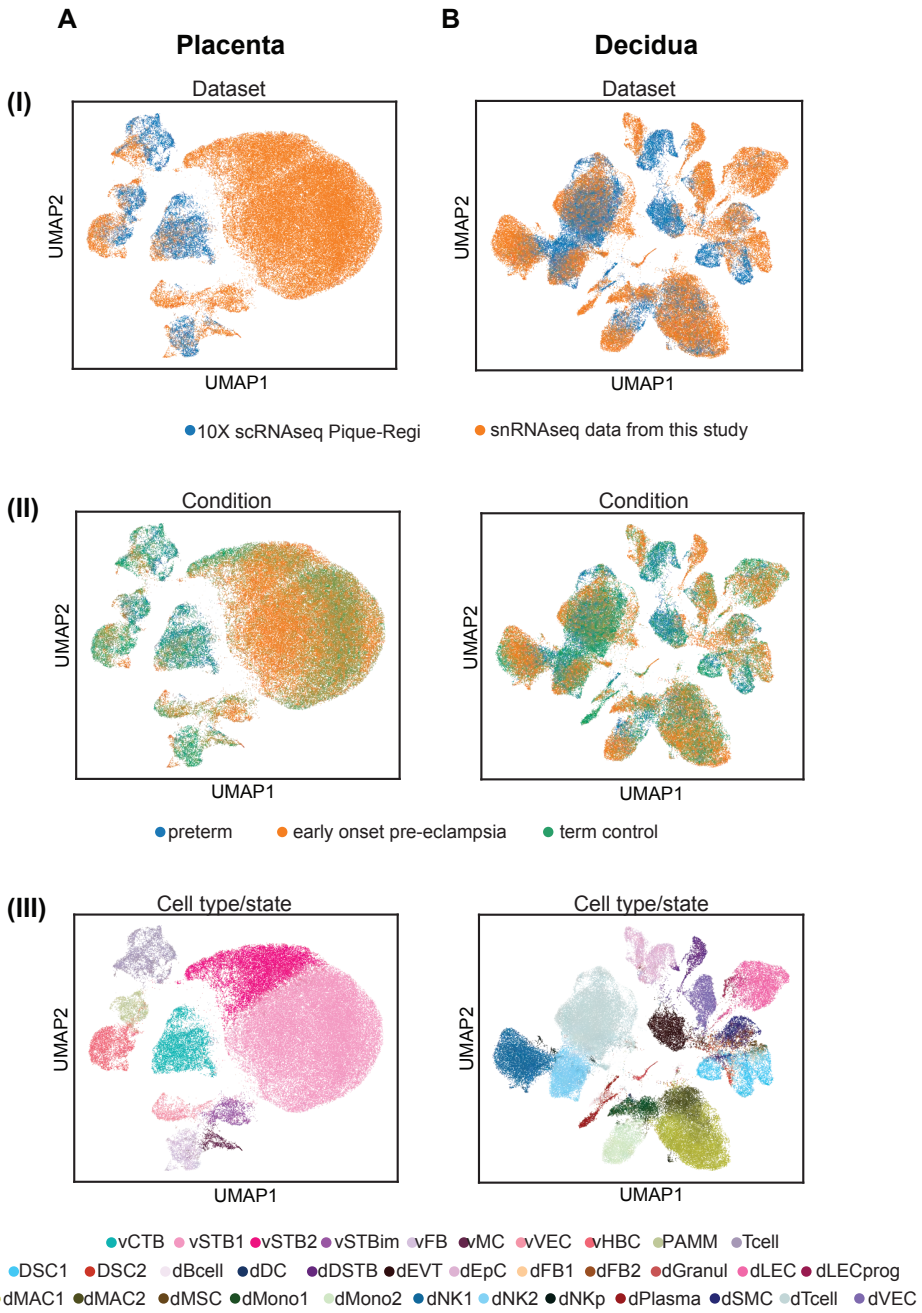
Supplemental Figure S1 | Batch effect assessment from snRNAseq samples in this study

(A) Violin plots illustrating the detected (log1p per sample) total UMI counts representative of total RNA transcripts (left) and number of total genes having at least one positive count in a cell (right) for technical replicates 557_1 and 557_2 (n= 6185 and 6081 nuclei respectively). **(B)** UMAP embeddings split by technical replicates 557_1 and 557_2 visualizing distribution and concordance of cell-types/states. **(C)** Kernel density estimation revealing similar composition in technical replicates 557_1 and 557_2. Values are scaled from 0-1 for comparison and high-density values suggest strong contribution of cells to the overall dataset. **(D)** Stacked bar-plot depicting similar relative composition of cell types/states in technical replicates 557_1 and 557_2. **(E)** Violin plots illustrating the detected (log1p per sample) total UMI counts representative of total RNA transcripts (left) and number of total genes having at least one positive count in a cell (right) for early (first trimester) pregnancy samples split by library condition- 10X V2 (n=3; 12848 nuclei) and 10X V3 chemistry (n=7; 67037 nuclei). **(F)** UMAP embeddings of early pregnancy samples split by 10X V2 and 10X V3 library samples visualizing distribution and concordance of cell-types/states. **(G)** Kernel density estimation reflecting contribution of cells to overall composition in 10X V2 and 10X V3 library samples (early). Values are scaled from 0-1 for comparison and high-density values suggest strong contribution of cells to the overall dataset. **(H)** Stacked bar-plot depicting similar relative composition of cell types/states in 10X V2 and 10X V3 library samples (early) with more vHBCp, vCTBp, vFB, vMC, vHBC, and vPAMM in 10x V3 samples. **(I)** Violin plots illustrating the detected (log1p per sample) total UMI counts representative of total RNA transcripts (left) and number of total genes having at least one positive count in a cell (right) for late term controls split by sampling site- Graz (n=3; 13969 nuclei) and Oslo (n=3, 15733 nuclei). **(J)** UMAP embeddings of term control samples split by sampling site- Graz and Oslo visualizing similar distribution and concordance of cell-types/states in the two categories. **(K)** Kernel density estimation reflecting contribution of cells to overall composition in Graz and Oslo term control samples. **(L)** Stacked bar-plot depicting similar relative composition of cell types/states in Graz and Oslo term controls showing increased proportions of PAMM, vT-cells, vFB and vMC in Graz samples. **(M)** Violin plots illustrating the detected (log1p per sample) total UMI counts representative of total RNA transcripts (left) and number of total genes having at least one positive count in a cell (right) for eoPE samples split by sampling site- Graz (n=2; 12365 nuclei) and Oslo (n=3, 17604 nuclei). **(N)** UMAP embeddings of eoPE samples split by sampling site- Graz and Oslo visualizing similar distribution and concordance of cell-types/states in the two categories. **(O)** Kernel density estimation reflecting contribution of cells to overall composition in Graz and Oslo eoPE samples. **(P)** Stacked bar-plot depicting similar relative composition of cell types/states in Graz and Oslo eoPE samples showing slight depletion in immune cells (PAMM & vTcells) in Graz samples relative to Oslo. CTB, villous cytotrophoblast; STB, syncytiotrophoblast; DSTB, deported STB; CCT, cell column trophoblast; EVT, extravillous trophoblast; VEC, vascular endotheli-al cell; LEC, lymphatic endothelial cell; LECP, LEC progenitor; SMC, smooth muscle cell; MC, myocyte; FB, fibroblast; EpC, epithelial cell; MSC, mesenchymal stem cell; DSC, decidual stromal cell; EB, erythroblast; HBC, Hofbauer cell; PAMM, placenta-associated maternal macrophage; Mono, monocyte; MAC1, M1-like macrophage; MAC2, M2-like macrophage; NK, natural killer cell; PC, plasma cell; DC, dendritic cell, Granul, granulocyte; v, villous; d, decidual; p, proliferative; im, immature; pf, pre-fusion.



Supplemental Figure S2 | Doublet nuclei inference in villous cell types and states

(A) Violin plot depicting doublet score distribution across villi cell types. Scores are ranged from 0 to 1. **(B)** UMAP embeddings reflecting cell type/state distribution. **(C)** Predicted doublets using a threshold of 0.35 from doublet score histogram as predicted by Scrublet. **(D-F)** Dot plots showing expression of key marker genes for predicted doublets against singlets for (D) vSTB1, (E) vSTB2, (F) vSTBim. Absence of contradictory lineage specific genes reveals lack of biological doublets in our cell types or states. CTB, villous cytotrophoblast; STB, syncytiotrophoblast; DSTB, deported STB; CCT, cell column trophoblast; EVT, extravillous trophoblast; VEC, vascular endothelial cell; MC, myocyte; FB, fibroblast; EB, erythroblast; HBC, Hofbauer cell; PAMM, placenta-associated maternal macrophage; v, villous; p, proliferative; im, immature; pf, pre-fusion.

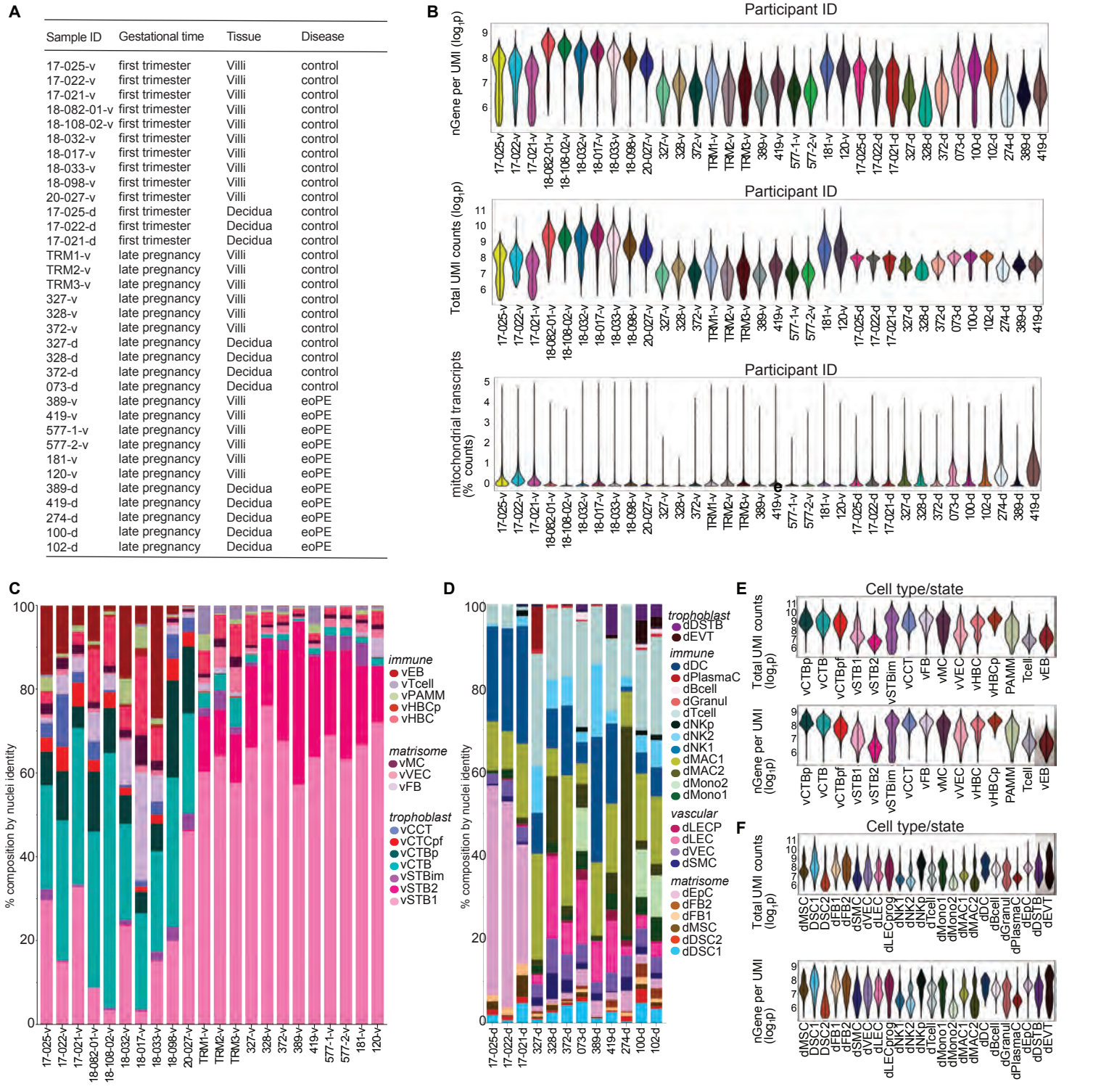


Supplemental Figure S3 | Integration visualisation summary for gestational age correction in differential gene expression inference

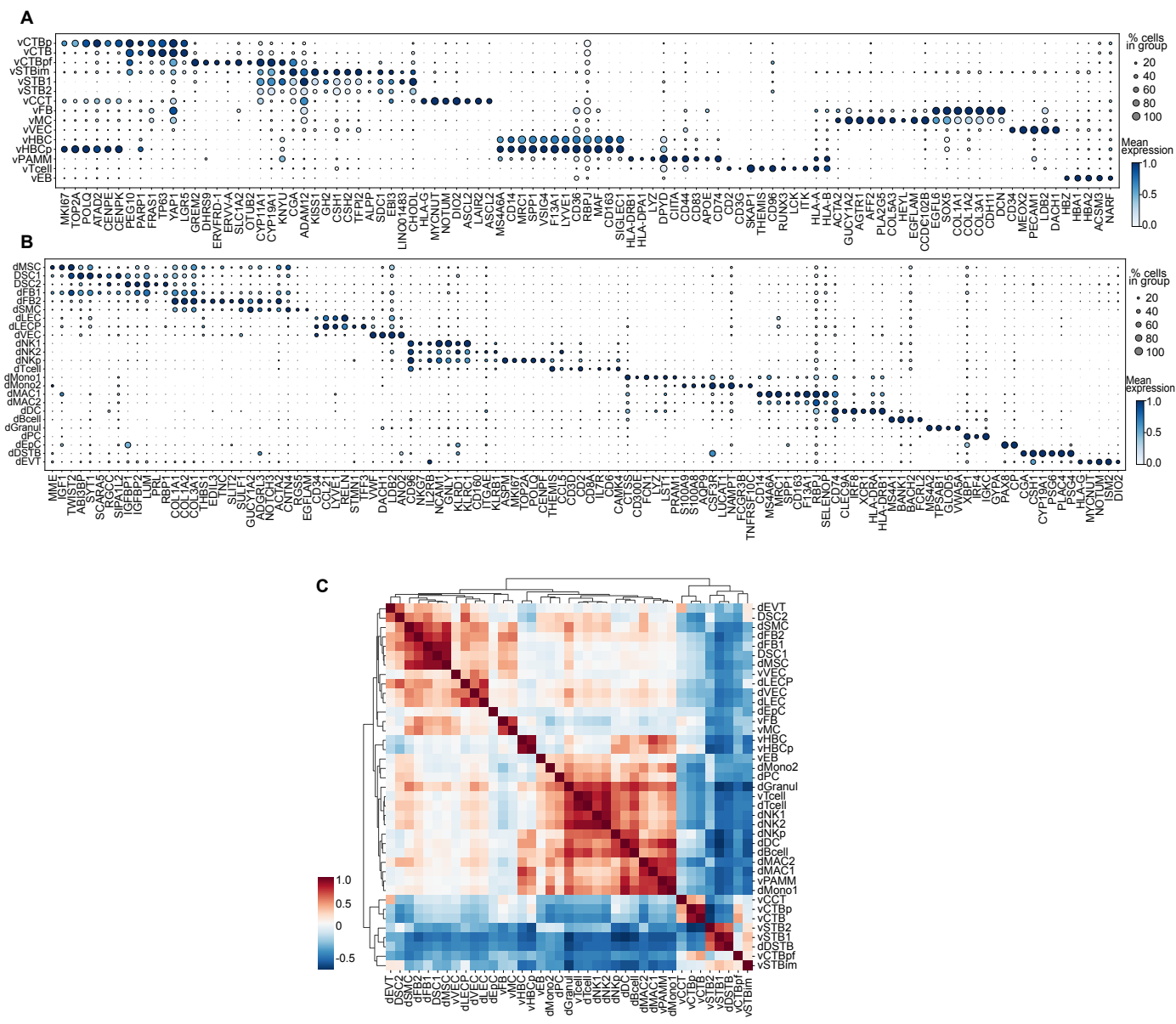
UMAP visualization of the 10X snRNA-seq from this study integrated with 10X scRNA-seq Pique-Regi data [18] of **(A)** placenta villi and **(B)** decidua respectively colored by **(I)** dataset of origin, **(II)** condition and **(III)** cell types/states. Data harmonization and label transfer was performed using scANVI. Since vSTB(s) were not profiled by Pique-Regi et al., downstream analysis were restricted to conserved cell types/states in both placenta and decidua. Batch-effect investigation using statistics are tabulated in Supplementary Table 6.

A

Sample ID	Gestational time	Tissue	Disease
17-025-v	first trimester	Villi	control
17-022-v	first trimester	Villi	control
17-021-v	first trimester	Villi	control
18-082-01-v	first trimester	Villi	control
18-108-02-v	first trimester	Villi	control
18-032-v	first trimester	Villi	control
18-017-v	first trimester	Villi	control
18-033-v	first trimester	Villi	control
18-098-v	first trimester	Villi	control
20-027-v	first trimester	Villi	control
17-025-d	first trimester	Decidua	control
17-022-d	first trimester	Decidua	control
17-021-d	first trimester	Decidua	control
TRM1-v	late pregnancy	Villi	control
TRM2-v	late pregnancy	Villi	control
TRM3-v	late pregnancy	Villi	control
327-v	late pregnancy	Villi	control
328-v	late pregnancy	Villi	control
372-v	late pregnancy	Villi	control
327-d	late pregnancy	Decidua	control
328-d	late pregnancy	Decidua	control
372-d	late pregnancy	Decidua	control
073-d	late pregnancy	Decidua	control
389-v	late pregnancy	Villi	eoPE
419-v	late pregnancy	Villi	eoPE
577-1-v	late pregnancy	Villi	eoPE
577-2-v	late pregnancy	Villi	eoPE
181-v	late pregnancy	Villi	eoPE
120-v	late pregnancy	Villi	eoPE
389-d	late pregnancy	Decidua	eoPE
419-d	late pregnancy	Decidua	eoPE
274-d	late pregnancy	Decidua	eoPE
100-d	late pregnancy	Decidua	eoPE
102-d	late pregnancy	Decidua	eoPE

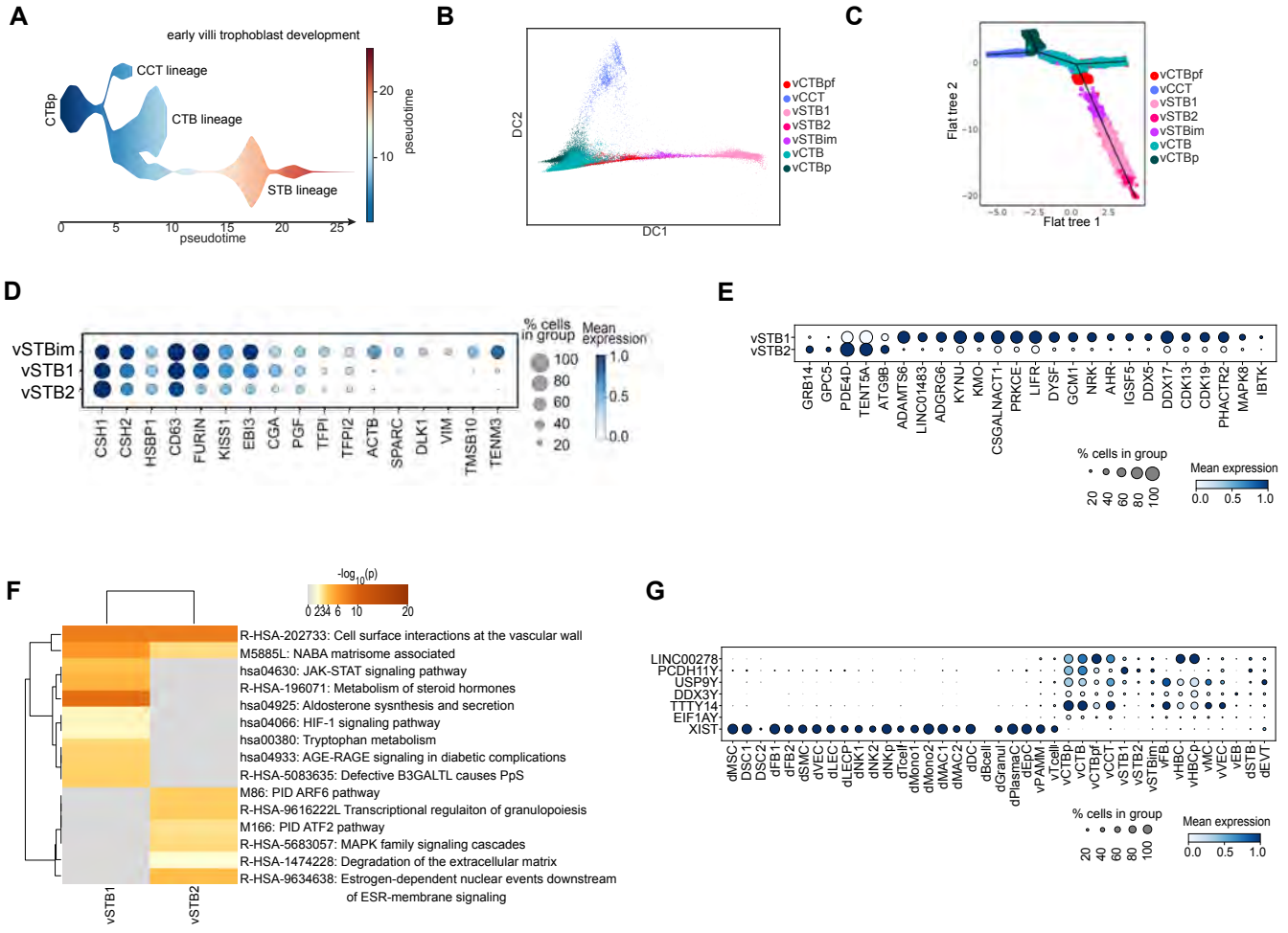


Supplemental Figure S4 | Quality control and characteristics of the single nuclei samples used for analyses
(A) Pseudonymised patient ID of maternal(decidua)-foetal(villi) interface sequenced samples with information regarding sample gestational age and preeclampsia status. **(B)** Violin plots illustrating the detected (log1p per sample) number of total genes having at least one positive count (UMI) in a cell; total UMI counts representative of total RNA transcripts; and percentage mitochondrial transcript per nuclei. Note that 577-1-v and 557-2-v are technical replicates. **(C)** Annotated cell type or state composition (%) per individual biological sample in villi illustrated as stacked bar plots. n = 21 villi (10 early, 6 late control, 5 eoPE) **(D)** Annotated cell type or state composition (%) per individual biological sample in decidua illustrated as stacked bar plots. n = 12 deciduas (3 early, 4 late control, 5 eoPE). **(E)** Violin plots illustrating the detected (log1p) and number of total genes having at least one positive count (UMI) per nuclei in villi annotated cell types and states. **(F)** Violin plots illustrating the detected (log1p) and number of total genes having at least one positive count (UMI) per nuclei in villi annotated cell types and states. **(G)** Box plots comparing composition per cell type or state compared between early and healthy term control (trm.ctrl) pregnancy samples. CTB, villous cytotrophoblast; STB, syncytiotrophoblast; DSTB, deported STB; CCT, cell column trophoblast; EVT, extravillous trophoblast; VEC, vascular endotheli-al cell; LEC, lymphatic endothelial cell; LECP, LEC progenitor; SMC, smooth muscle cell; MC, myocyte; FB, fibroblast; EpC, epithelial cell; MSC, mesenchymal stem cell; DSC, decidual stromal cell; EB, erythroblast; HBC, Hofbauer cell; PAMM, placenta-associated maternal macrophage; Mono, monocyte; MAC1, M1-like macrophage; MAC2, M2-like macrophage; NK, natural killer cell; PC, plasma cell; DC, dendritic cell, Granul, granulocyte; v, villous; d, decidual; p, proliferative;im, immature; pf, pre-fusion.



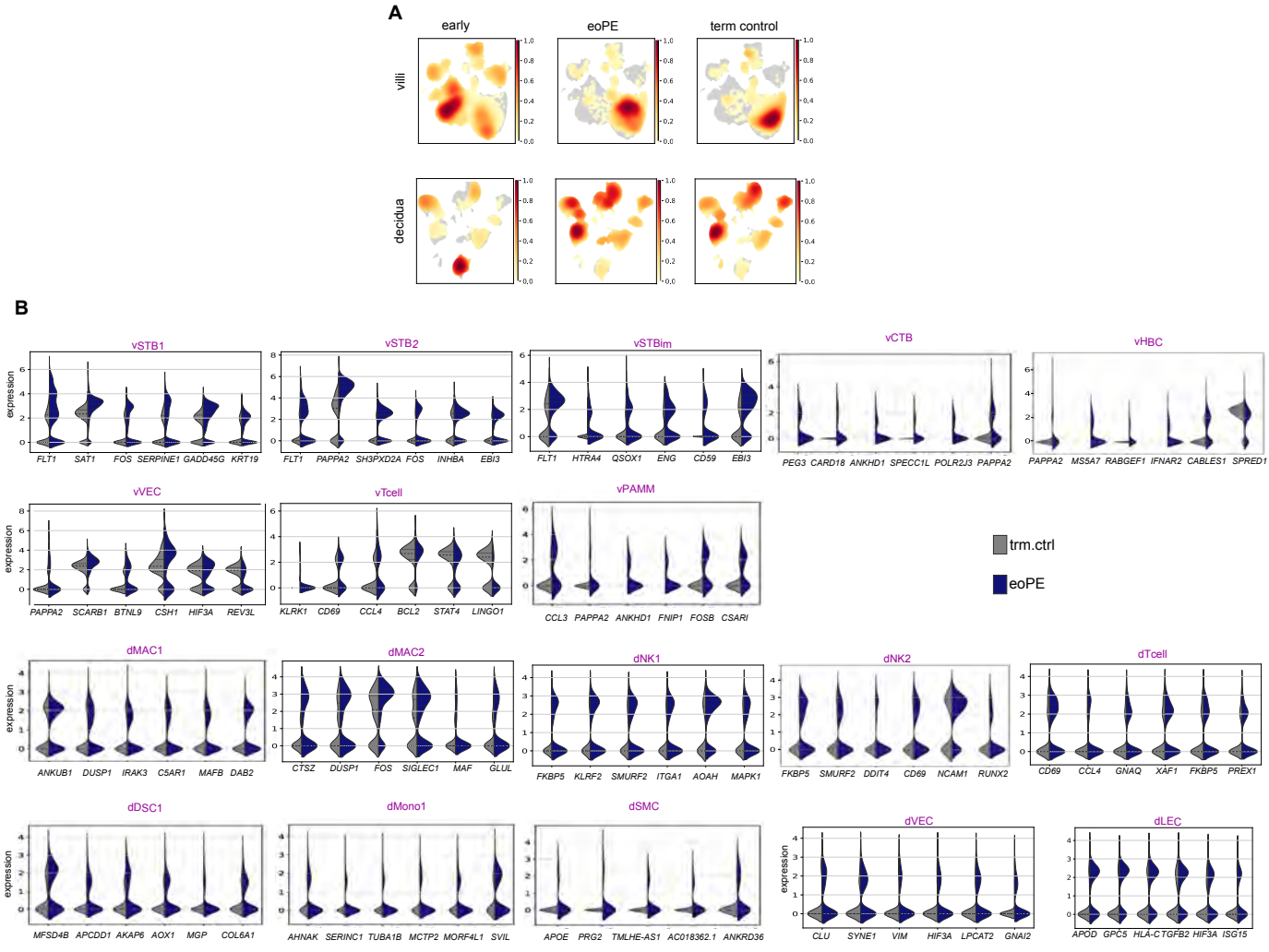
Supplemental Figure S5 | Known and novel cell types and states at the maternal-fetal interface are characterised by discrete and functionally relevant markers.

(A,B) Dot-plots featuring known and specific novel markers characterising each cell type or state in the villi (A) and decidua (B); computed by Logistic Regression based Generalized Linear model (Bonferroni adjusted two-sided, $P < 0.05$ and log2FC cut-off of ±0.25). Genes are scaled across clusters. Placental F13A1+/FGF13+ resident macrophages (Hofbauer cells, vHBC) uniquely express hyaluronan receptor LYVE1 in the immune cell subset, suggested to maintain arterial tone and have pro-angiogenic functions [18,24]. We additionally identify antigen presenting HLA-DRA+ placenta associated maternal monocytes/macrophages (vPAMM) which are villi-associated and are extra-embryonic or maternal in origin [24]. We identify villi myocytes through their expression of AGTR1. **(C)** Pearson correlation map of the highly variable genes (n = 6000) for identified cell-types and states. The clear uniqueness shared by STB groups also reveals clear differences in transcriptome between the novel vSTBim nuclei state and other STB nuclei subgroups. Because traditional characterization markers (such as CGA, CYP19A1, KISS1) used to describe STB are fulfilled by all groups, this suggests that vSTBim has additional unknown functions in pregnancy. Expression data of early late control and preeclampsia samples is shown. CTB, villous cytotrophoblast; STB, syncytiotrophoblast; DSTB, deported STB; CCT, cell column trophoblast; EVT, extravillous trophoblast; VEC, vascular endotheli-al cell; LEC, lymphatic endothelial cell; LECP, LEC progenitor; SMC, smooth muscle cell; MC, myocyte; FB, fibroblast; EpC, epithelial cell; MSC, mesenchymal stem cell; DSC, decidual stromal cell; EB, erythroblast; HBC, Hofbauer cell; PAMM, placenta-associated maternal macrophage; Mono, monocyte; MAC1, M1-like macrophage; MAC2, M2-like macrophage; NK, natural killer cell; PC, plasma cell; DC, dendritic cell, Granul, granulocyte; v, villous; d, decidual; p, proliferative; im, immature; pf, pre-fusion.



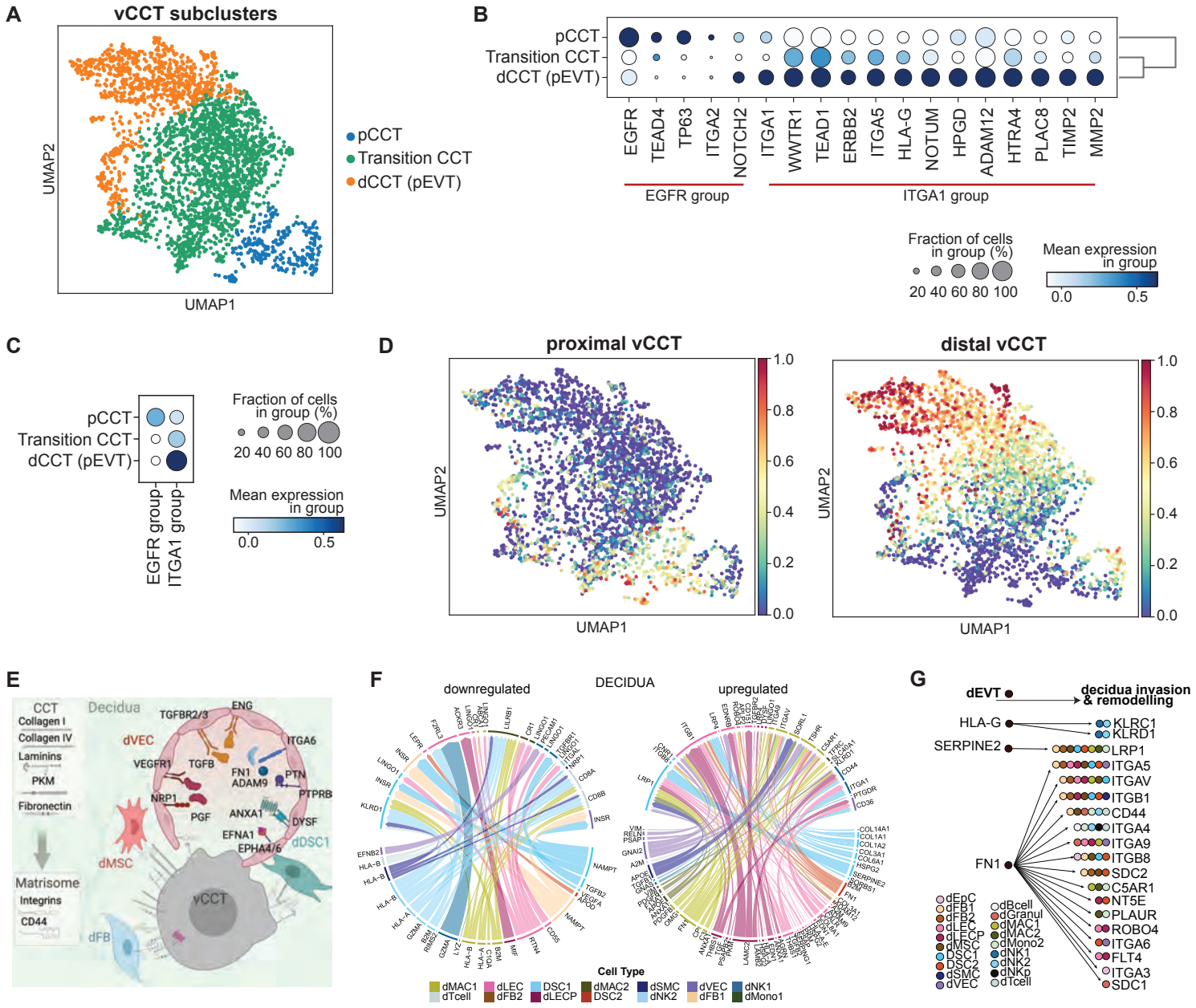
Supplemental Figure S6 | Early human trophoblast differentiation and the syncytiotrophoblast cell fate nuclear phenotypes

(A) Developmental trajectory of early villi trophoblast differentiation coloured by pseudotime; a numeric value is assigned to each nucleus to depict its hierarchical progression along the differentiation process. The width of each branch is proportional to the total number of cells. Pseudotime is a proxy of developmental time. **(B)** Diffusion map of early trophoblasts reflecting important markers for lineage commitment paths, and a developmental continuum for each differentiating branch based on ordering cells by transcriptomic similarities. **(C)** Flat tree plot, where trajectory branches are represented as straight lines and each dot represents a single nucleus. Assigned branches, branch lengths, and distances between nuclei are preserved from the space from which the trajectory was inferred. At first, the tree structure learned from a higher-dimension UMAP space was approximated by linear segments (each represents a branch) and subsequently, mapped to a 2D plane based on an improvised version of the force-directed layout Fruchterman-Reingold algorithm. **(D)** Dot-plot with markers characterising the villous syncytiotrophoblast (vSTB) nuclei states; computed by Logistic Regression based Generalized Linear model (Bonferroni adjusted two-sided, $P < 0.05$ and log2FC cut-off of ±0.25). Genes are scaled across clusters. **(E)** Dotplot featuring gene expression of markers characterising the two predominant nuclear states in the syncytial layer. **(F)** Comparative heatmap showing enriched pathways/processes in vSTB1 and vSTB2 differential genes; colored by -log(p-value), as measured by Metascape [48]. **(G)** Dotplot featuring gene expression of key sex-associated genes on male fetus placenta samples only. Decidual cell groups are the only ones with high expression of XIST, suggesting decidual samples are of maternal orgin. XIST expression in the villous derived T cell & placenta-associated maternal macrophage (PAMM) groups suggests these are also maternal in origin, and captured in the villi as invaded or adhered cells.
CTB, villous cytotrophoblast; STB, syncytiotrophoblast; DSTB, deported STB; CCT, cell column trophoblast; EVT, extravillous trophoblast; VEC, vascular endotheli-al cell; LEC, lymphatic endothelial cell; LECP, LEC progenitor; SMC, smooth muscle cell; MC, myocyte; FB, fibroblast; EpC, epithelial cell; MSC, mesenchymal stem cell; DSC, decidual stromal cell; EB, erythroblast; HBC, Hofbauer cell; PAMM, placenta-associated maternal macrophage; Mono, monocyte; MAC1, M1-like macrophage; MAC2, M2-like macrophage; NK, natural killer cell; PC, plasma cell; DC, dendritic cell, Granul, granulocyte; v, villous; d, decidual; p, proliferative; im, immature; pf, pre-fusion.



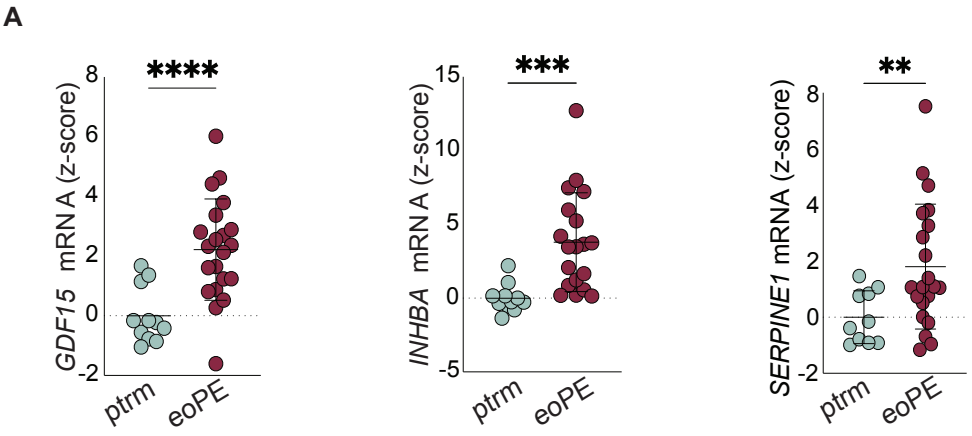
Supplemental Figure S7 | Transcriptomic dysregulation in pre-eclampsia has global and cell-specific patterns likely reflecting specific functional implications

(A) Nuclei from villi (top row) and decidua (bottom row) included in this study visualised in UMAP gaussian kernel density estimations. High density values indicate strong contribution of the cells to the overall dataset clusters. Density values are scaled between 0 and 1 to enable comparison between gestational timepoints and conditions. n = 12 deciduas (3 early, 4 term control, 5 eoPE), n = 21 placentas (10 early, 6 term control,5 eoPE). **(B)** Violin plots illustrating the most cell-type specific dysregulated genes in preeclampsia compared to late control. n = 12 deciduas (3 early, 4 term control, 5 eoPE), n = 21 placentas (10 early, 6 term control,5 eoPE). CTB, villous cytotrophoblast; STB, syncytiotrophoblast; DSTB, deported STB; CCT, cell column trophoblast; EVT, extravillous trophoblast; VEC, vascular endotheli-al cell; LEC, lymphatic endothelial cell; LECP, LEC progenitor; SMC, smooth muscle cell; MC, myocyte; FB, fibroblast; EpC, epithelial cell; MSC, mesenchymal stem cell; DSC, decidual stromal cell; EB, erythroblast; HBC, Hofbauer cell; PAMM, placenta-associated maternal macrophage; Mono, monocyte; MAC1, M1-like macrophage; MAC2, M2-like macrophage; NK, natural killer cell; PC, plasma cell; DC, dendritic cell, Granul, granulocyte; v, villous; d, decidual; p, proliferative; im, immature; pf, pre-fusion.



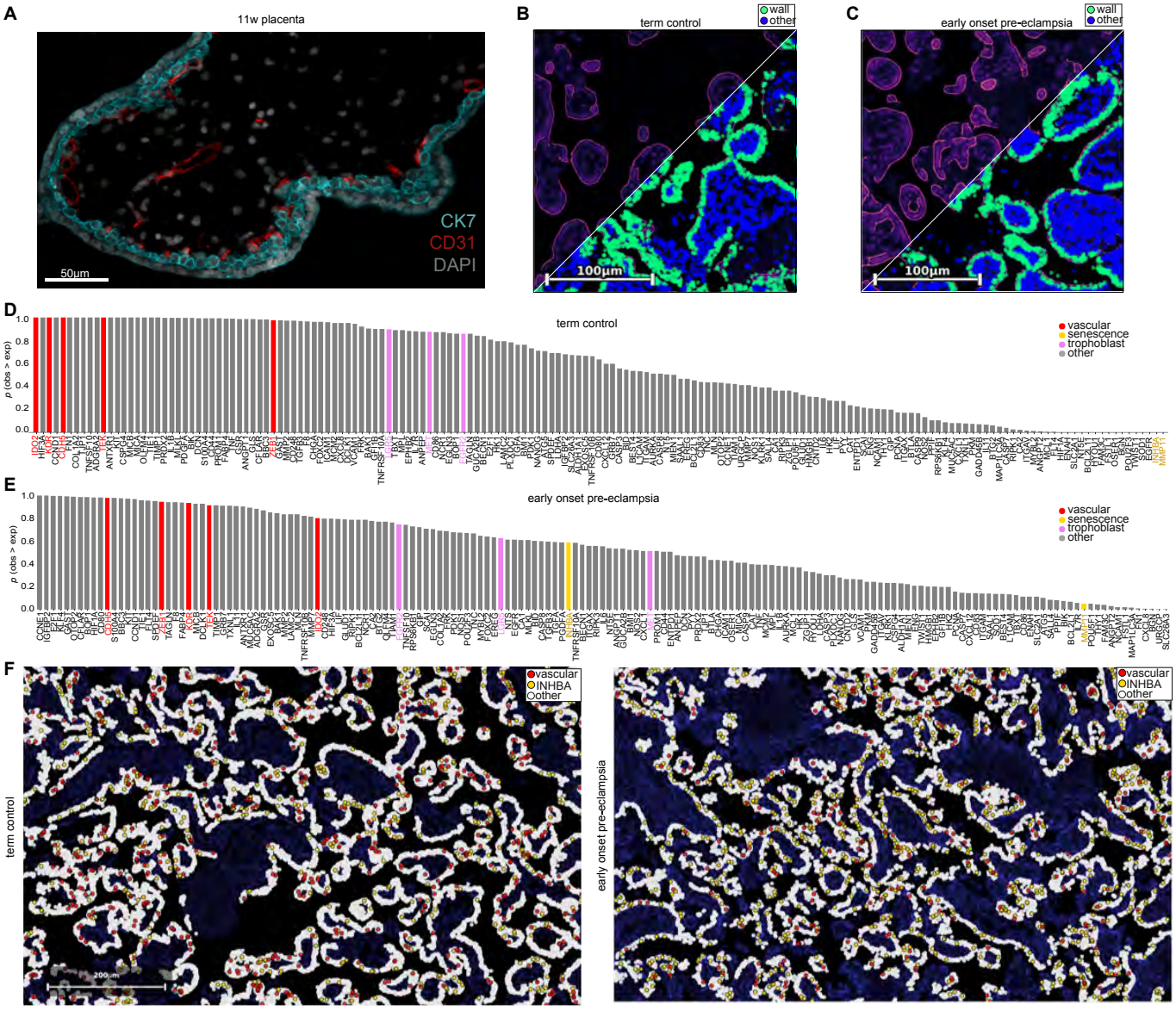
Supplemental Figure S8 | Subcluster analysis of invasive-phenotype cell column cytotrophoblast cell type

(A) UMAP embedding visualizing vCCT subclusters based on robust and specific marker profiles. **(B)** Dotplot depicting expression of key genes relevant for each subgroup. Fraction of cells per group expressing a gene is size coded; normalized mean expression is color-coded. **(C)** Module scores of two signature sets (derived from computational and curated subcluster markers) revealing relatively robust expression of Set1 (EGFR, TEAD4, TP63, ITGA2, NOTCH2) in pCCT and that of Set2 (ITGA1-MMP2). **(D)** Feature plot showing the signature scores of CCT subclusters using a gene expression score using the EGFR-set (left) and ITGA1-set (right) as described in (C). **(E)** Schematic summary of the cell interactions that occur during CCT from the villi onto the vascular and matrisome cell compartments of the decidua in early pregnancy (See supplementary Table 6 for a detailed summary of interactions). Schematic created with BioRender.com. n = 4 early control deciduas, n = 10 early control placentas. **(F)** Overview of selected receptor-ligand interaction network revealing dysregulation during early onset pre-eclampsia (eoPE) within immune and vascular matrisome compartments in decidua. Interactions were analyzed using Connectome and differentially expressed receptors/ligands were profiled using multivariate Logistic Regression (p-values adjusted for multiple testing). Ribbon color represents the respective source (or, ligand sending) cell types/states, and is proportional to the edge-weight score (product of normalized LFC of both partners). Networks shown are limited to edges where both receptor and ligand log2FC ± 0.25 between conditions (positive indicates upregulated in eoPE and vice-versa). The minimum percentage of cell groups expressing a receptor (or, ligand) is 10%. **(G)** Summary of selective receptor-ligand interactions for invasive extravillous trophoblasts (dEVT) onto decidual receptors at the maternal-fetal interface analyzed using Connectome. The interaction edges for whom ligand and receptor z-score > 0.25 were included. Further edges were filtered using the Diagnostic Odds Ratio of dEVT cut-off of 5, mean ligand expression 1.5, edge score > 3, and a minimum percentage of source expressing the ligand: 50%. n = 9 deciduas (4 term control, 5 eoPE). Values are scaled from 0-1 for comparison. CCT, cell column trophoblast; EVT, extravillous trophoblast; p, proximal; d, distal.



Supplemental Figure S9 | Transcriptomic synyctiotrophoblast dysregulation and senescence in preeclampsia is captured at the whole tissue lysate level

(A) mRNA expression of senescence-associated markers *GDF15, INHBA* and *SERPINE1* which are significantly upregulated in eoPE placental tissue in a multi-centre cohort comparing eoPE with gestational age matched preterm controls (ptrm: n=11; eoPE: n=23)(sig.level **<0.01, ***<0.001, ****<0.0001; unpaired two-tailed t-test with Welch's correction).



Supplemental Figure S10 | Within tissue communication is disturbed in early onset preeclampsia leading to disorganised STB senescence signalling

(A) Immunofluorescence image of a human placenta highlighting the main villi wall (CK7 positive trophoblasts) and the fetal vessels in both the stroma and close proximity to the trophoblast layer. **(B,C)** Illustration of the villi wall detection algorithm. The signal intensity of the DAPI stain was smoothed using a Gaussian filter with a bandwidth of 2 pixels. The canny edge detection algorithm was used to retrieve the villi walls (shown in the upper left figure halves as violet lines). Molecules detected within a radius of 5µm to any edge point were considered part of the villi wall, all other molecules were discarded for further analysis. (d,e) Spatial enrichment analysis around vessel markers in the villi walls for term control (B) and early onset pre-eclamptic (C) slides. The bar plots show the gene-wise p-values of a binomial test for a null hypothesis of equal count distribution among the in-situ sequencing (ISS) spots classified as 'vascular-proximal' (within 5µm to another vascular marker) and 'non-vascular-proximal' (outside 5µm to another marker), which can be violated towards the proximal (p>0.5, indicating enrichment around vessels) or the non-proximal (p<0.5, indicating reduction around vessels) side. **(D, E)** Notably, the occurrence of senescence markers INHBA and MMP11 is significantly lower in the vicinity of vascular markers (both p<0.01, with the lowest p-values out of all genes tested) in the term control sample (D). This effect is lost in the eoPE sample (E), where the count decrease of MMP11 is above the significance threshold (p~0.05) and INHBA is enriched around vascular markers (p~0.59). **(F)** Topography of wall molecules in the term control and pre-eclamptic samples used for the analysis presented in panels (D) and (E) over a DAPI stain background. n = 2 (1 slide per group). Yellow circles indicate INHBA molecules, red circles indicate vascular marker molecules, and all other molecules in the wall are rendered in white.