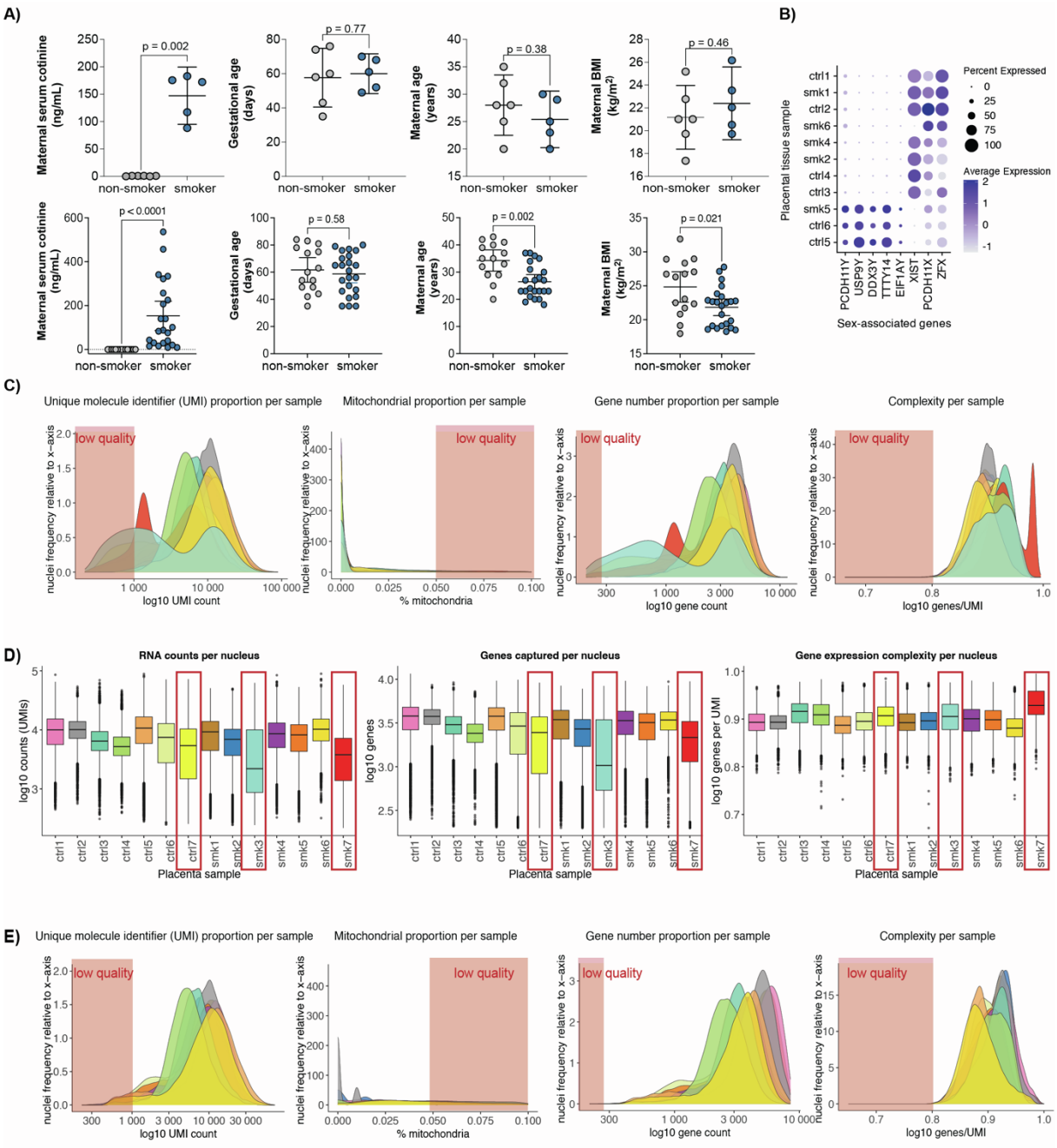
Cell Systems, Volume *17*

Supplemental information

Maternal smoking in early pregnancy disrupts placental function through syncytiotrophoblast and macrophage dysregulation

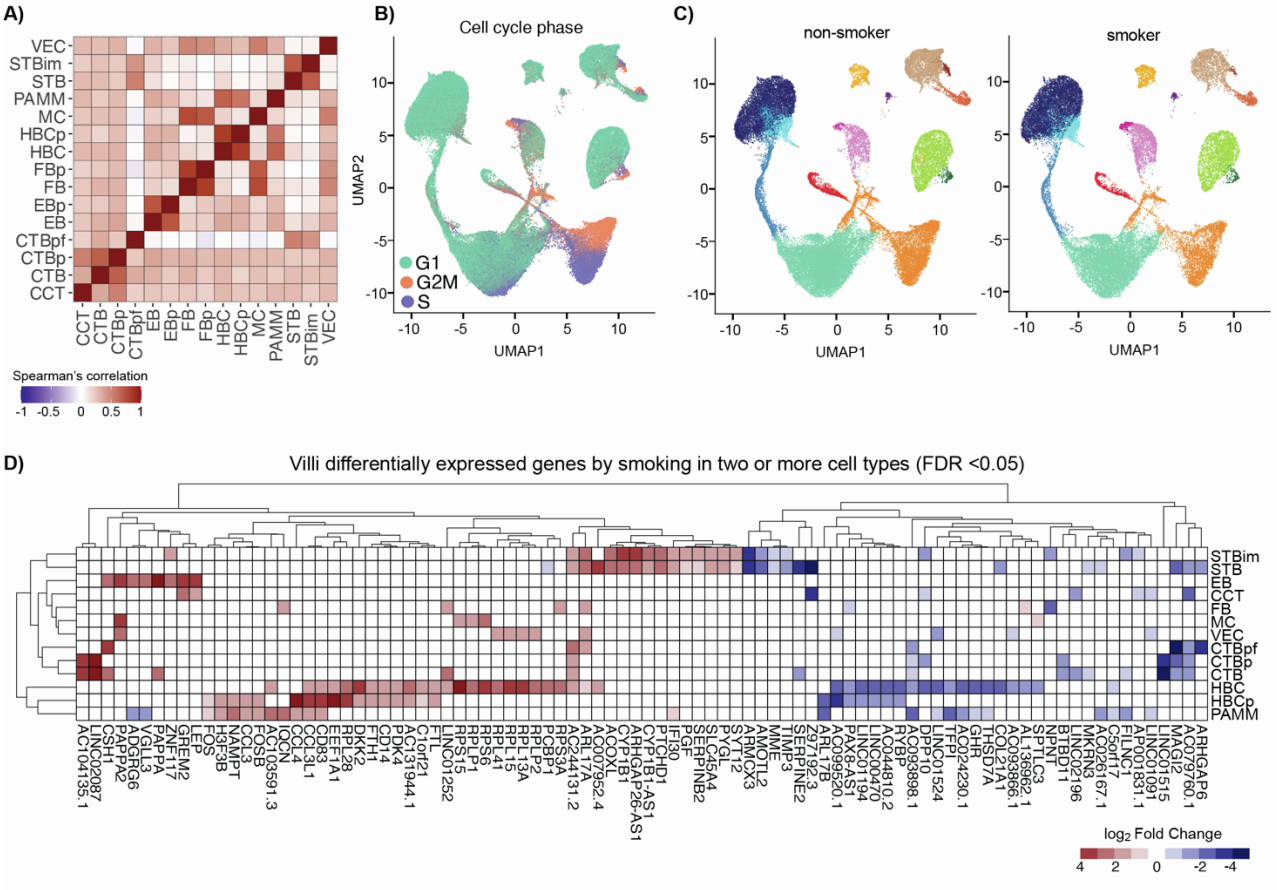
Daniela S. Valdes, Jose Nimo, Olivia Nonn, Juliane Ulrich, Désirée Forstner, Sandra Haider, Miriam Ressler, Marc Pignitter, Barbara Obermayer-Pietsch, Dominik N. Müller, Ralf Dechend, Martin Gauster, Fabian Coscia, and Florian Herse

SUPPLEMENTARY INFORMATION
Valdes et al. Maternal smoking in early pregnancy disrupts placental function through syncytiotrophoblast and macrophage dysregulation. 2026



Supplementary Figure 1. Baseline characteristics of included tissue cohorts and quality control of snRNAseq placental samples.
(A) Top: Baseline characteristics of cohort used for RNA sequencing and spatial proteomics. Participants were matched for gestational age, maternal age and maternal body mass index (BMI) (*n* = 5 smokers, *n* = 6 non-smokers). Bottom: Baseline characteristics of validation cohort. Participants were matched for gestational age, whilst maternal age and maternal body mass index (BMI) were lower in smokers, consistent with epidemiological data (*n* = 23 smokers, *n* = 14 non-smokers). Differences between groups assessed via two-tailed unpaired t-tests with Welch's correction.
(B) Dotplot with expression values of sex-associated genes per placental tissue used to infer fetal sex in the dataset. Expression plotted is scaled gene-wise. Samples with expression of male-associated gene *PCDH11Y, USP9Y, DDX3Y, TTTY14* and *EIF1AY* were classified as

male, whilst those expressing female-associated genes *XIST, PCDH11X* and *ZFX* were classified as female.
(C) Density plots of key quality control metrics for snRNAseq samples used. These include UMI counts per samples as an indicator of transcriptional activity captured; percentage of mitochondrial reads per sample as indicator of sample integrity; gene number per sample as insight into transcriptional diversity; the ratio between genes captured per UMI as measure for transcriptional complexity. Low quality thresholds are indicated by red boxes.
(D) Boxplots summarizing key metrics across individual samples. Samples falling outside of the expected range that did not pass quality control thresholds are highlighted in red boxes and excluded from further analyses.
(E) Density plots of consistent and reliable quality control metrics from included high-quality samples, ensuring dataset robustness for downstream analyses. Thresholds set excluded nuclei from high-quality samples that had less than 1000 UMI count, more than 0.5% percentage mitochondrial counts, less than 200 gene counts and a ratio of less than 0.8 log10 genes per UMI.



Supplementary Figure 2. Dysregulated cell types and states in the early human placenta display minimal global dysregulation patterns.

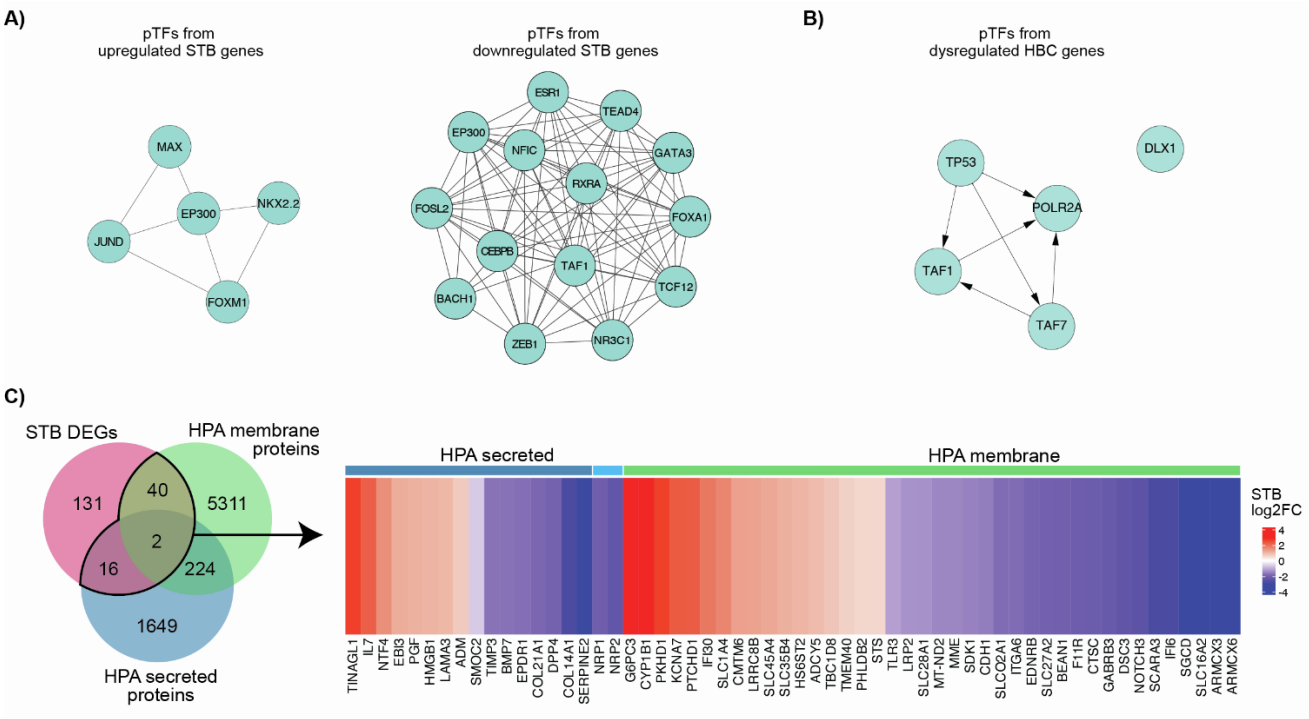
(A) Spearman correlation heatmap of whole-transcriptome correlations between maternal smokers and non-smokers.

(B) Uniform manifold approximation projection (UMAP) embedding of calculated cell cycle phase scores per nucleus. As expected, proliferating states map to the G2M and S phases, while non-proliferating states primarily map to the G1 cell cycle phase.

(C) UMAP of harmonized dataset colored by annotated cell types and states separated by maternal smoking status (left, non-smoker; right, smoker). Colors map to annotated cell types in Figure 1. n = 50,052 non-smoker nuclei, n = 38,756 smoker nuclei.

(D) Heatmap of dysregulated genes (log2FC) by maternal smoking present in two or more annotated cell type or state groups at a 5% FDR.

CTB, cytotrophoblast; STB syncytiotrophoblast; CCT, cell column trophoblast; HBC, Hofbauer cell; PAMM, placental-associated maternal macrophage; EB, erythroblast; FB, fibroblast; MC, myocyte; VEC, vascular endothelial cell; p, proliferating; pf, pre-fusion; im, immature.

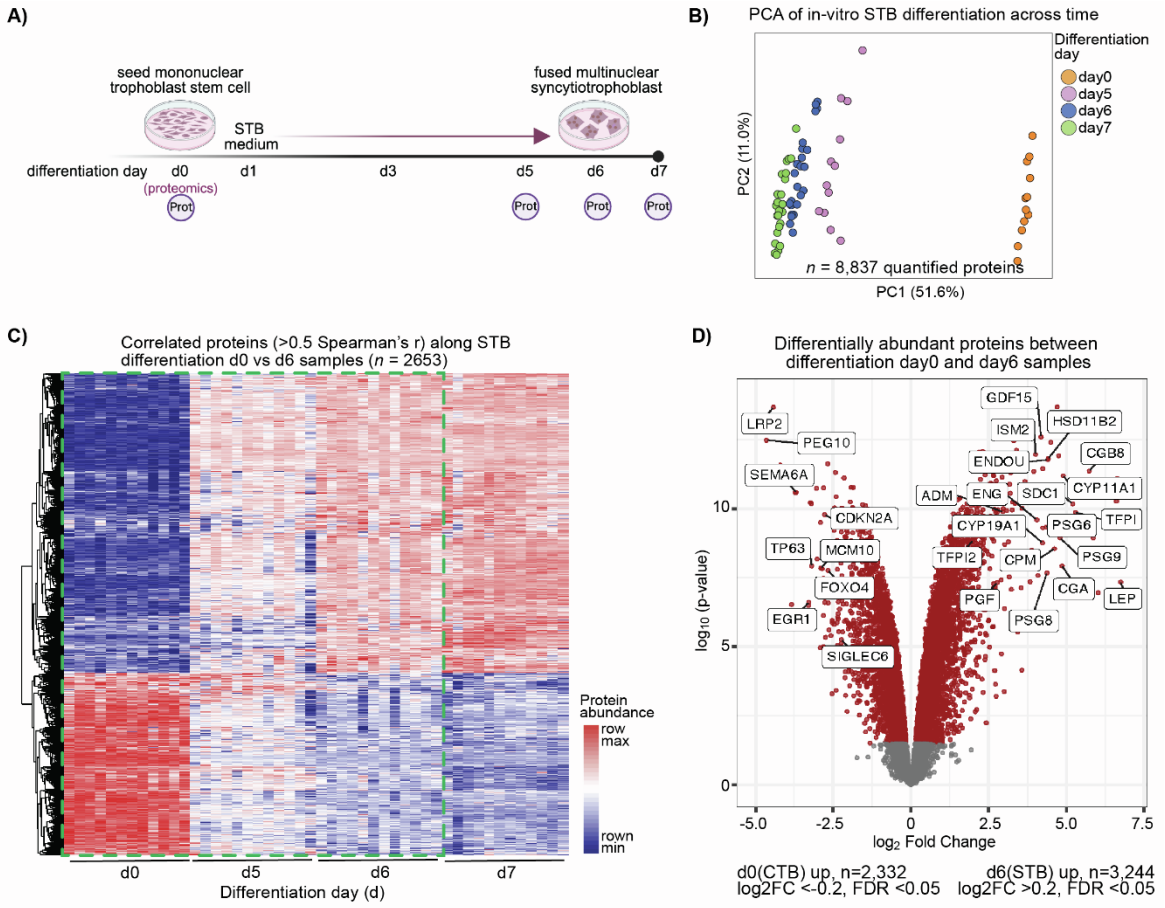


Supplementary Figure 3. Predicted transcription factors modulating orchestrated dysregulation by smoking.

(A) Predicted STB transcription factors (pTFs) modulating up (left) and down (right) hub gene networks calculated using the iRegulon cytoscape plugin. pTFs are visualized as a protein-protein interaction network (confidence level 0.15, no added proteins) using stringDB.

(B) Predicted HBC transcription factors (pTFs) modulating dysregulated hub gene networks calculated using the iRegulon cytoscape plugin. pTFs are visualized as a protein-protein interaction network (confidence level 0.15, no added proteins) using stringDB.

(C) Left: Venn diagram of STB DEGs (absolute $log_2FC > 0.25$, FDR < 0.1) overlapping with human membrane and secreted proteins as annotated by the human protein atlas (HPA) database. Right: Heatmap of overlapping STB dysregulated proteins as highlighted in (B) ordered by their HPA annotation membership. Heatmap color represents measured log_2FC between groups. Dark blue top legend represents HPA secreted annotated genes, light blue represents genes annotated as both secreted and membrane, green represents HPA membrane annotated genes.



Supplementary Figure 4. Proteomics characterizes in-vitro primary human trophoblast progenitor and differentiated states.

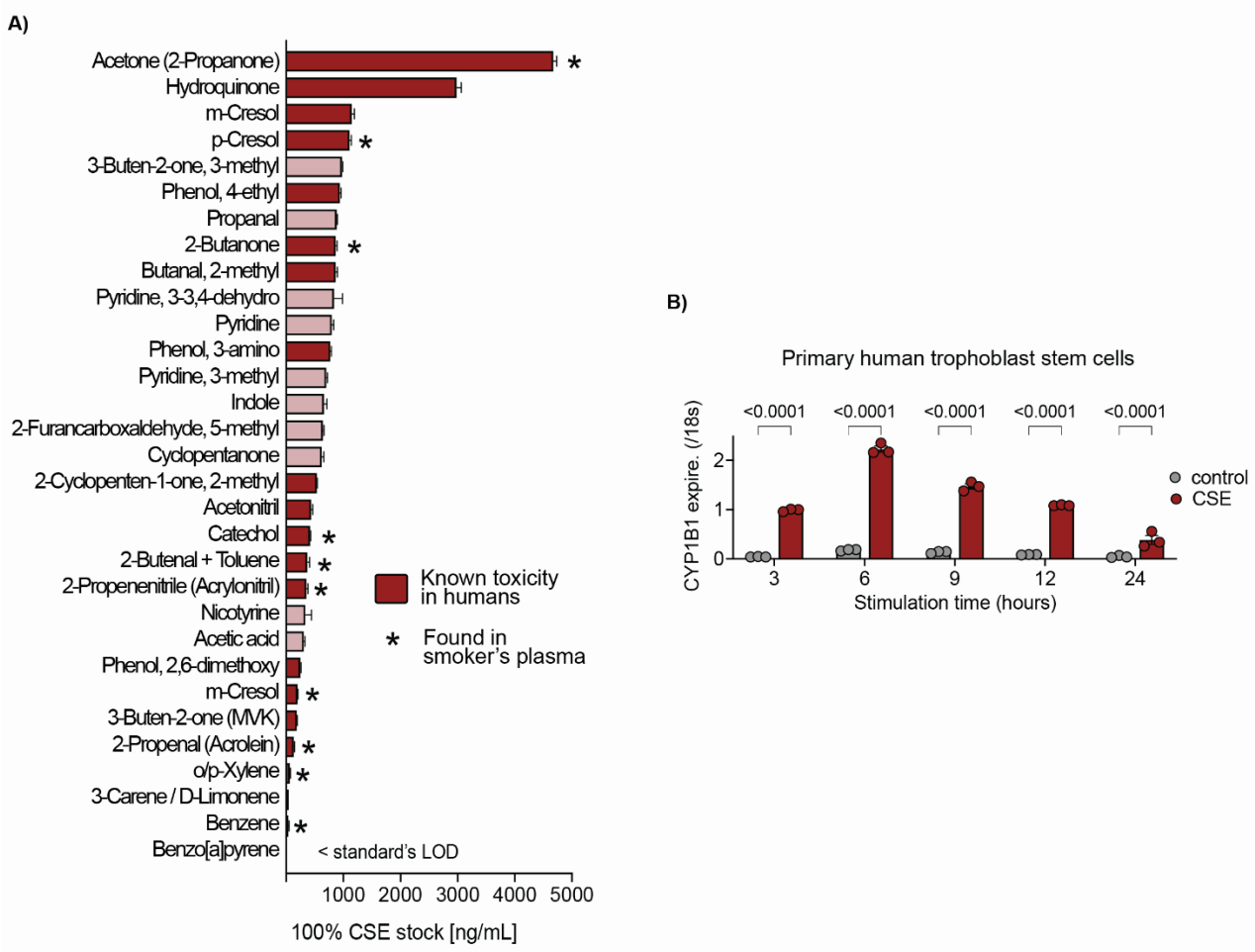
(A) Schematic representation of published protocol for the six day (d) directed differentiation of TSCs towards the STB phenotype *in-vitro*[16]. High-sensitivity mass-spectrometry based proteomics was used to evaluate proteomic profiles of undifferentiated cells (d0), incompletely fused (d5) and fused (d6, d7). (*n* = 3 independent experiments with technical quadruplicates per timepoint). Schematic created with BioRender.com.

(B) Principal component graph of quantified relative protein abundances across all timepoints in the experiment (*n* = 8,837 proteins after quality control). There is clear separation of proteomic profiles across time with no considerable batch effects. (*n* = 3 independent experiments with technical quadruplicates per timepoint).

(C) Heatmap of Spearman correlated proteins above an absolute r value of 0.5 between day 0 (TSC) and day 6 (STB) of differentiation as marked by the dotted green line (*n* = 4,363 proteins). Protein abundances are scaled row-wise.

(D) Volcano plot of differentially abundant proteins between d0 and d6 of differentiation, illustrating the deep and successful phenotypic shift across time, where each dot represents a protein. Dots in red represent proteins under a 5% FDR threshold and an absolute log_2FC threshold > 0.2. Labelled proteins are canonical to each phenotype. Differential abundance was calculated on bulk abundances per cell type or state using a generalized linear model in the edgeR package.

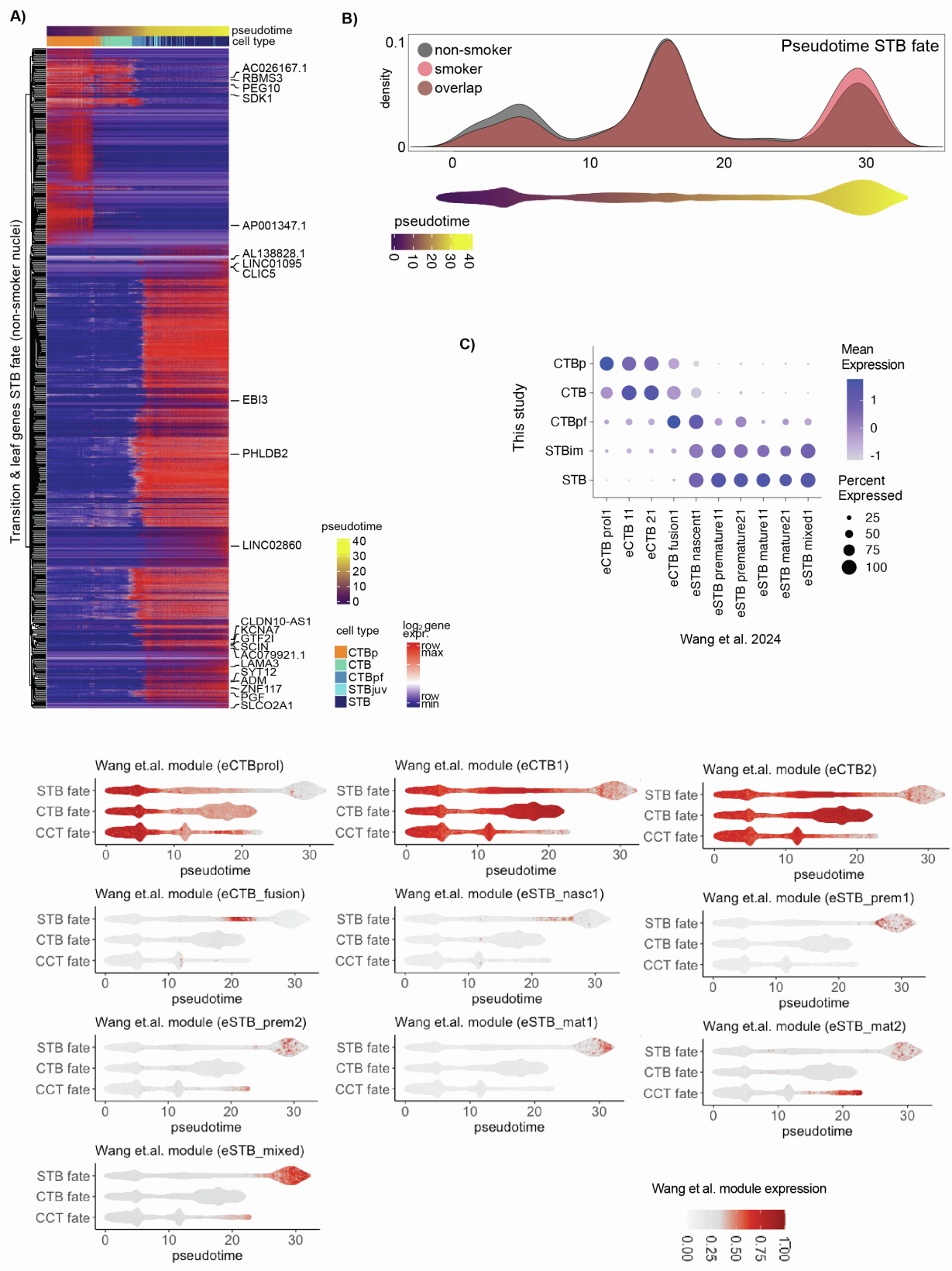
TSC, trophoblast stem cell; STB, syncytiotrophoblast; CTB, cytotrophoblast.



Supplementary Figure 5. Validation of cigarette smoke extract for in-vitro assessment of smoke component effects on cellular function.

(A) Bar plot of mean concentration of volatile and semi-volatile organic compounds present in a pooling of all 4x 100% cigarette smoke extract (CSE) batch stocks quantified by high-sensitivity gas-chromatography mass-spectrometry. Whiskers represent standard deviation. Red color indicates components known to have toxic cellular effects in humans, stars indicate components that have been measured in the saliva and circulation of adult smokers [179–182]. (*n* = 3 independent measurements in triplicate).

(B) Gene expression of markers associated with smoking by quantitative PCR on primary derived human trophoblast stem cells (bottom). (*n* = 3 independent experiments with pooled expression of 3 technical replicates per timepoint). Significance tested via 2-way ANOVA with Tukey's multiple comparisons.

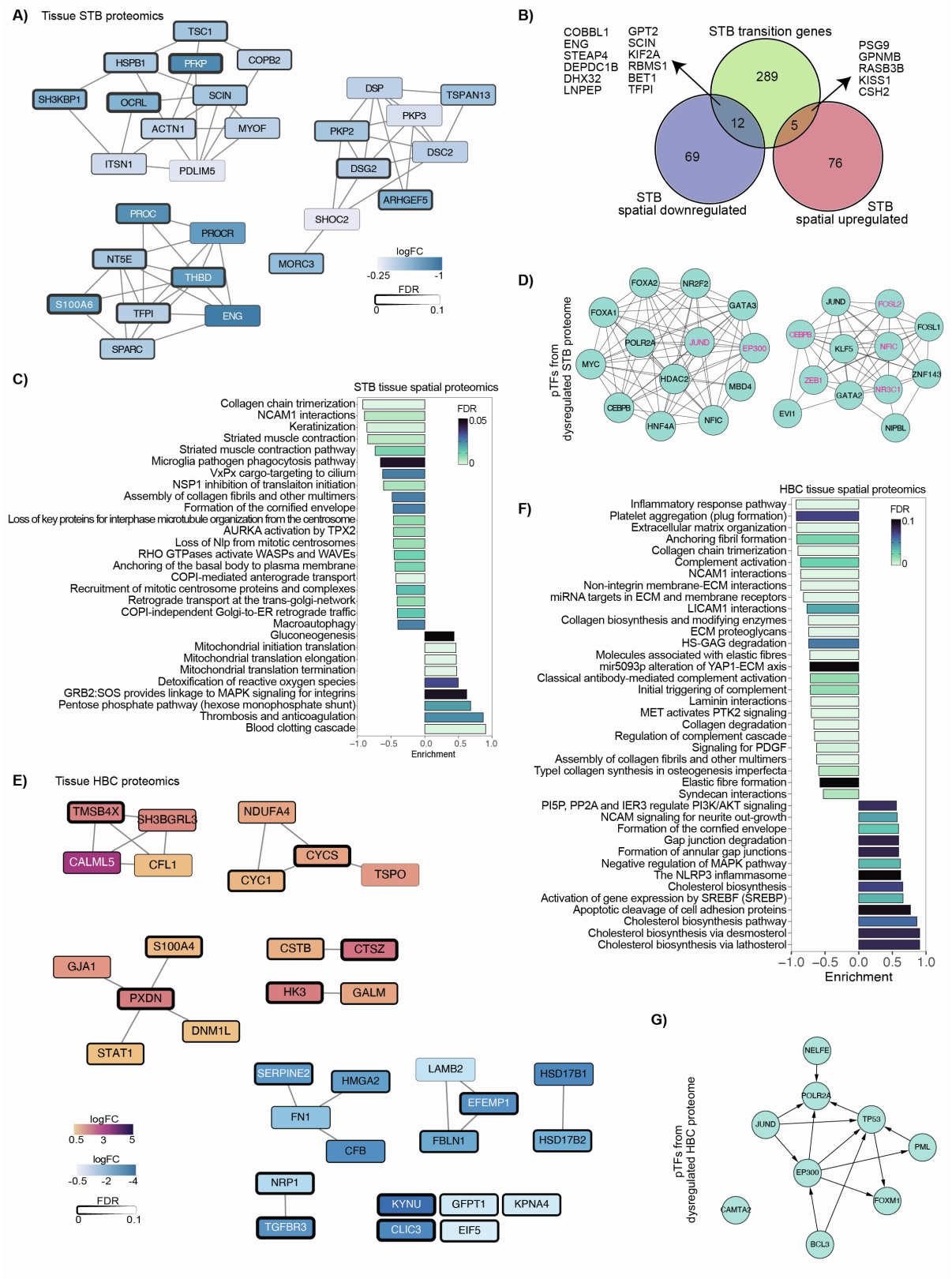


Supplementary Figure 6. Preferential differentiation towards the syncytiotrophoblast lineage by smoking.

(A) Heatmap of genes demarcating terminally differentiated cell states (leaf) and dynamic across pseudotime (transition) in non-smokers (n = 36,022 nuclei). Leaf genes were calculated as differential expression between nuclei corresponding to the first 20% and the last 20% values of pseudotime, with a concomitant absolute $\log_2$FC > 0.25 between timepoints. Pseudo-

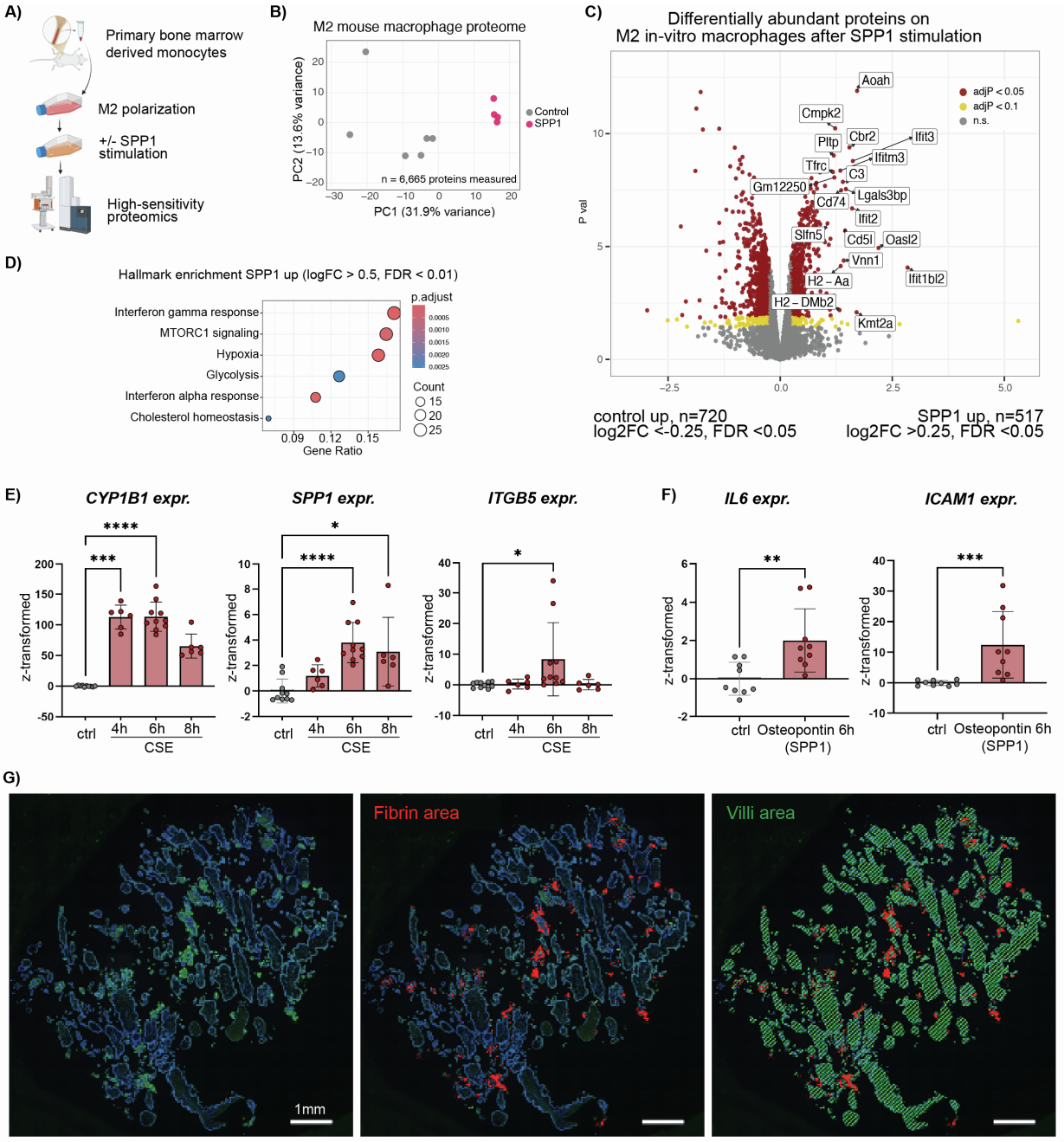
bulk counts per cell type or state using a generalized linear model in the edgeR package that corrected for effects of sequencing batch and fetal sex were used. Transition genes were calculated based on scaled expression counts for genes with an absolute $log_2FC > 0.25$ between start and end of the trajectory and having an absolute Spearman correlation coefficient greater than 0.4.
(B) Density plot highlighting pseudotime distribution in smokers and non-smokers for the STB lineage. The progression colored by pseudotime modelled is shown as reference.
(C) Dotplot with mean expression of module scores calculated using the top 20 most differential genes per group for trophoblast subgroups as recently described by Wang et al.[33]
(D) Scatter plot of modelled pseudotime values separated by inferred lineage. Each dot represents a nucleus, colored by its scaled module score expression as described in (C).

CTB, cytotrophoblast; STB syncytiotrophoblast; CCT, cell column trophoblast; p, proliferating; pf, pre-fusion; im, immature.



Supplementary Figure 7. Master regulators of placental macrophage and barrier dysregulation in maternal smoking at the proteo-transcriptomic level.
(A) Hub genes of upregulated (upper panel, $log_2FC > 0.25$, FDR < 0.1) tissue placental STB proteins with maternal smoking in early pregnancy, based on protein-protein interaction background stringDB network followed by topological analysis of ten different methods, clustered into communities based on Newman and Girvan's algorithm. Network analysis was

performed in the cytoscape software. Color represents log_2FC, node border thickness corresponds to differential abundance analysis' FDR.
(B) Venn diagram of STB tissue dysregulated proteins (logFC ± 0.25, FDR < 0.1) overlapping with transition genes from the trophoblast trajectory modelling (logFC ± 0.25, Spearman's r ± 0.4).
(C) Pathway enrichment analysis of dysregulated tissue STB proteins (absolute FC > 0.25, FDR < 0.1). Enrichment was done based on ranked log fold changes spanning the gene set enrichment analysis (GSEA), reactome and the WikiPathways databases. Positive enrichment values represent pathways enriched in smokers; negative values represent those enriched in non-smokers. Bar color is mapped to dysregulation FDR.
(D) Predicted tissue proteomics STB transcription factors (pTFs) calculated based on up- and down- regulated hub gene networks in (E) using the iRegulon cytoscape plugin. pTFs are visualized as a protein-protein interaction network (confidence level 0.15, no added proteins) using stringDB.
(E) Hub genes of upregulated (upper panel, $log_2FC > 0.25$, FDR < 0.1) and downregulated (lower panel, $log_2FC < -0.25$, FDR < 0.1) placental HBC tissue proteins with maternal smoking in early pregnancy, based on protein-protein interaction background stringDB network followed by topological analysis of ten different methods, clustered into communities based on Newman and Girvan's algorithm. Network analysis was performed in the cytoscape software. Color represents log_2FC, node border thickness corresponds to differential abundance analysis' FDR.
(F) Pathway enrichment analysis of dysregulated tissue HBC proteins (absolute FC > 0.25, FDR < 0.1). Enrichment was done based on ranked log fold changes spanning the gene set enrichment analysis (GSEA), reactome and the WikiPathways databases. Positive enrichment values represent pathways enriched in smokers; negative values represent those enriched in non-smokers. Bar color is mapped to dysregulation FDR.
(G) Predicted tissue proteomics HBC transcription factors (pTFs) calculated based on up- and down- regulated hub gene networks in (E) using the iRegulon cytoscape plugin. pTFs are visualized as a protein-protein interaction network (confidence level 0.15, no added proteins) using stringDB.



Supplementary Figure 8. In-vitro stimulation with SPP1 upregulates immune response markers in macrophages and trophoblast cells

A) Schematic experimental design. Mouse bone marrow derived monocytes were differentiated to M0 macrophages (BMDM) and polarized towards the M2-phenotype by IL-4 (10 ng/mL) and IL-13 (10 ng/ml) for 24h. Macrophages were stimulated with osteopontin (1ng/ul) for 24hours before being analysed at the proteomic level by mass-spectrometry based proteomics. Created with Biorender.com.
B) Principal component analysis of the proteome of M2-phenotype mouse bone marrow derived macrophages (n = 665 proteins measured).
C) Volcano plot of differentially abundant proteins in M2-phenotype mouse bone marrow derived macrophages following in-vitro SPP1 stimulation. The top 20 most dysregulated proteins are shown. Differentially abundant proteins were calculated by linear regression and Benjamini-Hochberg correction in the limma package.
D) Summary of pathway enrichment analysis of in-vitro upregulated proteins by SPP1 stimulation of mouse macrophages (absolute log2FC > 0.5, FDR < 0.05). Enrichment was

done based on ranked log2fold changes spanning the gene set enrichment analysis (GSEA) and Hallmark databases.
E, F) Gene expression of selected markers as measured via quantitative PCR of primary derived cytotrophoblast (CTB) cells following exposure to cigarette smoking extract (CSE [1:200 of 100%], E) and osteopontin (SPP1 [2.5ug/ul], F) in-vitro for variable timepoints as indicated. Differences calculated by ...
G) Placental perivillous fibrin deposition and platelet adhesion software-based analysis. Images of double-stained placenta tissue (Fibrin, CD42b) were obtained using an Evident Olympus VS200 Slide Scanner and images were afterwards subjected to Visiopharm (Version 2021.09) for a software-based quantification of fibrin. Villous regions were detected within each sample by applying a threshold on a feature combining a polynomial gradient filter of FITC (Fibrin) and a mean filter feature of the DAPI channel. Fibrin within the villous regions was detected by a threshold within the FITC channel. Adherent platelets were detected by measuring CD42b positive signals using the Cy3 channel. Fibrin area (μm^2) and platelet area (μm^2) was normalized to total villi area (μm^2). Scale bars represent 1mm.

Supplementary Tables

Supplementary Table 1. Participant characteristics. Demographic and clinical characteristics of the snRNA-seq/DVP omics cohort and the independent validation cohort (used for qPCR and PlGF ELISA), stratified by smoking status.

Supplementary Table 2. snRNA-seq differentially expressed genes for all cell types. Differential expression results (log2FC, FDR) between smokers and non-smokers for every annotated placental cell type/state in the single-nucleus RNA-seq data, computed from pseudobulk profiles with edgeR.

Supplementary Table 3. Trajectory leaf genes. Genes marking the terminal, differentiated end-points (leaves) of the three modelled trophoblast lineages (STB-, CTB-, and CCT-fate), calculated by pseudobulk edgeR DEG testing within each lineage using non-smoker nuclei only, capturing genes that define each terminal cell identity (FDR<0.05).

Supplementary Table 4. Trajectory transition genes. Genes whose expression correlates with pseudotime along each trophoblast differentiation trajectory (Spearman's $|r| \geq 0.4$, non-smoker nuclei), representing putative regulators of the progressive shift between progenitor and differentiated states.

Supplementary Table 5. DVP dysregulated proteins. Differential abundant protein results (log2FC, FDR) identified by deep visual proteomics in laser-microdissected STB, CTB, and HBC regions of interest, comparing smokers to non-smokers.

Supplementary Table 6. In-vitro proteomic comparison. Proteomic comparison of in-vitro differentiated primary trophoblast (STB) pellets harvested at differentiation day 7, contrasting 48-hour cigarette smoke extract (CSE) exposure versus untreated controls.

Supplementary Table 7. Cell-type resolved receptor-ligand interaction. Results inferred with CellChat, calculated separately per omics modality (snRNA-seq and DVP) and per smoking condition, then compared across the STB-CTB, CTB-HBC, and within-cell-type (autocrine) interaction sets, reporting communication probability and significance for each ligand-receptor pair.