RESEARCH

Disruption of maternal vascular remodeling by a fetal endoretrovirus-derived gene in preeclampsia

Xiaoli Gong[1†], Wei He[2†], Wan Jin[3,7†], Hongwei Ma[4,5†], Gang Wang[6,7,8,9†], Jiaxin Li[1†], Yu Xiao[6,7,8,9], Yangyu Zhao[1], Qiong Chen[3], Huanhuan Guo[3], Jiexia Yang[2], Yiming Qi[2], Wei Dong[10], Meng Fu[11], Xiaojuan Li[3,12], Jiusi Liu[3], Xinghui Liu[4,5*], Aihua Yin[2*], Yi Zhang[3*†] and Yuan Wei[1*]

†Xiaoli Gong, Wei He, Wan Jin, Hongwei Ma, Gang Wang, and Jiaxin Li contributed equally to this work.

†Yi Zhang is the lead contact.

*Correspondence:
xinghuiliu@163.com;
yinaiwa@vip.126.com;
zy@eulertechnology.com;
weiyuanbysy@163.com

¹ Department of Obstetrics and Gynecology, Peking University Third Hospital, Beijing, China
² Medical Genetic Center, Guangdong Women and Children Hospital, Guangzhou, China
³ Euler Technology, Beijing, China
⁴ Department of Obstetrics and Gynecology, West China Second University Hospital of Sichuan University, Chengdu, China
Full list of author information is available at the end of the article

Abstract

Background: Preeclampsia, one of the most lethal pregnancy-related diseases, is associated with the disruption of uterine spiral artery remodeling during placentation. However, the early molecular events leading to preeclampsia remain unknown.

Results: By analyzing placentas from preeclampsia, non-preeclampsia, and twin pregnancies with selective intrauterine growth restriction, we show that the pathogenesis of preeclampsia is attributed to immature trophoblast and maldeveloped endothelial cells. Delayed epigenetic reprogramming during early extraembryonic tissue development leads to generation of excessive immature trophoblast cells. We find reduction of de novo DNA methylation in these trophoblast cells results in selective overexpression of maternally imprinted genes, including the endoretrovirus-derived gene PEG10 (paternally expressed gene 10). PEG10 forms virus-like particles, which are transferred from the trophoblast to the closely proximate endothelial cells. In normal pregnancy, only a low amount of PEG10 is transferred to maternal cells; however, in preeclampsia, excessive PEG10 disrupts maternal vascular development by inhibiting TGF-beta signaling.

Conclusions: Our study reveals the intricate epigenetic mechanisms that regulate trans-generational genetic conflict and ultimately ensure proper maternal–fetal interface formation.

Background

Preeclampsia (PE), a pregnancy-specific disease, is characterized by endothelial dysfunction and unmanageable hypertension that leads to multi-organ damage in the expectant mother. Each year, PE affects 3–5% of pregnancies, leading to at least 42,000 maternal deaths globally [1, 2]. The curative effect of removing the placenta and fetus on PE indicates that the placenta may be the origin of all maternal syndromes. As one

of the "Great Obstetrical Syndromes" [3], PE exhibits a placenta that shares pathological features with other diseases, such as fetal growth restriction (FGR) and preterm birth [3]. Anatomically, PE is characterized by incomplete development of the villous tree [4], reduced vascularization into the terminal villus [4, 5], and incomplete remodeling of maternal uterine spiral arteries [6, 7].

Population genetic studies have shown that susceptibility to PE involves a genetic component. PE is strongly associated with both the maternal and fetal genomes but less so with environmental factors [8, 9]. For the fetal genome, it was proposed that PE risk is inherited from the paternal copy [9]. Genome-wide association scans (GWAS) have revealed that PE risk is associated with various loci that are implicated in maternal hypertension susceptibility, endothelial cell development, and the expression of the fetal vascular epidermal growth factor (VEGF) receptor gene fms-related receptor tyrosine kinase 1 (*FLT1)* [10–16], suggesting that the genetic factors influencing PE center on vascular development and function. A genetic association of PE to endothelial development genes in the maternal genome [15] and loci neighboring FLT1 in the fetal genome [13] suggests that PE results from the interaction between fetal and maternal cells.

Biochemically, PE is associated with elevated levels of angiotensinogen (AGT) [17–19], s-FLT1 [20–22], and soluble endoglin (s-Eng) [21, 23, 24], as well as reduced levels of PAPPA [25] and PlGF [21, 26–28] in the maternal serum, which collectively result in maternal peripheral vasoconstriction [29], decreased arterial compliance [30], and organ damage such as glomerular endotheliosis [31, 32]. Studies involving animal models [33, 34] and humans [6, 35] have revealed that these biochemical changes are attributed to trophoblast cells that invade the maternal uterus myometrium, suggesting that the interactions between trophoblast and endothelial cells may be central to PE pathogenesis.

Clinical, genetic, and anatomical data provide compelling evidence that the primary PE syndrome is due to a developmental defect in the maternal–fetal interface, i.e., the placenta. As an organ of the extraembryonic lineage, the placenta originates from the trophectoderm of the blastocyst. The specification and differentiation of placental cells are driven by genome-wide epigenomic reprogramming [36–39]. This overall epigenomic reprogramming during placental development is pivotal for proper placentation. This phenomenon not only establishes the proper cellular composition of the placenta but also regulates trans-generational conflict by modulating the gene expression levels that differentially affect maternal and fetal health. Shortly after zygote formation, global erasure of DNA methylation and H3K27me3 repressive marks results in a genome that is permissive for transcriptional activation during the zygote genome activation (ZGA) stage [40–43]. The reinstatement of repressive epigenetic marks on lineage specification genes during subsequent development is instigated by the priming of maternal and ZGA-active transcription factor (TF) [44–46], the binding of polycomb group protein [47–49], and the de novo DNA methylation of CpG-island promoters [50]. Recent studies have shown that genomic DNA from PE placentas tend to be hypomethylated compared with that from non-PE placentas [51–53], suggesting the occurrence of defects in DNA methylation during placenta development in PE. However, the precise molecular mechanism underlying the epigenetic and developmental alterations in PE remains unclear.

Gong *et al. Genome Biology*   (2024) 25:117

Page 3 of 46

In this study, through multi-omic integrative analysis of placentas from PE, non-PE, and twin pregnancies with selective intrauterine growth restriction, we revealed complex epigenetic processes regulating trans-generational genetic conflict to ensure proper maternal–fetal interface formation. Our results showed how disruption of such processes results in PE.

Results

Identification of pathogenic cells in PE

We sequenced single cells and bulk tissues of the placenta from both non-PE and PE pregnancies (Fig. 1a). Using $10 \times$ Genomics single-cell RNA (scRNA) and single-cell chromatin accessibility (scATAC), we sequenced fetal and maternal surfaces of placentas obtained from 11 donors, including those with normal pregnancy ($N=2/1/1$ [sample/placenta/donor]), gestational diabetes mellitus (GDM) pregnancy ($N=2/1/1$), preterm birth ($N=5/3/3$), PE and GDM (PE-GDM) pregnancy ($N=6/3/3$), and PE and FGR (PE-FGR) pregnancy ($N=4/2/2$), and a pair of placentas from a dizygotic twin pregnancy with selective intrauterine growth restriction (DCDA sIUGR, $N=2/2/1$, only fetal surface available) ("Methods," Fig. 1a,b, Additional file 1: Fig. S1a-c; Additional file 2: Tables S1 and Additional file 3: Table S2). Two (2) of these donors, a non-PE control donor

(See figure on next page.)

Fig. 1 Pathogenic cells underlying preeclampsia. **a** Schematic overview of multi-omic study for PE pathogenic mechanism. 10 × Genomics single-cell RNA and ATAC sequencing were performed with 11 donors (normal pregnancy: 1; GDM: 1; preterm birth: 3; PE and GDM (PE-GDM): 2; PE and FGR (PE-FGR): 3; DCDA slUGR: 1). Additionally, we supplemented this dataset with those of early-stage cells, i.e., public scRNA datasets pertaining to first-trimester placentas (PRJEB28266 and PRJNA492324) [54–57] and induced trophoblast stem cells (iTSCs, GSE150578 [58, 59]). Bulk genome-wide CpG capture bisulfite sequencing (DNAm-Seq) and ATAC sequencing were performed with 43 placenta samples (including 24 control, 5 GHT, 14 PE). By integration of single-cell and bulk sequencing, we annotated the major cell types in placenta; identified that trophoblast and endothelial cells were the most PE-associated cell types; reconstructed the developmental trajectory of trophoblast; revealed master transcriptional factors during trophoblast development, uncovered causal effector for PE pathological phenotype by genetic tracing and immunohistochemistry, and identified their downstream target cell. **b** UMAP projection of major cell types in scRNA cells, including trophoblasts. **c** UMAP projection of cell origin in scRNA. Trophoblast (including VCTp, VCT, SCT, and EVT) and Hoffbauer cell (HB) are fetal originated (blue); immune cells (including monocyte, M2, NK, T, and B cell) are from the maternal (orange); endothelial and fibroblast/stromal cells are contributed by both fetal and maternal. **d** UMAP projection of phenotype in scRNA. Cells of control and PE placentas are intermingled in each cell type. iTSC cells from GSE150578 [58, 59] are majorly integrated with trophoblast progenitors (VCTp). The cell number of each type in control and PE was labeled. **e** UMAP projection of scATAC cells. VCT/VCTp cells are mixed in this modality, while CD4 and CD8 T cells are clearly segregated. **f** Known marker gene expression of scRNA cell clusters. **g** Chromatin accessibility on marker genes of scATAC cell clusters. Color bar on the left side indicates the cell type in Fig. 1e. **h** PE-associated cell types identified by both phenotype-gene expression association and GWAS-chromatin accessibility association. Large panel: prevalence of Scissor-inferred PE-associated cells with scRNA (*x* axis) and the prevalence of SCAVENGE-inferred PE-associated cells with scATAC (*y*-axis). VCTp/VCT clusters are associated with PE in terms of hereditary risk but not gene expression, suggesting their implication in PE might due to abnormal differentiation but not their own function. In contrast, both endothelial cells and EVT are associated with PE hereditary risk loci as well as gene expression, implying the direct pathogenic role of them in PE. Small panel: (left) Scissor-inferred single-cell association with RNA expression profile from control or PE placenta, showing that PE is mostly associated with endothelial cells, EVT, SCT, VCTp, and monocytes. (right) SCAVENGE-inferred scATAC association with PE-associated GWAS loci, showing that PE GWAS loci are associated with VCT, VCTp, EVT, SCT, macrophages, and endothelial cells. *Abbreviation*: DNAm-seq: genome-wide CpG capture bisulfite sequencing; GHT: gestational hypertension; VCTp: trophoblast progenitor; VCT: villous trophoblast; SCT: synciotrophoblast; EVT: extravillous trophoblast; Mono: monocyte; NK: natural killer cell; M2: type-2 macrophage; HB: Hoffbauer cell; Fibro: fibroblasts/stromal cells; Endo: endothelial cells
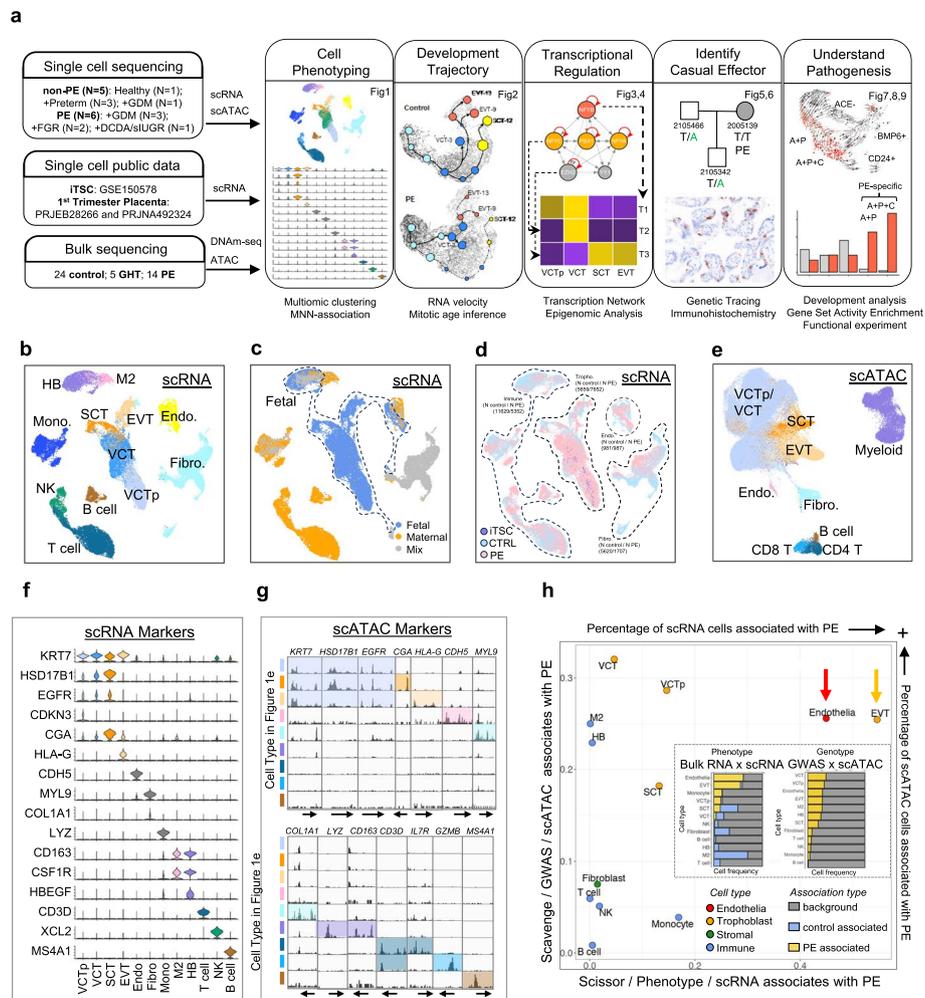
Gong *et al. Genome Biology* (2024) 25:117

Page 4 of 46

Fig. 1 (See legend on previous page.)

(2100042) and a DCDA sIUGR donor (2200314) (Additional file 2: Table S1), were not enrolled in the first study period (before 2021) but enrolled latter (after 2021) for internal validation. Hence, in Fig. 1b–h, we only used 9 donors out of 11. Additionally, we supplemented this dataset with those of early-stage cells, i.e., public scRNA datasets pertaining to first-trimester placentas [54–57] and induced trophoblast stem cells (iTSCs) [58, 59]. We also sequenced non-PE control (*N*=24), gestational hypertension (GHT) (*N*=5), and PE (*N*=14) placentas by genome-wide CpG capture bisulfite sequencing (DNAm-seq) and ATAC (Fig. 1a). These data were subjected to a multimodal integrative analysis. In the single-cell analysis, all non-PE samples were defined as control.

For scRNA analysis, datasets were integrated with potential batch effects removed by Harmony [60] prior to cell cluster identification using Seurat [61]. By identifying differentially expressed genes in each single-cell cluster (Fig. 1f), we annotated the major cell populations in the maternal–fetal interface in the scRNA dataset ("Methods," Fig. 1b and Additional file 1: Fig. S1d-e). These included the major cell populations in trophoblast, i.e., progenitor villous cytotrophoblast (VCTp) cells, villous trophoblast (VCT) cells, syncytiotrophoblast (SCT) cells, and extravillous trophoblast (EVT) cells; stromal

Gong *et al. Genome Biology*   (2024) 25:117

Page 5 of 46

cells with mixed arterial and venous endothelial cells, including vascular smooth muscle cells, pericytes, and fibroblasts; and immune cells, including macrophages, monocytes, NK cells, T cells, and B cells. These cells could be classified into three distinct groups based on their origin: fetal-originated cells (trophoblasts and fetal macrophage Hofbauer cells), maternal-originated cells (including T lymphocytes, B lymphocytes, maternal macrophages, monocytes, and NK cells), and mixed-origin cells (stromal cells, including endothelial cells, fibroblasts, smooth muscle cells, and pericytes) [55]. We further analyzed the expression levels of fetal- and maternal-cell-specific gene sets reported in Vento-Tormo et al. [55] in our scRNA dataset (Fig. 1b) to determine the cell origin of the mixed-origin cells (Fig. 1c and Additional file 1: Fig. S2a–l). All cell lineages described in Fig. 1b are present in the majority of the samples in this study (Additional file 1: Fig. S1e and Additional file 3: Table S2). Cells from control and PE placentas were well intermingled in each population (Fig. 1d and Additional file 3: Table S2). As expected, the iTSC population cultured in vitro [58, 59] was primarily colocalized with trophoblast progenitors (VCTp, Fig. 1d). Compared with control placentas, PE placentas were depleted in terminally differentiated SCT and EVT, were relatively enriched in VCT, and exhibited an increased count of stromal cells (Additional file 1: Fig. S1d, e). We integrated scATAC data with scRNA data (Methods) to use scRNA labels as a guide for annotating these cells (Fig. 1e). Cell clustering and annotation were validated by analyzing chromatin accessibility around known marker genes (Fig. 1g). In general, scATAC-classified cell types showed consistent chromatin opening on known marker gene promoters (Additional file 1: Fig. S3a, b).

To identify cell types that are associated with PE pathogenesis, we first compared cell type composition between PE and control donors. PE placentas exhibited a disproportionate composition of cell types, specifically trophoblasts, endothelial cells, and immune cells (Additional file 1: Fig. S1d). Using Scissor [62], we performed RNA expression similarity correlation analysis against public RNA-seq dataset of PE and control placentas [63, 64] (Fig. 1h inset, Additional file 1: Fig. S4 and S5) to search for scRNA cells whose expression profiles were associated with a specific phenotype. Orthogonally, we evaluated the association between single cells and the hereditary risk of PE by performing an enrichment analysis of PE-associated GWAS loci across scATAC profiles using SCAVENGE [65]. In the scRNA-based Scissor analysis, the PE phenotype was associated with endothelial cells, trophoblasts, and monocytes, whereas most other immune cells were associated with normal pregnancy (Fig. 1h, inset). In contrast, in the scATAC-based SCAVENGE analysis, hereditary risk loci for PE were associated with endothelial cells, trophoblasts, and macrophages (Fig. 1h, inset). Together, the two results identified EVT and endothelial cells were most significantly associated with the PE phenotype in terms of both hereditary risk and phenotypic association (Fig. 1h), indicating that they were likely the key cell types underlying PE pathogenesis.

PE-associated trophoblast, including VCTp, VCT, and SCT, were predominantly enriched in PE placentas (Additional file 1: Fig. S4a, b). The relatively higher enrichment of hereditary PE risk loci in VCT/VCTp than phenotypic enrichment suggests that PE risk loci might affect the differentiation of these cells rather than their function (Fig. 1h).

Gong *et al. Genome Biology*   (2024) 25:117

Page 6 of 46

Developmental delay of PE trophoblast

To investigate the role of trophoblasts in PE, we used RNA velocity to time the developmental age of trophoblasts in control and PE. A total of 13,301 trophoblast cells were identified and classified into four distinct types: VCTp, characterized by high proliferation and division activity (Additional file 1: Fig. S6a, b), VCT, SCT, and EVT, each defined by canonical markers (Fig. 2a and Additional file 1: S7a-b). The "developmental age" of trophoblasts was determined by RNA velocity in terms of latent time [66] and clustered using a bimodal Gaussian mixture model [67] (Fig. 2b). Control trophoblast cells (indicated as gray in Fig. 2b) were classified into two populations: a "juvenile" population, exhibiting a mean latent time of 0.15, and an "adult" population, exhibiting a mean latent time of 0.56 (Fig. 2b). However, PE trophoblast cells predominantly exhibited an intermediate phenotype, i.e., a phenotype between the juvenile and adult stages (Fig. 2b). Using CytoTRACE [68], we assessed trophoblast differentiation and identified significant maturation delay in PE undifferentiated progenitors (VCTp-4) and terminally differentiated trophoblasts (VCT-3, SCT-6/12, and EVT-9/13) (Fig. 2c and Additional file 4: Table S3). We performed an orthogonal analysis using *EpiTrace* [69] to infer trophoblast "mitosis aging rate" (age increase per gestational week) from scATAC data, showing that PE trophoblast cells exhibited a significantly decreased proliferation rate in each lineage (Fig. 2d).

By correlating the latent time with CytoTRACE scores, we determined that the juvenile cells correspond to the immaturely differentiated while adult cells are maturely differentiated (Fig. 2e, f). Majority of trophoblasts were terminally differentiated in control placentas; however, there was a significant increase of immature trophoblasts and decrease of mature ones (Fig. 2e, f and Additional file 1: Fig. S8a). Cell cycle scoring based on scRNA data showed that actively dividing cells (G2M/S phase) were slightly increased in terminally differentiated trophoblasts, i.e., VCT-3/SCT-12/EVT-9 (Additional file 1: Fig. S8b), in PE.

In PE, excessive immature trophoblasts exist in the placenta (Fig. 2g). This phenomenon is not solely explained by earlier delivery of PE pregnancies, because the "adult" population of trophoblast emerge as early as 6–12 gestational week (gw) in control pregnancies (Additional file 1: Fig. S9a and S10). Trophoblasts in control pregnancies could be clearly separated into juvenile or adult population and the majority of juvenile populations in 3rd control placentas enrich in fetal sides (Additional file 1: Fig. S9b). Trophoblast on maternal side of control placentas show greater maturity than those on fetal sides (Additional file 1: Fig. S9b). However, in all PE pregnancies, trophoblasts were developmentally stalled at an intermediate stage (Additional file 1: Fig. S9c and Additional file 1: Fig. S10). In the comparison groups, 29gw control vs 32gw PE/30gw control vs 35gw PE/ 38 gw control vs 36 gw PE, though control placentas are at earlier gestational week than PE, they exhibit a higher proportion of adult populations (Additional file 1: Fig. S9b-c and Additional file 1: Fig. S10).

On the basis of the mitosis rate, mitosis activity, and developmental maturity data, we concluded that trophoblast development is slowed in PE, especially in terminally differentiated SCT and EVT. Validation with an external dataset GSE173193 [70] showed similar trophoblast developmental delay in PE by CytoTRACE analysis (Additional file 1: Fig. S11).
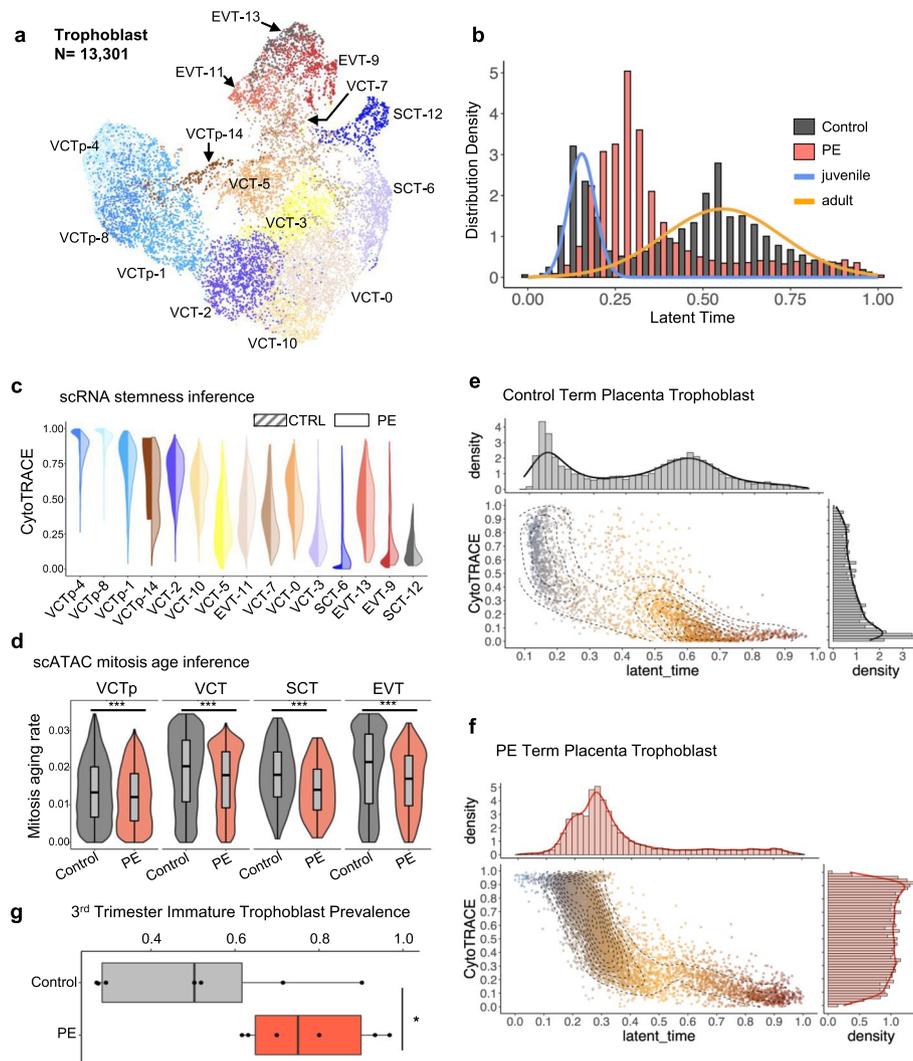
Fig. 2 Developmental delay in PE trophoblasts. **a** UMAP projection of trophoblasts in scRNA. **b** RNA velocity derived latent time distribution for PE (pink) or control (gray) trophoblasts. Control trophoblasts were segregated into a juvenile (blue) population and an adult population (orange), while PE trophoblasts were stuck in the middle. **c** Stemness prediction by CytoTRACE of scRNA clusters of control (solid) and PE (transparent) trophoblasts. Significant delayed maturation was found for PE EVT and SCT. **d** Mitosis aging rate (age increase per gestational week) inferred by *EpiTrace* of scATAC clusters of control (gray) and PE (pink) trophoblasts, indicating significantly reduced cell proliferation/division in PE trophoblasts. *P*-values were tested by Wilcox test, raw *P*-value. **e** Scatter plot of stemness prediction (*y*-axis) and latent time of control trophoblasts, showing that the juvenile population is immature, and the developed population is mature. Distribution of latent time and CytoTRACE score were shown at the *x*- and *y*-axis, respectively. **f** In PE, developmentally delayed trophoblasts were found to be immature. **g** The frequency of immature trophoblast in termed placentas is significantly increased in PE compared to control ($P = 0.03$, *P*-value was tested by *t*-test, raw *P*-value). Immature trophoblasts were defined as CytoTRACE score < 0.2 and Latent Time value < 0.4, based on a Gaussian mixture model segregation

We then delineated the developmental trajectory of trophoblast cells by RNA velocity [66] and subsequently determined three major developmental trajectories culminating in VCT-3, SCT-12, and EVT-13/9, respectively (Additional file 1: Fig. S12a-b). In PE trophoblast, the trajectory separated from control since early stage

Gong *et al. Genome Biology*   (2024) 25:117

Page 8 of 46

(Additional file 1: Fig. S12a-b), leading to a significant reduction in terminally differentiated cells (EVT-13/SCT-12) in PE (Additional file 1: Fig. S8a). These observations were consistent with orthogonal analysis with trajectories generated using Slingshot [71] ("Methods") on diffusion map projection (Additional file 1: Fig. S12c-d). Together, these results indicate a developmental trajectory switch in PE trophoblasts, characterized by early-onset developmental delay, resulting in impaired maturation of terminally differentiated EVT and SCT.

Early transcriptional defect in trophoblast development in PE

To determine the molecular mechanism underlying trophoblast developmental delay in PE, we analyzed transcriptional network among "velocity genes" identified by RNA velocity. Six master transcriptional binding site motifs regulating trophoblast developmental velocity genes were identified, including ZGA-related *NFYA*/*NFYB*/*NFYC* [46, 72], the polycomb repressor complex 2 (PRC2) complex members *EZH2* [73] and *YY1* [74], and a homeobox TF, *PBX1* (Fig. 3a). By examining TF binding sites in each TF's promoter, we found that NFYB served as the foremost upstream regulator, binding to the promoters of all the other five master TFs. NFYA/NFYC/PBX1 functioned as intermediate layer regulators, regulating both themselves and EZH2/YY1, which were determined to be downstream, bottom layer regulators (Fig. 3b).

The enrichment analysis of TF binding sites in trophoblast cell type-specific velocity genes reveals that majority of VCTp-expressed velocity genes were controlled by NFYB (Fig. 3c), whereas those in VCT were regulated by NFYA/C and PBX1 and velocity genes in terminally differentiated SCT/EVT were exclusively controlled by PRC2 (Fig. 3c). Thus, trophoblast development was modulated by a cascade of master TFs. Despite no changes in RNA expression of master TFs (Additional file 1: Fig. S13a), we observed altered TF activities in PE with scATAC data: decreased NFYB/PBX1 activities in VCTp and VCT clusters; decreased NFYA/NFYC/YY1 activities in VCT clusters (Fig. 3d and Additional file 1: S13b). Due to a technical reason, we were unable to directly infer EZH2 activity in the scATAC dataset together with other TFs. However, the activities of two of its upstream master regulators E2F4 and E2F7 were downregulated in PE trophoblasts (Additional file 1: Fig. S14). Changes in activities of TFs related to trophoblast differentiation, including GCM1, TEAD4, and TFAP2C, were also detected (Fig. 3d and Additional file 1: Fig. S13b).

We further validated the relative expression level of gene sets under control of the master TFs with trophoblast in GSE173193 dataset [70]. We found that for all but one (NFYA gene set in SCT) cases, master TFs controlled genes were expressed in lower levels in PE cells compared to control cells (Additional file 1: Fig. S15), confirming that the master TFs-regulated gene network is disrupted in PE trophoblasts.

A complementary assay was conducted to profile the transcriptional regulation through bulk ATAC sequencing of additional frozen placenta samples for validation ("Methods," Additional file 1: Fig. S16 and Additional file 5: Table S4).
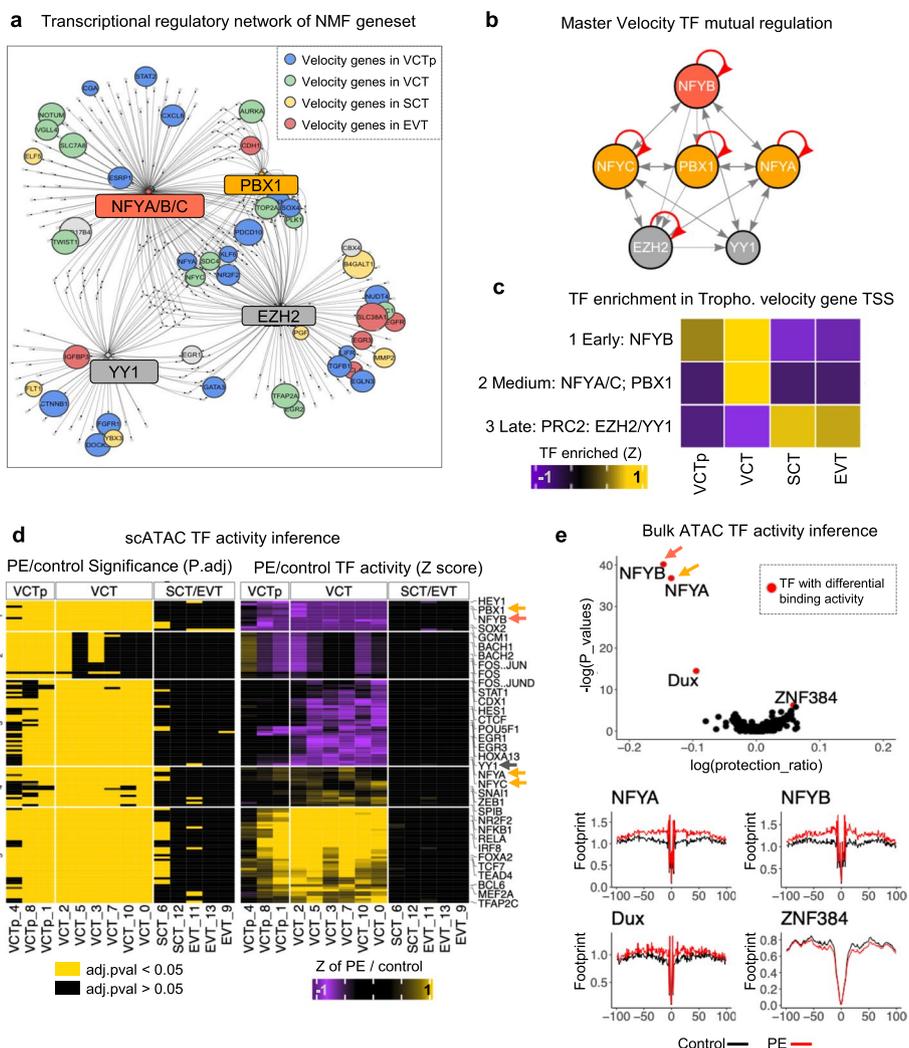
Fig. 3 Early transcriptional defect during trophoblast development in PE. **a** The promoters of developmental switch genes ("velocity gene") during trophoblast development predicted by RNA velocity are enriched with six core transcription factors in this gene set: NFYA, NFYB, NFYC, PBX1, EZH2, and YY1. **b** Transcriptional regulation network between the core transcription factors shows three tiers of regulation with NFYB controlling middle layer NFYA/C and PBX1, which in turns controls EZH2 and YY1. **c** TF binding activity, inferred by scATAC, of core transcription factors on the promoter of velocity genes, showing that NFYB is most active in VCT progenitor, whereas NFYA/C and PBX1 are most active in VCT, and EZH2/YY1 are most active in terminally developed EVT/SCT. **d** Heatmap of differential TF binding activity between PE and control in each trophoblast clusters. Left: Statistical significance (yellow: significant; black: non-significant), *P*-values were tested by Wilcox test and adjusted with Holm method; Right: *Z*-score of TF binding activity. Differential TF binding is most evident in early-stage VCTp and intermediate-stage VCT but not the terminally developed trophoblasts (SCT/EVT), suggesting an early developmental defect primes the pathogenic phenotype in terminally differentiated trophoblasts in PE. **e** Differential TF binding activity (protection ratio) between PE and control inferred by bulk ATAC-seq, showing significantly downregulated ZGA-associated transcription factors NFYA/B and DUX activity, and elevated ZNF384 activity in PE placenta. Bottom: averaged genome-wide chromatin accessibility profile around NFYA, NFYB, DUX and ZNF384 TFBS

From bulk ATAC data, we inferred TF binding activity (protection score) by comparing control and PE placentas. In line with scATAC findings, ZGA-active TF, including DUX [44], NFYA/NFYB [46, 72], showed significantly decreased binding activity in PE

(Fig. 3e). Collectively, these results collaboratively revealed a disrupted core transcription network in PE, participating trophoblast lineage determination is disrupted in PE.

Defective de novo DNA methylation on PRC2-controlled regulatory loci results in imbalanced imprinted gene expression in placenta

Transcription network dysregulation during early trophoblast development in PE suggested that either the expression or the function of the dysregulated master TFs were defective in PE. The expression of master TFs, however, did not differ between PE and control trophoblasts (Additional file 1: Fig. S13a). Notably, both the NFY family and PRC2 are closely associated with DNA methylation; the NFY family of TFs binds to methylated retrotransposon sequences [45] and the PRC2 complex binds to methylated CpG islands [73, 75]. Therefore, we hypothesized that, in PE, the dysregulation in the activity of master TFs in trophoblast may be due to defective DNA methylation. We then analyzed the epigenomic profiles of PE and control placentas to determine the mechanism underlying early transcriptional dysregulation.

The placenta develops from extraembryonic tissue derived from the trophectoderm in the blastocyst while undergoing extensive epigenetic reprogramming, including de novo DNA methylation. By profiling whole-genome DNA methylation patterns in PE and control placentas, we identified a total of 60,515 differentially methylated loci in PE, including 2710 hypomethylated (PE-hypo) continuously differentially methylated regions (DMRs) and 1271 hypermethylated (PE-hyper) continuously DMRs (Additional file 6: Table S5 and Additional file 1: Fig. S17a-b). PE-hypo DMRs were significantly hypomethylated in the maternal side, but not the fetal side, of PE placentas (Additional file 1: Fig. S18a). In contrast, hypermethylation of PE-hyper DMR was more significant in the fetal side of PE placenta (Additional file 1: Fig. S18b). These results suggest that DNA hypomethylation sites are specific for PE differentiated trophoblasts, whereas DNA hypermethylation sites are associated with all PE trophoblasts.

By averaging the DNA methylation level per DMR, we projected DNAm-seq data pertaining to sequenced control and PE placentas together with a public single-cell WGBS data set pertaining to early embryonic development [42] as a UMAP. PE placentas stalled along the trajectory of trophoectoderm development toward trophoblast cells, indicating that DNA methylation in PE placentas was delayed (Fig. 4a). Trophoectoderm developmental trajectory based on chromatin accessibility at all DMRs in placental tissues and single cells of early embryonic developmental stages from publicly available data set [45, 76] suggested a similar scenario, i.e., PE placentas were situated midway along the trajectory from trophectoderm toward trophoblast (Fig. 4b). Together, these results indicate that the epigenome reprogramming during extraembryonic development is delayed in the PE placenta.

We defined open DMRs as regions with trophoblast ATAC peak and differential DNA methylation. Open DMRs were significantly enriched only in placenta-accelerated genomic regions [77] but not human-specific accelerated regions [78] and vertebrate-specific ultraconserved noncoding elements [79] (Fig. 4c), suggesting the functional relevance of DMRs in placenta development. We then annotated PE-specific DMRs using chromVAR [80]. Compared with PE-hyper DMRs, PE-hypo DMRs were enriched

in PRC2 complex-related TF binding sites, including EED/EZH2/JARID2/SUZ12/YY1 (Fig. 4d), implying that defective DNA methylation in PE placenta affects PRC2 complex binding and function. Furthermore, PE-hypo DMRs covered loci that underwent de novo DNA methylation during epigenetic reprogramming in extraembryonic tissues [50], as these loci were de novo methylated during extraembryonic tissue development (Additional file 1: Fig. S17b) and were extensively methylated in cancer tissue (Fig. 4e). DNA methylation on PE-hypo DMRs was severely aberrant in PE placentas (Fig. 4e), indicating a specific defect in extraembryonic tissue-specific de novo methylation. Conversely, PE-hyper DMRs (PE-Hyper cluster 4) were enriched in retrotransposon sequences (Fig. 4d). Typical hypomethylation in maternally imprinted retrotransposons, particularly those of the long terminal repeat 12 (LTR12) family of endogenous retroviruses (ERVs), was disrupted in PE placentas (Fig. 4d and Additional file 1: Fig. S19a). These maternally imprinted and paternally hypomethylated LTR12C elements were transiently active during ZGA (Additional file 1: Fig. S19b), possibly owing to the binding of NFYA/B/C in their terminal repeat sequences [45]. In fact, hypomethylated LTR12C in human sperm are more likely to be hypermethylated in PE (Additional file 1:

(See figure on next page.)

Fig. 4 De novo DNA methylation defect on PRC2-controlled placenta regulatory loci results in imbalanced imprinted gene expression. **a** Cell evolution trajectory from zygote to placenta inferred by genome-wide DNA methylation sequencing data, showing delayed development of PE placenta in terms of DNA methylation. **b** Cell evolution trajectory from zygote to placenta inferred by chromatin accessibility on PE-specific differentially methylated region (DMR), showing delayed development of PE placenta in terms of chromatin accessibility on PE DMR. **c** Enrichment of opened (accessible) PE DMR in different types of genomic elements: ultraconserved noncoding elements (UCNE), placenta-accelerated genomic region (PAR), and human-specific accelerated region (HAR). Opened PE DMR is enriched in PAR but not HAR or UCNE, suggesting that they are functionally relevant to placental development, as the PAR regions undergone Darwinian positive selection in placental animals. *P*-value and odd's ratio (OR) were calculated by Fisher's exact test. Inset: schematic diagram of animal evolution and the time point where genomic regions (UCNE/PAR/HAR) were subjected to evolutionary selection. **d** Enrichment heatmap of TFBS and chromVAR-annotated genomic regions in PE DMR, showing that PRC2 complex-related genomic regions (H3K27me3 in trophectoderm and ESC, and the PRC2 complex components SUZ12, EZH2, JARID2, EED) were more likely to be subjected to DNA hypomethylation instead of hypermethylation in PE. On the other hand, NFYB binding sites and LTR12C elements are more likely to be hypermethylated in PE. **e** Sample-wise CpG methylation levels (box-and-whisker plot with mean, 0%-25%-75%-100% quantiles) of conserved PE-hypomethylated loci (top), human-specific PE-hypomethylated loci (middle), and PE-hypermethylated loci (bottom) of normal tissue (light blue, left), cancer tissue (dark blue, left), and placenta of normal pregnancy (dark green, right) and preeclampsia (light green, right) showing that PE-hypomethylated loci undergone similar hypermethylation in both placenta and cancer, suggesting these regions are the ones subjected to de novo DNA methylation during extraembryonic tissue development. In contrast, DNAm levels on PE-hypermethylated regions do not segregate cancer from non-malignant tissues, suggesting that they are under control of a different molecular mechanism. **f** Significant enrichment of all PE DMR on imprinted genes and PE-hypo DMR on h1 hESC EZH2 binding sites ($P = 0.014$, $Z = 3.4795$, PE DMR vs imprinted genes; $P = 0.0099$, $Z = 26.9084$, PE-hypo DMR x EZH2; Z test. Basal distribution was done by permutation performed × 1000 times. Expected number of overlap: black vertical line; alpha = 0.05 number of overlap: red vertical line; observed number of overlap: blue vertical line) on imprinted genes, indicating that imprinted genes were affected by DNA methylation defects in PE. **g** Differential expression of imprinted genes between PE and control single trophoblasts. Blue: maternal allele expressed (paternally imprinted); Pink: paternal allele expressed (maternally imprinted). *P*-values were calculated by Wilcox test and adjusted based on Bonferroni correction. **h** Log fold-change of imprinted gene expression between PE and control trophoblasts, showing that expressed maternal alleles (pink) is likely to be downregulated, while expressed paternal alleles (blue) is likely to be upregulated in PE. **i** Frequency of imprinted genes that has shown PE-specific upregulation (red: PE-enhanced) or downregulation (white: PE-attenuated) in trophoblasts, showing dysregulated imprinted gene expression in differentiated trophoblasts, while the paternal alleles were significantly more likely to be upregulated in SCT ($P = 0.033$, Fisher's exact test)
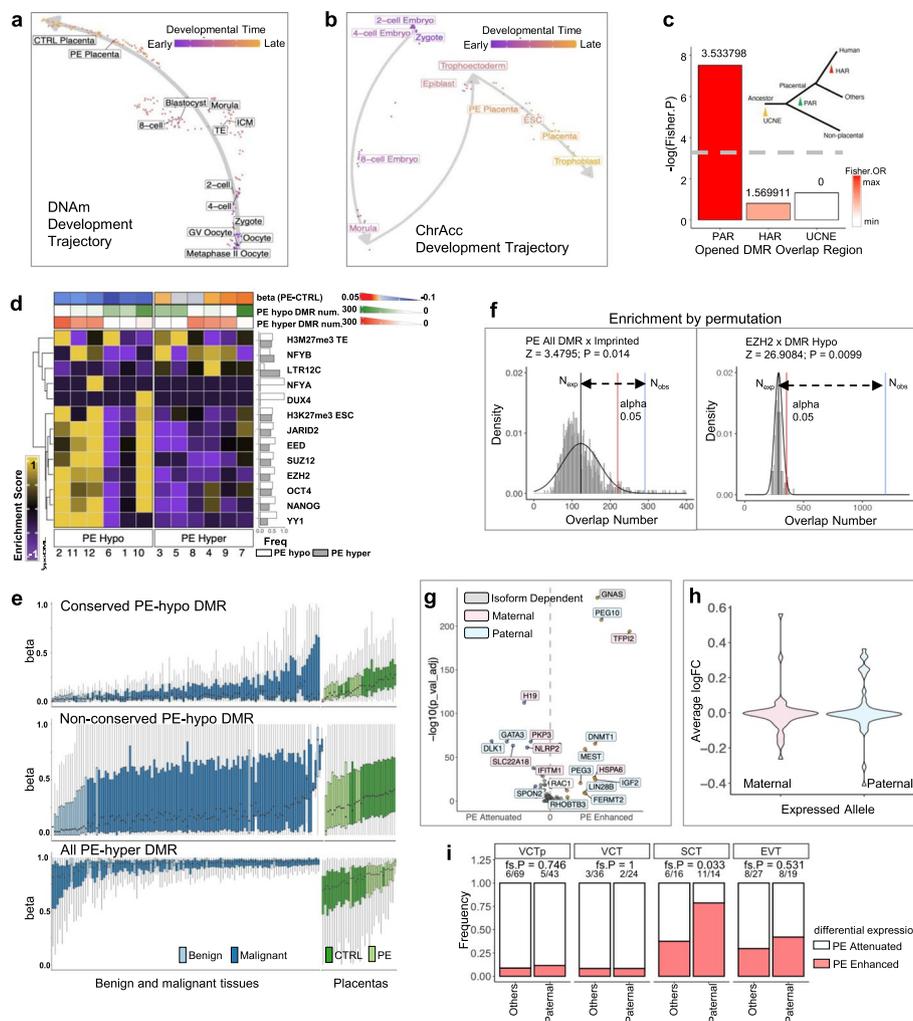
Fig. 4 (See legend on previous page.)

Fig. S20), suggesting that the establishment or post-zygotic maintenance/refinement of DNA methylation imprinting may be disrupted in PE.

Imprinting is modulated by DNA methylation and H3K27me3 modifications [81–83]. In PE, DMRs were significantly enriched in imprinted genes and binding sites of PRC2-related genes, such as *EZH2* (Fig. 4f and Additional file 1: Fig. S21), suggesting that de novo DNA methylation may further regulate imprinted gene expression in addition to the DNA methylation profile established during germline development. If such a case, we expect dysregulated imprinted gene expression in the PE placenta. This was validated by analyzing RNA expression in the scRNA dataset, which showed significant differential expression of imprinted genes in PE trophoblast (Fig. 4g, h and Additional file 7: Table S6). Notably, maternally imprinted genes were more likely to be upregulated in PE, whereas paternally imprinted (maternal allele expressed) genes were more likely to be downregulated in PE (Fig. 4g, h and Additional file 7: Table S6). The overexpression of imprinted genes was mostly exhibited by terminally differentiated SCT and EVT (Fig. 4i). In SCT, significantly more paternal allele-specific genes were overexpressed in PE (Fig. 4i, Additional file 1: Fig. S22 and Additional file 7: Table S6).

Imprinting could also be regulated by H3K27me3 modification. We additionally profiled H3K27me3 in PE and control placenta using CUT&Tag assay. In PE placentas, we identified 32,409 genomic regions with increased H3K27me3 modification (PE-gain), and 13,284 genomic regions with decreased H3K27me3 modification (PE-lost), compared to control placentas (Additional file 1: Fig. S23a-c). The PE-lost regions are highly enriched around EZH2 binding site (Additional file 1: Fig. S23d). Overall, EZH2 loci are associated with decreased H3K27me3 modification in PE placenta (Fig. S23c). Many of these PE-lost, EZH2-bound regions were located on imprinted genes (Additional file 1: Fig. S24a). Compared to the maternally imprinted loci (expressing the paternal allele), paternally imprinted loci (expressing the maternal allele) are more likely to lost H3K27me3 modification in PE placentas (Additional file 1: Fig. S24b). Together, our results indicate that both molecular mechanisms regulating imprinted gene expression, namely H3K27-trimethylation and extraembryonic tissue-specific de novo DNA methylation, are defective in PE.

Excessive maternally imprinted gene product PEG10 in PE

We performed differential gene expression analysis in the trophoblast scRNA dataset. PE trophoblasts overexpressed *PEG10*, *GADD45A*, *CCND1*, *SIGLEC6*, and *H2AZ1* (Fig. 5a and Additional file 7: Table S6). On the other hand, the lncRNA *MALAT1* and the canonical mature trophoblast marker genes *CGA*, *HLA-G*, and *PGF* are downregulated in PE trophoblasts (Fig. 5a and Additional file 7: Table S6). GSEA enrichment of differentially expressed genes showed that PE trophoblasts exhibited increased gene set activities in "G2M checkpoint," "E2F targets," and "MYC targets" pathways (Additional file 1: Fig. S25a). These results are in concordance to the function of PE upregulated genes such as *CCND1* (an important factor that regulates cell cycle) and *GADD45A* (an important factor that regulates chromatin accessibility and cell cycle). The results suggest a scenario that in PE trophoblast, *GADD45A* expression is induced to arrest the cells at G2/M checkpoint from further division and differentiation, which is supported by cell cycle gene set expression scoring on single-cell RNA-seq dataset (Additional file 1: Fig. S25b).

Notably, we found an imprinted gene *PEG10* was significantly upregulated in all trophoblast cells across various trophoblast lineages in PE (Fig. 5a and Additional file 1: Fig. S26). *PEG10* was the most highly upregulated maternally imprinted gene in PE (Fig. 4g, Additional file 1: Fig. S22 and Additional file 7: Table S6). *PEG10*, which is a retrotransposon gene involved in trophoblast stem cell specification [84], functionally implicated in placenta development and embryogenesis [85].

In control samples, *PEG10* expression gradually decreased during trophoblast differentiation. In detail, PEG10 expression was high in VCTps and VCTs, but low in SCTs and EVTs (Fig. 5b, top panel). In contrast, in PE, the decrease in *PEG10* expression was delayed, with significantly elevated expression in SCTs and EVTs both in our in-house sequenced trophoblast and external dataset GSE173193 [70] (Fig. 5b, top panel, Additional file 1: Fig. S11g and h). In the paired sIUGR placentas, the trend of *PEG10* expression in the trophoblasts in normal fetus placenta was similar to control, whereas the FGR fetus placenta trophoblast phenocopied PE placenta (Fig. 5b, bottom panel). Since
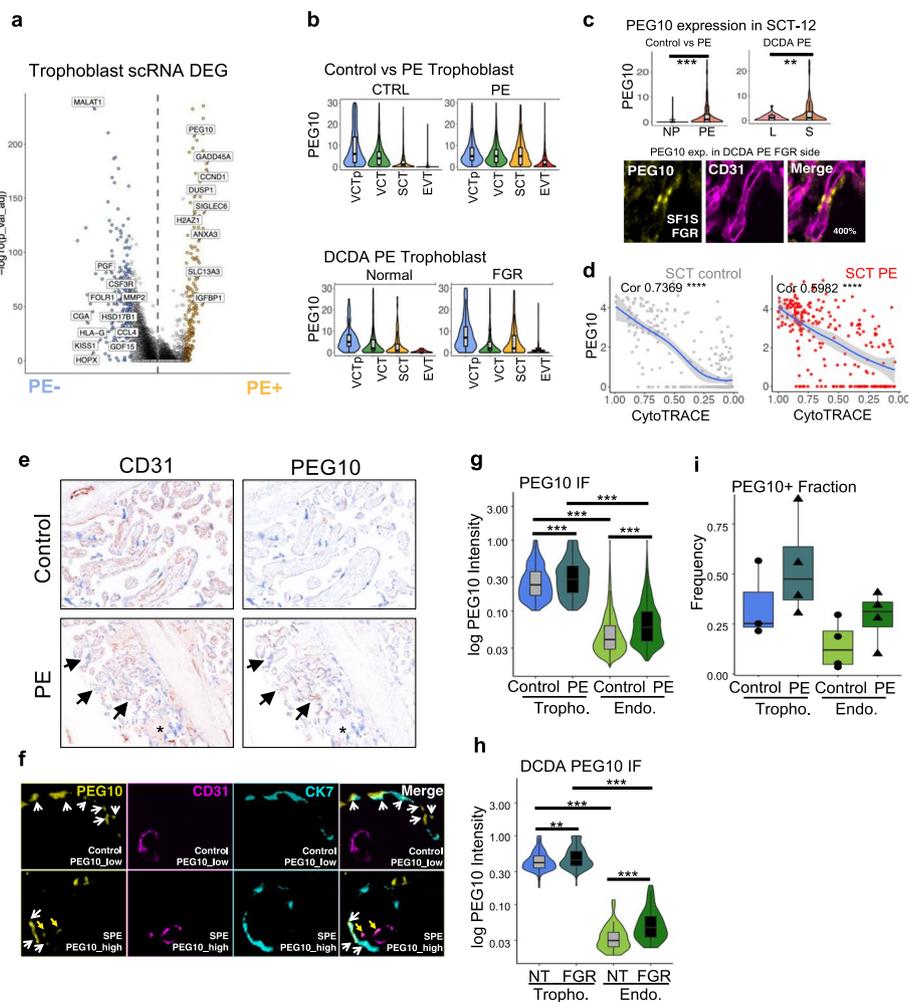
Gong *et al. Genome Biology*   (2024) 25:117

Page 14 of 46

Fig. 5 Excessive maternal imprinted gene product PEG10 in PE. **a** Differential expression of genes between PE and control trophoblasts, showing PEG10 among the most significantly upregulated genes in PE trophoblasts. *P*-values were calculated by Wilcox test and adjusted based on Bonferroni correction. **b** Expression of PEG10 in control, PE, and normal/FGR placentas from DCDA sIUGR twins, showing that PEG10 expression is gradually downregulated as trophoblast differentiate but this trend is blocked in PE or FGR placenta. **c** Significantly upregulated PEG10 expression in terminal SCT (SCT_12) cluster in PE or sIUGR FGR placenta (S) compared to their controls (L). *P*-value was determined by *t*-test, **: $P < 0.01$, ***: $P < 0.001$. Staining of PEG10 (yellow) is found in SCT lining the blood vessel (CD31+, purple) but we noticed a faint staining of PEG10 in CD31+ endothelial compartment in this FGR placenta sample. **d** Single-cell PEG10 expression (*y*-axis) and cell maturation (1-CytoTRACE, *x* axis) in SCT of PE and control placenta. PEG10 expression is significantly correlated with cell maturity in an inverse manner (Cor.coef: 0.7369 (control)/0.5982 (PE); *P*-value: both $< 2.2*1e-16$), indicating that PEG10 overexpression in PE is due to developmental delay of trophoblasts. **e** OPAL multiplex immunohistochemistry staining (brown: IHC stain; blue: DAPI) of CD31 (left) and PEG10 (right) in control (top) and PE (bottom) placentas, showing villous void of (asterisk) or with narrow, deformed fetal capillary vessel (arrows) in PE, together with excessive PEG10 expression. **f** OPAL multiplex immunohistochemistry staining of PEG10 (yellow), CD31 (purple), CK7 (cyan) in control (top) and PE (bottom) placenta, showing that PEG10 is found not only in non-CD31 trophoblast (CK7+) cells in control but is additionally found in endothelial (CD31+) cells in PE. **g** Immunofluorescence intensity of PEG10 in trophoblasts (T) or endothelial cells in normal (control) or PE placenta, showing significantly elevated PEG10 protein level in PE trophoblasts as well as endothelial cells. *P*-values were tested by Wilcox test, not adjusted. **: $P < 0.01$, ***: $P < 0.001$. **h** Immunofluorescence intensity of PEG10 in trophoblasts or endothelial cells from the paired normal (NT) and FGR placentas from the sIUGR pregnancy, showing significantly elevated PEG10 protein level in FGR cells. *P*-values were tested by Wilcox test, not adjusted. **i** Frequency of trophoblasts or endothelial cells that are positive with PEG10, in control and PE placentas, showing increased PEG10-positive cells in abnormal placenta. *Abbreviation*: DCDA: Dichorionic diamniotic twin pregnancy; sIUGR: selective intrauterine growth retardation. NT: normal fetus; FGR: fetal growth restriction; SPE: severe preeclampsia

sIUGR placentas share a similar maternal environment, these results indicate that cell-autonomous transcriptional regulation underlies *PEG10* expression dynamics.

PEG10 overexpression was most evident in terminally differentiated SCT-12 in either PE or sIUGR/FGR placenta (Fig. 5c). Immunohistochemistry validated the protein expression of PEG10 in perivascular SCTs, and weak but consistent PEG10 staining was observed within the CD31 + endothelial cells (Fig. 5c). Correlation analysis between the differentiation status deduced by CytoTRACE and *PEG10* expression levels in SCTs suggested that *PEG10* expression was a function of differentiation status in both control and PE placenta (Fig. 5d). Hence, the delayed maturation of trophoblasts, most evidently in SCTs, underlies *PEG10* overexpression.

Through OPAL multiplex immunohistochemistry, we determined the protein expression of *PEG10* in singleton PE, control, and DCDA PE placentas (Fig. 5e, f). The PE placenta exhibited an increased degree of PEG10 staining in villous that was void or contained deformed and narrow fetal capillaries (Fig. 5e). Consistent with the results of scRNA, PEG10 immunofluorescent intensity in trophoblasts (CK7 +), and the fraction of PEG10 + trophoblast were significantly increased in PE (Fig. 5e–i) and sIUGR placentas (Fig. 5h). In addition to PE-specific PEG10 overexpression in trophoblast, we observed a significant increase in PEG10 expression in CD31 + endothelial cells, in terms of both staining intensity (Fig. 5f–h) and cell count (Fig. 5i). These results indicate that trophoblasts exhibiting developmental delay in FGR (sIUGR) and PE placenta overexpress *PEG10.*

PEG10 is transferred from the trophoblast to maternal endothelial cells

PEG10 is an ancient retroviral Gag gene encoded by a domesticated endoretrovirus (ERV). It binds to 5′ and 3′ UTR of its transcript, leading to the formation of virus-like particles (VLPs) [86]. These VLPs are secreted via the exosome pathway and transferred into other cells. PEG10 protein and RNA transcripts levels were increased in PE trophoblasts and non-trophoblast cells, including endothelial cells (Fig. 5a, e–h and 6b, e). However, scATAC data showed that the *PEG10* locus was only accessible in trophoblast cells and inaccessible in all other cells (Fig. 6a and Additional file 1: Fig. S27a). The *PEG10* locus was differentially methylated in the gamete germline (termed germline differentially methylated region, gDMR [87], Additional file 1: Fig. S27a). We found that the *PEG10* gDMR was bound by the PRC2 complex component EZH2 and that it exhibited specific DNA hypomethylation and increased H3K27ac modification in the PE placenta (Fig. 6a and Additional file 1: Fig. S27b), indicating either a loss-of-imprinting or loss-of-de-novo-methylation around the region, resulting in reduced PRC2 binding and dysregulated expression of *PEG10* during trophoblast lineage differentiation in PE.

We further analyzed the structure of *PEG10* RNA transcript across various cell types. Compared with those in control, single cells in PE exhibited increased expression of *PEG10* RNA transcripts (Fig. 6b). In fetal trophoblast, the sequenced RNA reads were biased toward the 5′ UTR, suggesting active ongoing transcription (Fig. 6b). On the contrary, in other cells, including fetal and maternal endothelial cells and maternal immune cells, the sequenced RNA reads were biased toward the 3′ UTR. The fraction of reads corresponding to the 3′ UTR was increased in non-trophoblasts compared to trophoblasts (Fig. 6b, right panel), suggesting more mature transcripts in non-trophoblasts.
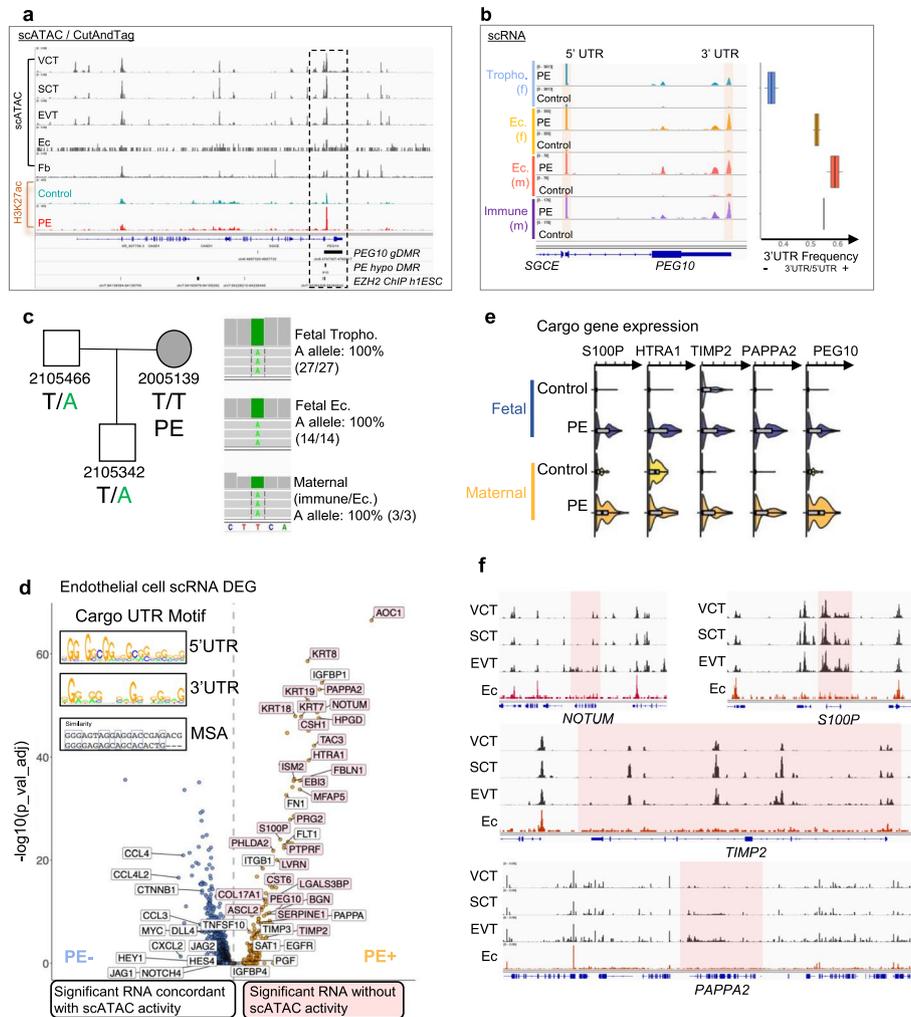
Fig. 6 PEG10 is transferred from trophoblast to maternal endothelial cells. **a** Chromatin accessibility of VCT, SCT, EVT, endothelial cells (Ec) and fibroblasts (Fb), and H3K27ac Cut-and-Tag signals (Additional file 1: Fig. S27) from control and PE placentas, around the PEG10/SGCE locus. Compared to trophoblasts, Ec and Fb have little chromatin accessibility on PEG10 locus, despite the significant upregulation of PEG10 transcripts and proteins in PE Ec. The PEG10 locus is subjected to imprinting (gDMR), controlled by EZH2, hypomethylated in PE, and shows PE-specific increase of H3K27ac modification (Additional file 1: Fig. S27), indicating PE-specific epigenomic activation of PEG10 locus. **b** scRNA sequencing reads from fetal trophoblasts, fetal endothelial cells, maternal endothelial cells, and maternal immune lymphoid (B/NK/T) cells on PEG10 genomic region. PE cells show significantly higher PEG10 expression. RNA reads covering PEG10 3′UTR in trophoblasts are significantly lower compared to those in non-trophoblast cells, suggesting most PEG10 transcript in non-trophoblast cell are mature and spliced. **c** A trio pedigree of a PE pregnancy, where the mother (2005139) suffered SPE during pregnancy. The father and progeny carried a paternal-specific PEG10 SNP allele (T/A) and the mother is wild-type (T/T). scRNA sequencing reads from the 2005139 placenta showing that all transcripts covering the SNP, regardless of their cell-of-origin, carried the paternal A allele, suggesting a fetal origin of these RNA in maternal cells. **d** Differential expressed genes between PE and control endothelial cells, labelled with scATAC-silent, scRNA-upregulated genes (pink). Transcripts of these genes shared common 5′UTR and 3′UTR sequence motifs with sequence similarity (inset), suggesting that they are likely to be bound with similar RNA-binding protein and might be transported from trophoblast to endothelial cells with this protein. *P*-values were calculated by Wilcox test and adjusted based on Bonferroni correction. **e** scATAC-silent, scRNA-upregulated genes in fetal and maternal endothelial cells from control and PE placenta, showing that these transcripts were similarly upregulated in fetal and maternal endothelial cells in PE. **f** Aggregated scATAC coverage tracks for trophoblasts (VCT: $N = 10299 + 47086$; SCT: $N = 230 + 3495$; EVT: $N = 1062 + 6366$, control + PE) and endothelial cells (Ec: $N = 293 + 395$, control + PE) on cargo genes *NOTUM*, *S100P*, *TIMP2*, and *PAPPA2*. Cargo gene region are highlighted in pink. Adjacent genomic regions were shown to control for scATAC validity in Ec

These results, in concordance with reported 3′ UTR-biased RNA levels in exosome [88] and the transfer of PEG10 VLPs between cells [86], suggest that *PEG10* gene product may be exported to non-trophoblast cells from trophoblasts.

In our cohort, we identified a family where the father and fetus were heterozygous for a *PEG10* SNP allele chr7:94665871 T > A (genotype T/A), whereas the PE mother was homozygous for the wild-type allele on this locus (genotype T/T) (Fig. 6c, left panel). All *PEG10* RNA reads (100%) in fetal trophoblast and endothelial cells harbored the paternal allele (A). Interestingly, *PEG10* transcripts in maternal immune and endothelial cells harbored the same paternal *PEG10* allele (A), indicating their fetal origin (Fig. 6c, right panel). Collectively, these results indicate cell-to-cell transfer of *PEG10* transcripts from trophoblast to other cells. The persistent monoallelic expression of the paternal allele indicates that DNA hypomethylation at the *PEG10* loci in PE results from defective de novo DNA methylation or H3K27me3 re-establishment, rather than germline imprinting defects in the maternal gamete.

PEG10 VLPs are known to be capable of carrying other RNA transcripts, bound by PEG10 on their UTRs, into other cells. We hence identified a series of putative PEG10 VLP "cargo genes", which were characterized by the presence of RNA without an accessible genomic locus in endothelial cells ("Methods," Fig. 6d–f and Additional file 8: Table S7). Similarities between the 3′ and 5′ UTRs of these cargo genes (0.68, $Z = 4.3$) suggested that they may be bound by a similar RNA-binding protein (Fig. 6d and Additional file 1: Fig. S28). Several of these genes, such as *HTRA1*, *TIMP2*, *ASCL2*, and *CSH1*, were known to be functional in placentation and/or could be transferred to cultured trophoblast by trophoblast debris in ex vivo culture medium [88], suggesting that cargo genes shuttle from trophoblast to both maternal and fetal endothelial cells via PEG10 VLP in PE placentas (Fig. 6d), similar to cell-free trophoblast "debris" described in a previous in vitro study [88].

We performed Pearson correlation analysis of the cargo transcript RNA expression profiles between endothelial cells and trophoblast (Additional file 1: Fig. S29) to determine whether PEG10 VLP-mediated RNA cargo transfer between cells was possible. Although we observed a high correlation between VCT/VCTp and early endothelial cell populations in control placentas, we also observed similarity between terminally differentiated trophoblast (EVT and SCT) and endothelial cells in PE (Additional file 1: Fig. S29), indicating increased cell-to-cell transfer of PEG10 VLP in PE. Collaboratively, these results suggest that PEG10-containing VLP production is amplified in PE, which results from defective de novo DNA methylation in trophoblast. The majority of transcriptional alterations in endothelial cells in PE result from the trans-cellular transfer of RNA molecules from trophoblasts, potentially carried by PEG10 VLPs.

Excessive PEG10 disrupts maternal artery development to phenocopy arteriovenous malformation (AVM)

In PE, vascular development in both the fetus and mother was disrupted. Specifically, the remodeling of maternal uterine spiral arteries was incomplete, and fetal capillary vessels were underdeveloped, thereby leading to placental villus that was devoid of blood circulation. We analyzed angiogenesis by delineating the developmental trajectory of arterial endothelial cells by RNA velocity [66] (Fig. 7a and Additional file 1: Fig. S2a). The
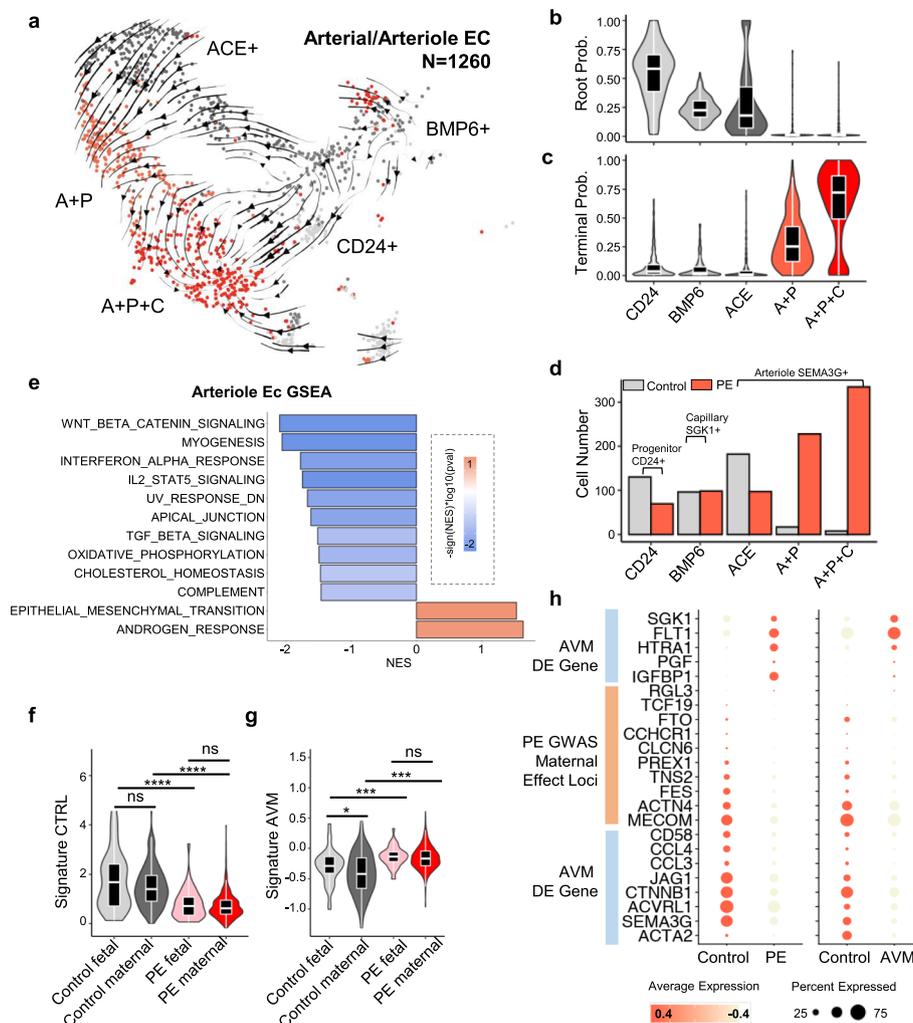
Fig. 7 Excessive PEG10 disrupts maternal artery development to phenocopy AVM. **a** Single-cell evolution trajectory inferred by RNA velocity of arterial/arteriole endothelial cells (Ec), projected on UMAP space. Arterial Ec evolve from the CD24 + cluster (Additional file 1: Fig. S2b) into BMP6 + (SGK1 + , capillary, Additional file 1: Fig.S2b) cluster or ACE + (SEMA3G + , arteriole, Additional file 1: Fig. S2b) cluster. ACE + populations. ACE + population give rise to two ACE + /PAPPA2 + populations that differs by CLIC3 expression. **b** Root and **c** terminal cell probability of single cells in each Ec cluster shows that the PAPPA2 + clusters of Ec are *induced* in PE. **d** Cell number of each Ec cluster in control and PE placenta, showing that the PAPPA2 + clusters are PE-specific. **e** Normalized enrichment score (NES, *x* axis) of gene pathways with differential activity between PE and control ACE + population in GSEA analysis. Only significantly different pathways (raw *P*-values < 0.05) were shown. Differential activities are shown with color (red: upregulation in PE; blue: downregulation in PE). **f** Non-arteriovenous malformation (AVM) brain endothelial signature (Signature control) expression in fetal and maternal-originated placental Ec in PE and control. **g** AVM signature [89] expression in fetal and maternal-originated placental Ec in PE and control. *P*-values were tested by Wilcox test, not adjusted, *: $P < 0.05$; ****: $P < 2.2*10e-6$. **h** Gene expression of AVM-associated genes, including the hereditary AVM pathogenic gene ACVRL1 which interacts and is inhibited by PEG10, and PE-associated genes by GWAS study [15] in PE/control placenta Ec (left panel) or AVM/control brain Ec (right panel). *Abbreviation*: A + P: ACE2 + , PAPPA2 + ; A + P + C: ACE2 + , PAPPA2 + , CLIC3 +

results revealed that an arterial CD24 + cluster served as a progenitor cell population with heightened cell division activity (Fig. 7b,c and Additional file 1: Fig. S30a-b) that differentiated into BMP6 + /SGK1 + capillaries ("BMP6" in Fig. 7a–d and Additional file 1: Fig. S2a-b; "Capillary" in Additional file 1: Fig. S30) or ACE + /SEMA3G + arteriole cells

("ACE," "A + P" [ACE and PAPPA2], "A + P + C' [ACE, PAPPA2, and CLIC3] in Fig. 7a–d and Additional file 1: Fig. S2a-b; "Arteriole" in Additional file 1: Fig. S30a-b). In PE, two novel arteriole PAPPA2 + endothelial cell clusters, namely arterial A + P and arterial A + P + C, emerged (Fig. 7b). RNA velocity transfer probability analysis revealed that the root cell probability was high in arterial CD24 + and BMP6 + endothelial cells (Fig. 7c) and that terminal cell probability was high in PE-specific PAPPA2 + cells (Fig. 7d). Although the cell cycle activity of the progenitor cells was higher than that of the capillary and arteriole endothelial cells in control, the cell cycle activity of capillary and arteriole endothelial cells was significantly elevated in PE (Additional file 1: Fig. S30b). Together, these results suggest the occurrence of induced differentiation of endothelial cells into a PAPPA2 + state in PE.

PEG10 protein product RF2 inhibited the TGF-beta co-receptor ACVRL1 (ALK1) and TGF-beta signaling in vitro [90]. In vivo, PEG10 knockout resulted in placenta development failure [85] and defective fetal vascular development [91, 92]. As TGF-beta signaling is essential for angiogenesis, and mutation in these genes results in major vascular developmental defects, such as congenital cardiovascular malformation or arteriovenous malformation [93–96] (OMIM #131300, #609192, #187300, #600376, #178600, #174900, #175050), we hypothesized that excessive PEG10 transferred to maternal endothelial cells may disrupt its development by affecting the TGF-beta signaling pathway. Indeed, gene set enrichment analysis of differentially expressed genes in the arteriole indicated significant attenuation of TGF-beta signaling, together with Wnt, IL-2, and IFN-a signaling in arteriole endothelial cells in PE (Fig. 7e and Additional file 1: Fig. S31). Meanwhile, EMT activity and androgen signaling pathway were upregulated in arteriole endothelial cells in PE (Fig. 7e).

TGF-beta signaling is not only implicated in vascular development but is also known to be modulated by trophoblast-derived molecules. For example, *MMP9/MMP2* is specifically expressed in EVT and SCT (Additional file 1: Fig. S7a) and regulates TGF-beta signaling intensity [97]. To validate TGF-beta signaling defects in endothelial cells in PE, we took advantage of the interesting duality of vascular malformation disease, namely AVM, and PE. AVM is also caused by loss-of-TGF-beta signaling [93–95]. Mutation in either endoglin (*Eng*) [12], an accessory TGF-beta receptor, or its co-receptor *ACVRL1* [98], which interacts with and is inhibited by PEG10 [90], causes AVM. *Eng* is known to be implicated in PE; excessive solutable endoglin (s-Eng) has been observed in the serum of patients with PE [21, 23, 24], and the overexpression of s-Eng has been observed to result in a PE-like phenotype in animal models [99]. Single-cell analysis of brain vascular endothelial cells has revealed similar downregulation of TGF-beta signaling in AVM arteries [89, 100]. We thus compared the gene expression profile between AVM and PE. Fetal and maternal endothelial cells exhibited an AVM-like gene expression pattern, i.e., the downregulation of control-specific arteriole genes and the upregulation of AVM-specific gene activities (Fig. 7f, g).

The mirroring of gene expression profile between PE and AVM arterial endothelial cells was found across individual AVM-specific and control-specific genes, as well as the expression of GWAS-associated genes of PE (Fig. 7h), including *MECOM*, a key endothelial TF that is significantly associated with PE [101–103]. In control placentas, the AVM-specific gene activity was lower in maternal endothelial cells than in fetal

endothelial cells (Fig. 7g). Considering that PEG10 was produced by fetal endothelial cell-proximal VCT but not the maternal endothelial cell-proximal SCT/EVT in control, these results collectively suggest that PEG10 induced TGF-beta signaling switch between maternal and fetal endothelial cells to divert differential developmental trajectories for fetal and maternal blood vessels during placentation. The placental villous tree, which serves as an arteriovenous shunt between the remodeled maternal spiral artery and draining vein [104], is innervated by fine fetal capillaries [105, 106]. Coordinated enlargement of spiral artery opening and miniaturization of fetal capillary requires differential angiogenic signaling. Although the loss of PEG10 results in the collapse of the fetal capillary vessel in villous [92], our results revealed that maternal vascular remodeling was affected by the fetal overexpression of *PEG10*.

PEG10 perturbs endothelial cell proliferation and function by negatively regulating TGF-beta signaling

We experimentally validated whether PEG10 affects endothelial cell biology via TGF-beta pathway in a human endothelial cell line (HUVEC), which has basal PEG10 expression (Additional file 1: Fig. S32). Overexpression of PEG10 increased the PEG10 RNA level up to 20-fold compared to control (Additional file 1: Fig. S32a) and elevated PEG10 protein expression (Additional file 1: Fig. S32b). Knockdown of PEG10, on the other hand, significantly downregulated PEG10 on both transcriptional (Additional file 1: Fig. S32c) and protein levels (Additional file 1: Fig. S32d).

In the MTT assay, overexpression of PEG10 inhibited, while knockdown of PEG10 accelerated HUVEC cell proliferation over a course of 4 days (Fig. 8a). In concordance to this, scRNA sequencing of the same cells showed increased cell cycle gene set activity in PEG10 knockdown cells compared to control cells or PEG10 overexpression cells (Fig. 8b). Furthermore, PEG10 overexpression decreased tube formation activity (Fig. 8c,d), while PEG10 knockdown increased the tube formation activity of HUVEC cells compared to control cells (Fig. 8e,f). PEG10 overexpression decreased the tube formed by ~ 3 folds and junction formed by ~ 2 folds (Fig. 8c,d). On the contrary, PEG10 knockdown increased the tube formed by 50%, and junction formed by ~ 25% (Fig. 8e,f).

In concordance to the cell proliferation and function experimental result, PEG10 overexpression cells showed increased expression of cell cycle arrest-related gene such as *GADD45A*, and decreased expression of proliferation genes such as *MKI67* and *TOP2A* (Fig. 9a). Notably, the TGF-beta signaling pathway components *TGFB1*, *TGFBR2*, and *SMAD3* are upregulated in PEG10 knockdown cells (Fig. 9a). Gene set enrichment analysis showed that compared to PEG10 knockdown cells, PEG10 overexpression cells showed downregulated activities in cell proliferation (G2M-checkpoint, Mitotic spindle, E2F targets), TGF-beta-signaling, and angiogenesis pathways (Fig. 9b). The expression of angiogenesis and endothelia-related genes, such as *SGK1*, *GDF15*, and *ACTN4* are disrupted by PEG10 perturbation (Additional file 1: Fig. S33), suggesting possible downstream effector of PEG10-TGF-beta signaling and mimicking the phenotype observed in vivo for PE endothelial cells.

We then tested whether treating PEG10 overexpression cells with TGF-beta-1 protein could rescue their biological function. Compared to control condition, TGF-beta-1 protein restored tube growth and junction formation in the PEG10 overexpression cells
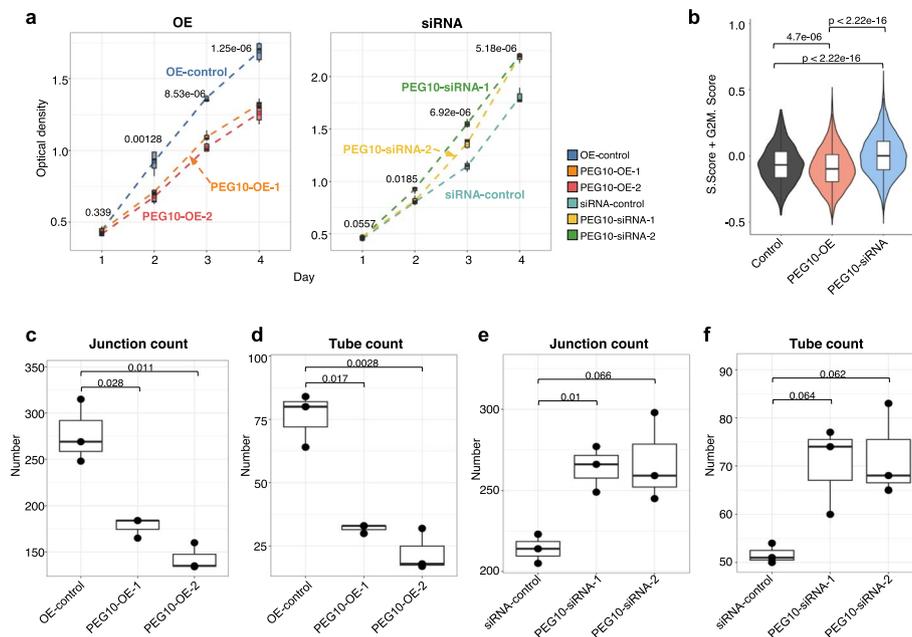
Fig. 8 PEG10 overexpression negatively impacts endothelial cell proliferation and function. **a** Optical density in MTT assay of PEG10 in HUVEC cells transfected with control vector (OE-control) or PEG10-overexpression (PEG10-OE-1, PEG10-OE-2) vectors (left), or HUVEC cells transfected with control scrambled siRNA (siRNA control) or PEG10-targeting siRNA (PEG10-siRNA-1, PEG10-siRNA-2) (right). *P*-values were examined by *t*-test and adjusted with "BH" method. **b** Cell cycle-related gene expression score (S and G2/M) in PEG10 RNAi, control, and PEG10 overexpression (OE) cells by single-cell RNA sequencing. **c** Junction count in tube formation assay for HUVEC cells transfected with control or PEG10-overexpression vectors. **d** Tube count in tube formation assay for HUVEC cells transfected with control or PEG10-overexpression vectors. **e** Junction count in tube formation assay for HUVEC cells transfected with control or PEG10-targeting siRNA. **f** Tube count in tube formation assay for HUVEC cells transfected with control or PEG10-targeting siRNA. *P*-values were examined by *t*-test, not adjusted

(Fig. 9c, d). In general, TGF-beta-1 protein significantly improved junction and tube formation number in the PEG10 overexpression cells (Fig. 9c,d), which restored junction and tube formation in PEG10 overexpression cells to a level not significantly different from the control cells (Fig. 9c,d). Together, these results showed that PEG10 disrupts endothelial cell biology via negatively regulating TGF-beta signaling.

Discussion

PE is clinically characterized as defective development of the placental villous tree, which is formed by trophoblasts [4, 5], the failure of uterine spiral artery remodeling, and deterioration of the endothelial vasculature [107]. Our results revealed a molecular account of PE pathogenesis.

Epigenetic mechanisms regulating trophoblast development in PE

The uterine spiral artery is the terminal branch of the uterine artery network and undergoes structural remodeling during implantation. The invasion of the spiral artery by EVTs blocks blood flow and replaces the inner wall of endothelial cells, resulting in the dilation of blood vessel termini [108]. The encroachment of the villous tree into the uterine cavity thus functions as an arteriovenous shunt between the uterine spiral
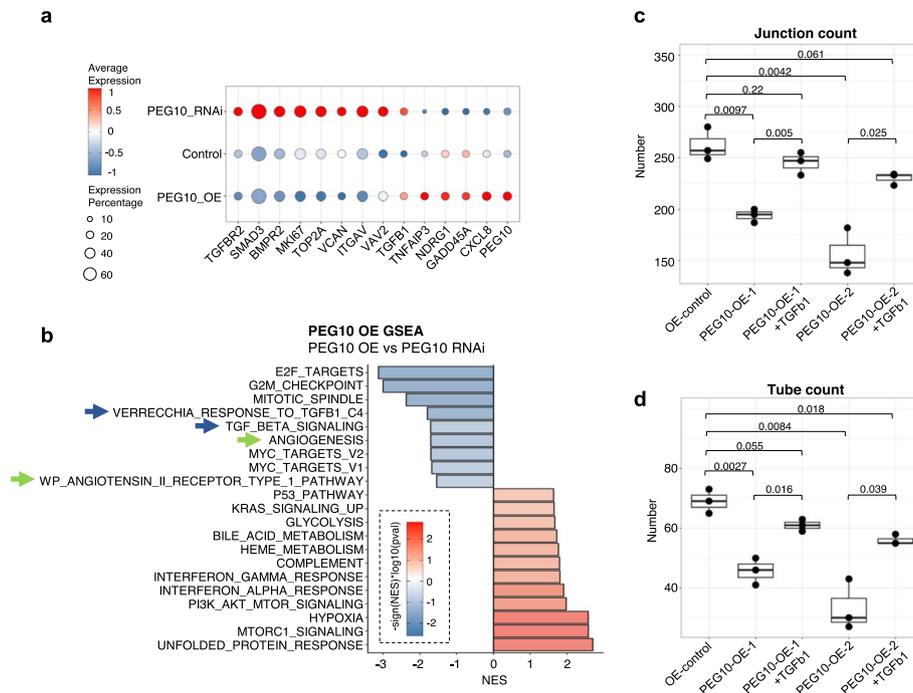
Gong *et al. Genome Biology* (2024) 25:117

Page 22 of 46

Fig. 9 PEG10 perturbs endothelial cell biology by negatively regulating TGF-beta signaling. **a** Differential gene expression between PEG10 overexpression (OE), PEG10 RNAi, and control cells. **b** Differential gene set enrichment activity between PEG10 OE and PEG10 RNAi cells. Blue arrows indicate TGF-beta signaling and green arrows show angiogenesis pathways. **c** The reduction of tube junction count in PEG10 overexpression cell is rescued by TGF-beta supplementation. **d** The reduction of tube count in PEG10 overexpression cell is rescued by TGF-beta supplementation. *P*-values were tested by *t*-test, not adjusted

artery ("arteriole") and venous system [104], with notably reduced blood flow velocity and increased total blood flow perfusing the villi that enable efficient gas and nutrient exchange [109]. In PE, the villous tree is underdeveloped and many terminal villi lack capillary [4, 5]. This is accompanied by reduced spiral artery remodeling in the maternal myometrium [6, 7].

By comparing individual cells derived from the placenta of PE and normal pregnancies, we corroborated and expanded upon our prior findings regarding alterations in cellular composition in PE [4, 7, 110–113]. Additionally, we identified previously unknown endothelial cells that are associated with PE. Particularly, we determined that in PE, defective epigenomic reprogramming including reduced de novo DNA methylation specific to extraembryonic tissue and dysregulated H3K27me3 modification coordinately stall the development of the trophoblast.

De novo DNA methylation of extraembryonic tissue-specific gene promoter has been extensively documented [50]. Our transcription network analysis in conjunction with DNA methylation profiling revealed that the PE phenotype may be pre-determined at the preimplantation stage as de novo DNA methylation occurs in the blastocyst stage. Additionally, we identified PE trophoblasts showing defective TF activities are primarily early-stage trophoblast progenitor cells and VCT but not terminally differentiated SCT and EVT. Hence, these findings suggest that an early imbalance of growth factor signaling between the inner cell mass and trophectoderm results in defective placenta

development. Additional investigations on organoids are necessary to clarify the precise molecular mechanisms involved in this process.

Retrovirus in placenta development

The emergence of placental structures is accompanied, and probably driven, by retrovirus domestication. Retroviruses exploit the placenta as their waypoint of trans-generational infection [114]. Due to host–pathogen co-evolution, most endoretrovirus genes are silenced during the adult stage through DNA methylation [115, 116]. However, these genes are only transiently activated during embryogenesis owing to global demethylation [42]. The genes and products of endoretroviruses are often expressed in the placenta [117–119], with the promoter of retroviral genes serving as species-specific enhancer elements that are specifically active in the placenta [120–123]. The incorporation of endoretrovirus-encoded envelop proteins, such as Syncytin I and II, promotes trophoblast fusion, resulting in the formation of multi-nuclei giant cells that line the maternal–fetal interface [124–129]. Interestingly, various animal taxa have exhibited the incorporation of envelop proteins from various retroviruses, suggesting convergent selection of retrovirus-gene-mediated placental features [117, 124, 130–138]. The incorporation of retroviral genes into trophoblasts probably facilitates not only the cell–cell fusion but also the invasiveness of trophoblasts, as well as immune suppression on the maternal side.

The adaptation of the placenta results in an active battleground for parental genes. Theoretically, such conflict may arise due to zero-sum-like competition between the mother and the fetus, such as for the nutritional supply. In the loci regulating such effects in the fetal genome, the paternal allele exhibits a proclivity towards fetal growth, whereas the maternal allele tends to protect the mother. Furthermore, the paternal endoretrovirus allele experiences reduced selection pressure for expression because it facilitates transmission of itself, in theory, to sibling littermates or the maternal body [139–141]. Examples that align with these theories include the LTR12 family of endoretrovirus, which are hypomethylated in the sperm and undergo ZGA-specific transcription. The transient expression of LTR12 endoretrovirus loci suggests a tight control over its chromatin accessibility, probably during the re-establishment of the global DNA methylation landscape. Although the functions of LTR12 endoretrovirus remain unclear, they may serve as an alternative 5′ promoter for downstream genes in certain cases [142]. On the contrary, two paternally imprinted endoretroviral genes, namely *RTL1* and *PEG10*, are expressed and functional in the placenta [143].

Epigenetic reprogramming regulates endoretrovirus expression

PEG10 protein and RNA transcript are prevalent in the PE placenta but not in control (Figs. 5 and 6). Our findings indicate that trophoblast maturation is accompanied by reduced levels of PEG10 expression (Fig. 5d), suggesting the maturation-dependent expression of PEG10 in early trophoblasts such as TSC (VCTp) and VCT. The PEG10 promoter is bound by PRC2 (Fig. 6a) and undergoes de novo DNA methylation during extraembryonic tissue development. In the PE placenta, LTR12C is commonly hypermethylated (Additional file 1: Fig. S18b), whereas PEG10 is maintained in a hypomethylated state. The imbalance of de novo methylation effect on these loci, albeit both

hypomethylated at the paternal germline, may suggest a redirected de novo DNA methylation effort from one type of endoretrovirus to another.

In PE, the concomitant increase in H3K27ac and reduction in de novo DNA methylation on the *PEG10* locus in the placenta (Fig. 6a and Additional file 1: Fig. S27) is concordant with the overexpression of *PEG10* in trophoblasts, especially in the terminally differentiated EVTs and SCTs (Fig. 5b, d). Although PEG10 is essential for villous blood vessel formation, the restriction of PEG10 expression to villous cytotrophoblasts may be essential to restrict its function exclusively to the fetal side instead of the maternal side. In the placenta of normal pregnancy, fetal capillaries in the villous tree are very fine, whereas the maternal spiral artery is deformed and wide open. The spatially adjacent blood vessels thus require differential regulation of angiogenesis.

PEG10 function in the regulation of angiogenesis via TGF-beta

PEG10 is subjected to strong purifying selection (loss-of-function [LoF] intolerance score pLI = 0.87, with only 1 LoF SNV in 249,162 sequenced alleles in gnomAD), suggesting its essentiality. In vitro analysis and gain-of-function experiments have shown that PEG10 binds to its own UTR to form VLPs [86], which can be transferred to other cells. A similar transfer of exosome-like particles between trophoblasts and HUVEC cells has been previously documented [88]. PEG10 attenuates TGF-beta signaling by interacting with TGF-beta co-receptor ACVRL1/ALK1, ultimately inhibiting its function [90]. Our results indicate that excessive PEG10 perturbs endothelial cell biology by negatively regulating TGF-beta signaling, to an extent that mirroring AVM formation. TGF-beta is a well-documented master regulator of angiogenesis. Mutations of genes in the TGF-beta pathway underlied many human vascular developmental disorders. Interestingly, in PE, s-Eng are notably elevated in the serum, suggesting that antagonism of TGFb signaling might be a common theme in regulating vascular remodeling and endothelial permeability in PE.

Conclusions

In our research, through single-cell analysis, we discovered developmental delays and transcriptional defects in PE trophoblasts. Our multimodal epigenomic studies showed that impaired extraembryonic tissue-specific de novo DNA methylation led to delayed development in PE trophoblasts. We observed excessive immature trophoblasts in PE placentas, which significantly upregulated maternally imprinted genes, notably PEG10. This gene can produce virus-like particles that transfer between cells. While PEG10 expression is limited to fetal vessel-interacting VCT in normal pregnancy, the maternal vessel-interacting SCT and EVT overexpressed PEG10 in PE pregnancy, enabling PEG10 VLP transfer into the maternal endothelial cells to induce ACE2 + /PAPPA2 + arteriole endothelial cells. These PE-specific endothelial cells are functionally defective marked by reduced TGF-beta and Wnt signaling. They share high transcriptomic similarity with other vascular malformation diseases caused by loss-of-TGF-beta signaling such as arteriovenous malformation. They also exhibit a gene expression profile indicative of impaired endothelial zonation. These results revealed how multi-layered epigenetic mechanism controls proper placentation, and how its disruption leads to a lethal pregnancy complication.

Methods

Human biospecimens

This study was conducted in accordance with the measures of the People's Republic of China on the administration of Human Assisted Reproductive Technology, the ethical principles of the Human Assisted Reproductive Technology, the Helsinki Declaration, and the internal ethic protocols of Peking University Third Hospital, Guangdong Woman and Children's Hospital, West China Second Hospital of Sichuan University and Zhongnan Hospital of Wuhan University.

The study protocol was approved by Peking University Third Hospital Medical Science Research Ethics Committee (approval number: (2020)YiLun(310–02)), Peking University Institutional Review Board (approval number: IRB0001052-16009), Institutional Review Board of Guangdong Woman and Children's Hospital (approval number: YiLun[201701044]), Ethics Committee of the West China Second Hospital of Sichuan University (approval number: YiXueKeYan2020(064)) and Institutional Ethics Committee of Zhongnan Hospital of Wuhan University (approval number: 2015029 and 2020102). Informed consent was obtained from the donors and their guardians. The study protocol was approved by the Institutional Review Board (IRB) of Peking University Third Hospital ((2020)YiLun(310–02) and IRB0001052-16009), IRB of Guangdong Woman and Children's Hospital (YiLun[201701044]), IRB of West China Second Hospital of Sichuan University (YiXueKeYan2020(064)), and IRB at Zhongnan Hospital of Wuhan University (approval number: 2015029 and 2020102). The human sample preservation by Department of Biological Repositories, Zhongnan Hospital of Wuhan University (official member of the International Society for Biological and Environmental Repositories-International Repository Locator, https://irlocator.isber.org/details/60) was approved by the Ethics Committee (approval number: 2017038) and China Human Genetic Resources Management Office, Ministry of Science and Technology of The People's Republic of China (approval number: 20171793).

For $10 \times$ single-cell RNA, single-cell ATAC, and immunostaining experiments, placenta samples were collected from Peking University Third Hospital and West China Second Hospital of Sichuan University. For placenta ATAC-seq and CUT&Tag experiments, placenta samples were collected from Peking University Third Hospital. For placenta methylation capture sequencing experiments, placenta samples were collected from Peking University Third Hospital and Guangdong Woman and Children's Hospital. FFPE tumor or normal tissue slides were collected from Zhongnan Hospital of Wuhan University, Wuhan, China.

Clinical assessment of human (pregnant female) phenotype was done according to ACOG guideline of hypertension in pregnancy [144]. Routine laboratory tests and pathology assessments were done according to the relevant Chinese clinical protocols.

Placentas were collected within 1 h of labor and transferred to laboratory in high glucose + 10% FBS in DMEM. Placenta samples were resected in PBS from the maternal or fetal surface of placenta, 2 cm*1 cm in size, and within 4–5 cm from the root of umbilical cord. For single-cell sequencing, CUT&Tag, or ATAC-seq, villous was further manually dissected from the maternal surface placenta with a pair of fine forceps for further processing. For methylation sequencing, placenta samples were flash frozen in liquid nitrogen and stored at -80 °C.

Cell culture

Human umbilical vein endothelial cells (HUVECs) were obtained and tested for mycoplasma non-contamination from the Chinese Academy of Sciences Cell Bank (Shanghai, China). The identification of the HUVEC cell line was conducted at the China Centre for Type Culture Collection in Wuhan, China. The cells were cultured and maintained in MEM medium supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin with 5% CO_2 at 37 °C.

siRNA and plasmid transfection

The PEG10-target specific siRNA and negative control siRNA were purchased from Shanghai GenePharma Co., Ltd. The sense sequence of the siRNAs were as follows: PEG10 siRNA-1 (5′-CCCACUACCUGAUGCACAATT-3′), PEG10 siRNA-2 (5′-GCA CUCGAUCUAUCGUCUUTT-3′), and negative control siRNA (siRNA control) (5′-UUCUCCGAACGUGUCACGUTT-3′). The plasmids for human PEG10, OE-PEG10 (or PEG10 OE-1) was synthesized based on the sequence of NM_001172437.2 by GenScript Biotech while OE-PEG10-2ORF (PEG10 OE-2) was synthesized based on the sequence in Segel et al. [86] by GenScript Biotech. Two plasmids encode the same PEG10 amino acid sequence. Negative control vector for overexpression is pCMV-GFP-HA.

HUVECs were transfected with siRNA or plasmids using Lipofectamine 3000 Transfection Kit (L3000-015, Invitrogen, USA), according to the manufacturer's protocol. The gene silencing efficiency and overexpression efficiency of transfected cells were confirmed by qRT-PCR and Western blot.

Quantitative real-time PCR (qRT-PCR)

HiPure Total RNA Mini Kit (R4111-03, Magen, China) was utilized to extract total RNA from cell lines. The ReverTra Ace qPCR RT Kit (FSQ-101, Toyobo, Japan) was used for the reverse transcription. Five-hundred-nanogram cDNA templates were added to a PCR system with a final volume of 20 μl. The primer sequences are listed in the following table:

Gene	Forward primer (5′–3′)	Reverse primer (5′–3′)
GAPDH	CTGGGCTACACTGAGCACC	AAGTGGTCGTTGAGGGCAATG
PEG10	GAGCACCAGGGATTTCTCAGT	GGTAGTTGTGCATCAGGTAGTG

Western blot analysis

The cells were lysed in RIPA buffer (containing protease inhibitor and phosphatase inhibitor) on ice for 30 min. The cell lysates were centrifuged at 12,000 *g* for 15 min and the supernatant was collected. For Western blots, total protein was separated using 10% SDS-PAGE gels, then transferred to PVDF membrane. Membranes were blocked with 5% skim milk in TBST buffer for 2 h at room temperature. The membranes were then incubated separately with the appropriate primary antibodies, including anti-PEG10 (Proteintech Inc, China, Cat. #4412-1-AP, 1:1000 dilution for WB and anti-GAPDH (Proteintech Inc, China, Cat. #0004-1-Ig, 1:5000 dilution for

Gong *et al. Genome Biology*   (2024) 25:117

Page 27 of 46

WB), for overnight at 4 °C. After washing three times with TBST buffer, membranes were incubated with secondary antibodies, including goat anti-mouse IgG (H + L) HRP (Sungene Biotech, China, Cat. #LK2003, 1:5000 dilution for WB) and goat anti-rabbit IgG (H + L) HRP (Sungene Biotech, China, Cat. #LK2001, 1:5000 dilution for WB), for 2 h at room temperature. Bands were detected using the Clarity™ Western ECL Substrate kit (1705061, Bio-Rad, USA), and images of blots were taken by the BioSpectrum 515 Imaging System (UVP, USA).

MTT assay
The transfected cells were seeded in 96-well plates at a density of 3000 cells per well. After incubation of 1–4 days, each well was added 20 μL 5 mg/mL MTT (M5655, Sigma, USA) for 4 h at 37 °C. The medium was removed and 150 μL of DMSO was added to dissolve the formazan precipitate in the 96-well plate. Finally, the absorbance of each well at 570 nm was tested using a microplate reader (Molecular Devices, USA).

Tube formation assay in vitro
Tubule formation assays were performed as previously described [145]. The 96-well plates were coated with 50 μL of Matrigel (40183ES10, YEASEN, China) per well and incubated for 1 h. The transfected HUVECs ($1.8 \times$ cells per well) were seeded on Matrigel and cultured for 6 h. The tube formation of HUVECs was observed and photographed using a microscope, and the counts of tubes and junctions were analyzed by ImageJ.

For evaluating the effect of TGFb1 treatment on the tube formation, the transfected HUVECs were pretreated with 5 ng/mL TGFb1 (HZ-1011, Proteintech, China) for 12 h, and then seeded on the Matrigel for tube formation assay described above.

Statistical methods
Clinical phenotypes were summarized as mean (range lowest–highest) or mean (percentage). All statistics for clinical phenotype was done with t-test (two-sided) or Wilcoxon Rank test (two-sided) if t-test was not applicable. Detailed statistical methods were briefly denoted in the figure legends or text accompanying. All statistical analyses in this study were performed using R (3.6.2) (http://CRAN.R-project.org).

Single-cell RNA and chromatin accessibility sequencing
Fresh tissues were processed immediately after being obtained from donors. Tissues were cut into tiny pieces (< 1 mm diameter) and then subjected to dissociation using collagenase II (Biofrox Ltd., #2275MG100) and 100 μl of DNase (Servicebio Ltd., #1121MG010) at 37 °C for 1 h. After dissociation, cells were filtered with 40 μm BD filter mesh and subsequently centrifuged at 250 g for 5 min. Cell pellets were washed in PBS twice and resuspended in 1 ml ice-cold RBC lysis buffer and incubated at 4 °C for 10 min. Ten milliliters of ice-cold PBS was added to the tube and subsequently centrifuged at 250 g for 10 min. After decanting the supernatant, the pellet was resuspended in 5 ml of calcium- and magnesium-free PBS containing 0.04% weight/volume BSA. Cells were counted using Trypan blue (Solarbio, Beijing, China). For chromatin accessibility sequencing, approximately 10^6 cells were used for nucleus extraction. Nucleus

Gong *et al. Genome Biology*   (2024) 25:117

Page 28 of 46

extraction were performed as $10 \times$ single-cell library preparation was done according to the manufacturer's protocol. Chrominum Single Cell 3' V3 kits and ATAC V2 kits were used. For RNA sequencing, single-cell suspensions were loaded onto a Chromium Single-Cell Controller Instrument ($10 \times$ Genomics) to generate single-cell gel beads in emulsions (GEMs) targeting ~ 8000 cells. After generation of GEMs, reverse transcription reactions were engaged to generate barcoded full-length cDNA, which was followed by disruption of emulsions using the recovery agent, and then cDNA clean-up was performed with DynaBeads Myone Silane Beads (Thermo Fisher Ltd.). Next, cDNA was amplified by PCR. Subsequently, the amplified cDNA was fragmented, end-repaired, A-tailed, and ligated to an index adaptor, and then the library was amplified. The scRNA libraries were sequenced aiming to have ~ 5000 reads per cell on Illumina Novaseq6000 with paired-end 150-bp reads. Sequencing was performed at Berry Genomics, Beijing, China. QC metrics is described in Additional file 9: Table S8.

For ATAC analysis, tagmentation was performed according to the manufacturer's protocol. After tagmentation reaction, nucleus suspensions were loaded a Chromium Single-Cell Controller Instrument ($10 \times$ Genomics) targeting ~ 10^4 nucleus in one reaction. After generation of GEMs, PCR reaction were performed to amplify the library. DNA clean-up was performed with size-selection XP beads. Libraries were sequenced aiming to have ~ 5000 reads per cell on Illumina Novaseq6000 with paired-end 50 bp reads. Sequencing was performed at Berry Genomics, Beijing, China. QC metrics is described in Additional file 9: Table S8.

Shared single-cell profiling of RNA and chromatin accessibility (SHARE-seq)

SHARE-seq experiment was performed with cryo-preserved transfected HUVEC cell line samples. The experiment was performed as described previously [146] with minor modifications: (1) After fixation, nuclei were isolated with Nuclei Lysis Buffer (10 mM Tris HCl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% Tween20, 0.1% NP40, 0.01% Digitonin, 0.75%BSA) on ice for 10 min; (2) Input 20,000 nuclei for ATAC Tn5 tagmentation per reaction; (3) Samples were multiplexed by using different R1-barcode in one single SHARE-Seq reaction. (4) ATAC library was purified with QIAquick Gel Extraction Kit (Cat. 28704, Qiagen); (5) Input 150 ng cDNA for RNA library tagmentation and purified with $0.7 \times$ AMPureXP beads (Cat. A63880, BECKMAN COULTER); (6) Each ATAC or RNA library from one single SHARE-Seq reaction was sequenced with MGI2000 sequencer with PE150 format to target approximately 500 M raw reads on average. Sequencing was performed at Euler Technology, Beijing, China. QC metrics is described in Additional file 9: Table S8.

Chromatin accessibility sequencing (ATAC-seq) on bulk placenta tissue

Twenty-milligram flash-frozen placenta tissue samples were minced using a double-sized douncer (Sigma, #D8938) in 1xHB (0.25 M sucrose, 0.06 M KCl, 0.005 M $MgCl_2$, 0.015 M NaCl, 0.01 M Tris HCl pH 7.5), added to 5 ml trypsin and 40 μl 5U/μl DNase I (Sigma, #D5025) and digested in 37 °C for 45 min, with two times of rotation in between to mix the reaction. The digested cells were then neutralized with equal volume DMEM (Thermo Fisher, #11995065) plus 10% FBS (Gibco, #16000044) and filtered through a 70-μm cell filter (BD Falcon, #352350). The homogenized sample was centrifuged at

500 *g*, 4 °C for 5 min. The sedimented cells were then resuspended in 400 μl 1xHB and washed once, transferred to 2 ml LoBind Tube (Eppendorf), and washed again. Cells were counted using Trypan blue (Solarbio, Beijing, China). After quantification, the cells were then added to a 30%-40%-50% iodixanol (Sigma, #D1556) gradient and centrifuged at 3000 *g*, 20 min at 4 °C. The cell layer at 30%-40% interface was collected for library preparation. DNA library were prepared with a Tn5 transposase kit (Vazyme, #TD501) using 1 million cells per reaction according to the manufacturer's protocol. After Tn5 transposition and PCR amplification, the sequencing library were quality-controlled with SYBR-green-based qPCR using primers for house-keeping gene (*GAPDH*) promoter and gene desert (chr5: 105187030–105190000) before sequencing. Each library was sequenced to 30 M reads on Novaseq6000 sequencer (Illumina, CA). Sequencing was performed at Berry Genomics, Beijing, China. QC metrics is described in Additional file 9: Table S8.

CutAndTag sequencing on dissociated cells

Digestion of placenta tissue follows the same protocol with bulk ATAC assay. After quantification, 50,000–100,000 cells were used for CutAndTag experiment. CutAndTag experiments were performed with NovoProtein CutAndTag 2.0 pAG-Tn5 kit (Novo-Protein, #N259) according to the manufacturer's protocol. Antibodies used in this study include the following: anti-H3K4me3 (Diagenode, #C154100003), anti-H3K27ac (Abcam, #ab4729), anti-CTCF (Abcam, #ab188408), anti-Histone H2A.Z (Abcam, #ab4174), anti-H3K27me3 (Abcam, Cat. #ab6002), goat anti-mouse IgG (Sangon, #D111024), and goat anti-rabbit IgG (Sangon, #D111018). Each library was sequenced to 2 × human genome coverage on Novaseq6000 sequencer (Illumina, CA) in PE150 format. Sequencing was performed at Berry Genomics, Beijing, China. QC metrics is described in Additional file 9: Table S8.

Nucleic acid preparation for DNA and methylation sequencing

Tissue genomic DNA was extracted from oral swab or placenta tissue with Qiagen Animal Tissue DNA Extraction Kit (Qiagen, #69504) according to the manufacturer's protocol. Genomic DNA from FFPE tissue slides were extracted using MagPure Tissue DNA DF Kit (Magen Inc., #MD5112-TL-06). Extracted DNA were quality-controlled by Qubit dsDNA HS assay (Thermo Fisher Ltd.) and Agilent 2100 Fragment Analyzer.

Single-stranded DNA methylation capture sequencing

Genomic DNA (200 ng) were bisulfite converted using EZ-DNA Methylation-Gold Kit (Zymo, #D5006) according to the manufacturer's protocol. After conversion, the DNA were subjected to a single-stranded library preparation protocol Tequila 7N (Euler Technology). In brief, the DNA were end-repaired using Klenow (NEB) and tailed with poly-A using TdT (Takara), ligated to a poly-T overhang adaptor using T4 DNA ligase (Enzymatics), and linearly amplified for 12 cycles using PhusionU (Thermo Fisher Ltd.). The linear products were then annealed to a 5' adaptor with 7 bp 3' random nucleotide overhang and amplified using adaptor oligos (Sangon, Shanghai, China) with Phusion (Thermo Fisher Ltd.), resulting in a library with proper Illumina sequencing adaptor ends ready for NGS. Hybridization was done with SeqCap EpiGiant Enrichment Probe

(Roche, #07138911001), oligos and SeqCap wash and binding buffers (Roche) following the manufacturer's protocol. After hybridization, the library was amplified using Phusion (Thermo Fisher Ltd.) for 8 cycles and sequenced on Novaseq6000 sequencer (Illumina, CA) to 100 M PE150 reads. Sequencing was performed at Berry Genomics, Beijing, China. QC metrics is described in Additional file 9: Table S8.

Genomic region lift-over

For genomic regions lift-over between UCSC GRCh38, GRCh37, mm9, and pantro2, the R package easyLift (version: 0.2.1, https://github.com/caleblareau/easyLift), the lift-over executable from UCSC Kent Utility (https://genome.ucsc.edu/cgi-bin/hgLiftOver) and the lift-over synteny chain files from UCSC Genome Browser were used.

Genomic element annotation

Genomic element annotation was done using ANNOVAR (http://annovar.openbioinformatics.org/) and bedtools (https://bedtools.readthedocs.io), where applicable. Differences on any class of genomic element were computed using Fisher's exact test.

Single-cell RNA data preprocessing and cell clustering

Loompy-Kallisto [147] was used for mapping the RNA data for gene expression analysis. Loom files were read in R by hdf5r (https://cran.r-project.org/web/packages/hdf5r/index.html) and preprocessed with Seurat 3.2.2 [61]. Quality control was done for every single sample individually to filter against gene counts, UMI counts, total reads, and mitochondrial reads. Generally, cells with > 10% mitochondrial reads, or with UMI < 600 or > 5000, or with gene counts > 5000 were filtered prior to subsequent analysis. Such quality control process might iterate at every subsequent step to ensure the stringency of analysis. Individual samples were processed through the Seurat pipeline. Data firstly passed DoubletFinder (2.0.3) [148] with standard parameters to filter against potential doublets. The filtered data were then normalized (Log Normalization by Seurat::NormalizeData), and top 2000 variable genes were identified. Ribosomal proteins, heat shock proteins, and chaperones were intentionally removed from the variable gene list because of the highly inconsistent nature of their behaviors between different tissue types. Gene expression profile were then scaled and reduced by PCA using Seurat. Generally, ≥ 30 PCA components were included in subsequent steps. FindClusters function were initially performed using a high resolution then gradually lowered to ensure the final clusters are less than PC components. SingleR [149] annotation with human reference and conventional markers were used to initially categorize the cell clusters. Differentially expressed (DE) genes were identified with Seurat FindAllMarkers function with Wilcoxon test for the cell clusters. Comparison of the found DE genes with conventional markers was performed to ensure that clusters contain relatively pure cell population. Cell type-annotated cells were then separated into different subsets based on their types. Detailed cell types were classified manually according to known canonical markers and comparison with reference single-cell datasets. In this study, trophoblasts, endothelial cells, fibroblasts, CD8 T cells, CD4 T cells, NK cells, B cells, and myeloid cells were considered: "pools" for subsequent analysis. After preprocessing, similar type of cells from different samples were merged and re-analyzed. Quality control, doublet

identification, dimensionality reduction, cluster identification, differentially expressed gene identification, and cell type identification were iteratively performed on these "pure" sets of cells. In such setting, cell type composition between samples were relatively homogeneous and usually it is unnecessary to perform data integration. When sample-driven variation is evident or in cases when dataset contains samples collecting from different sources, to control against technical variations, batch effect removal was performed using Harmony [60] with (vars.to.regress = 'source') option during data scaling. Contaminant cells which mis-segregated into large pools were identified and put back into the unprocessed pool. Iterative processing of cells was done semi-automatically until all cells were processed.

Cell cycle scoring

Single-cell RNA cell cycle gene set activity measurement and cell cycle classification was done with CellCycleScore function in Seurat [61] using the `cc.gene.2019` dataset as source of G2/M and S phase gene sets. For visualization, G2M.score and S.score of a single cell were added together to suggest the relative "actively proliferating" probability.

Developmental lineage analysis

Diffusion map of single-cell RNA expression data was computed with destiny [71] (version: 3.0.1). RNA velocity analysis of single-cell RNA sequencing data was performed with scVelo [66] using Reticulate [150] in R (3.6.2) with Python3. With RNA velocity defined the root cluster, slingshot was performed on diffusion maps to produce minimally spanning tree lineages.

Regulatory network inference with cisTarget

Velocity genes determined in RNA velocity analysis were categorized into co-regulatory modules by non-negative matrix factorization and transcriptional binding sites (TFBS) were extracted by RcisTarget [151] with the human GRCh38 -500 bp ~ +100 bp database downloaded from cisTopic [152] (hg38_refseq-r80_500bp_up_and_100bp_down_tss.mc9nr.feather). After extraction, high-confidence co-regulator TF of regulons with NES > 3.0 were extracted from the data. Visualization of co-regulated transcription network is performed with visNetwork (https://cran.r-project.org/package=visNetwork).

Differentiation potential assessment

Cellular differentiation potential was assessed by CytoTRACE (version: 0.1.0) [68]. Briefly, RNA expression matrix of single cells is extracted from Seurat object, with all transcripts regardless of their in-assay-variability. CytoTRACE analysis was performed on this matrix without downsampling. Per-cluster median CytoTRACE score was used as an index for the differentiation state of each cell cluster.

Differential expression in scRNA dataset

Differential expression (Additional file 7: Table S6) between any two sets of single cells were computed with FindAllMarkers function in Seurat with Wilcoxon test as default statistical method. The Log2FC threshold was set to be 0 and minimally expressed percentile was set to 0.1 (10%). Statistically, differentially expressed gene were determined

Gong *et al. Genome Biology* (2024) 25:117

Page 32 of 46

from the resulting table using adjusted *P*-value (Berfernorri method) < 0.05 and abs(Log2FC) > 0.5 as threshold.

Define maternal or fetal origin of placental cells

Placental myeloid, endothelial fibroblast/stromal cells from both maternal and fetal origins. Maternal- or fetal-specific marker gene sets in each cell type were calculated by FindAllMarkers based on scRNA data of Vento-Tormo et al. [55]. Significant marker genes were picked by the following criteria: (1) p_adj_val < 0.001; (2) avg_logFC > 1; (3) abs(pct.1-pct.2) > 0.3 between the annotated (WGS SNV based) fetal and maternal cells, which means the difference of cell fraction that positively expressing a certain gene between the given cluster and all other clusters is at least 30%. For example, if 60% of single cells in a cluster X expressed gene Y, compared to only 20% of single cells not in cluster X expressed gene Y, then the abs(pct.1—pct.2) is 0.4. This is followed by gene set activity measurement in single cells using Seurat::AddModuleScore. In-house sequenced single cells of each cell type are classified as maternal or fetal cells according to the gene set activities mentioned above.

Single-cell chromatin accessibility data analysis

Single-cell ATAC (chromatin accessibility, scATAC) raw reads were mapped with cellranger-atac [153]. The mapped fragment files were then processed by ArchR [154] (version:1.0.1). Quality control, doublet identification, LSI-based dimensionality reduction, clustering, gene expression inference, peakset identification, and marker peak finding were all performed in ArchR. Batch effect removal was not involved in this analysis. Basically, cells from all samples were pooled in the initial analysis, manually annotated with known markers, and separated into different subsets. The subset data were then subjected to scRNA integration in ArchR using Seurat CCA algorithm, with constrains for integration on large pools of cell type. Peaks were called for each single-cell dataset using MACS2. Chromatin accessibility on given genomic regions was also calculated in ArchR for single-cell chromatin accessibility differential analysis. Transcription factor footprinting (activity measurement) was done in ArchR with chromVAR. Co-accessible regions were calculated with 'addCoAccessibility' (ArchR) (maxDist = 1000000) and filtered with correlation > 0.1.

Extraction of cell-specific reads for SNV analysis

The 10 × Genomics cellranger (4.0) pipeline was used for mapping the scRNA for SNV analysis. The scATAC data were mapped as described above. Barcodes of defined scRNA or scATAC cell groups were firstly curated from Seurat object. BAM files was then processed by Rsamtools and reads of given specific cell barcodes were fetched. SNV were analyzed using samtools mpileup and visualized using IGV.

Definition of VLP cargo genes

Putative PEG10 cargo genes were defined as follows: (1) abs(LogFC) > 0.5 between PE PAPPA2 + endothelial cells compared to control PAPPA2-counterparts, or adjusted *P*-value < 0.01; 2. Completely inaccessible chromatin status in PE and control endothelial cells. The PE-specific PEG10 cargo genes were defined as follows: (1) LogFC > 0.5 in

Gong *et al. Genome Biology*   (2024) 25:117

Page 33 of 46

PE PAPPA2 + endothelial cells compared to control PAPPA2-counterparts, or adjusted *P*-value < 0.01; (2) LogFC > 0.5 in PE trophoblast cells compared to control, or adjusted *P*-value < 0.01; 3. Accessible chromatin peaks in trophoblast cells; (4) Completely inaccessible chromatin status in PE *and* control endothelial cells. For each putative cargo gene, manual visual inspection of scATAC tracks were performed to confirm the chromatin accessibility state in trophoblast and endothelial cells (Additional file 8: Table S7).

Sequence motif in cargo UTR

5′ and 3′ UTR of all potential transcripts from cargo gene loci were extracted from reference genome (R package org.Hs.eg.db, version: 3.6.2). Motif finding and visualization was done with R package rGADEM. Motif PWM were generated with R package TFB-Stools and similarity was calculated using Pearson's method. To evaluate statistical significance, background motifs were pulled from JASPAR (vertebrate non-redundant core motifs) or ORegAnno, and similarities between the background motifs and the found UTR motifs were calculated. *Z*-score deviation from the background distribution was then computed with *scale* function in R.

Bulk ATAC, Cut-and-Run, ChIP, and CutAndTag sequencing data preprocessing

Raw paired-end open chromatin tagmentation (ATAC), Cut-and-Run, ChIP, or CutAnd-Tag sequencing data were mapped to human reference genome GRCh38 (Cut-and-Tag) or GRCh37 (Cut-and-Run, ChIP, and ATAC) using Bowtie2 (-k 10 --very-sensitive -X 2000) (https://github.com/BenLangmead/bowtie2). All unmapped reads, non-uniquely mapped reads, reads with low mapping quality (MAPQ < 20), and PCR duplicates were removed. For CutAndTag sequencing libraries, data were used as is, because all CutAndTag libraries showed excellent concordance to reference peak sets in Encode library. As bulk placenta tissue ATAC-seq suffers greatly from cell debris containing free-floating cell-free DNA, we further filtered the libraries for quality control. The in-house ATAC-seq data were quality-controlled by assessing insertion size (using an in-house R script) and TSS enrichment (using an in-house R script with GenomicRanges package (https://github.com/Bioconductor/GenomicRanges) measuring the depth ratio at the promoter region (refFlat annotation from UCSC Genome Browser) (0 bp of TSS vs. 1kbp +/− of TSS). A QC-passed ATAC-seq library must have TSS enrichment of 6, mapped deduplicated sequencing fragments ≥ 20 M PE reads, PCB1 > 0.9, PCB2 > 3 (https://www.encodeproject.org/pipelines). Enrichment peaks were determined by intersecting peaks found from MACS2 callpeak (-f BAMPE, https://github.com/taoliu/MACS) and Genrich (-r -m 1 -j; for ATAC only; and standard parameter for CutAndTag) (https://github.com/jsh58/Genrich). Further quality control of ATAC-seq libraries including read length, V-plot, and TSS enrichment were done with custom R script and deeptools (https://github.com/deeptools/deepTools). Reliable peaks were identified with IDR (https://www.encodeproject.org/software/idr). Reliable ATAC peaks from different set of data were converged with 1 bp minimum overlap and extended to the largest width of overlapping peaks. Joining these operation results in a set of non-overlapping, varied-width peaks across the genome encompassing all reliable open chromatin region.

Gong *et al. Genome Biology* (2024) 25:117

Page 34 of 46

Correlation of bulk and single-cell ATAC-seq datasets

Coverage (RPKM) of ATAC-seq on IDR peaks were calculated for each sample, and Pearson correlation coefficient were calculated for each pair of samples in R (3.6.2).

Measurement of difference of bulk ATAC-seq peaks

Read coverage of sequencing library were collected over the repeatable ATAC enrichment peak mentioned above with Sambamba. Differential enrichment was performed with DESeq2 (https://github.com/mikelove/DESeq2) using standard parameters. For samples with few replicates, we adapted a general linear model approach for estimating difference following the method mentioned in Reilly et al. [155]. Differential ATAC peaks between preeclampsia and normal placenta are summarized in Additional file 5: Table S4.

Motif finding and transcription factor footprint analysis in bulk ATAC-seq

Transcription factor binding motifs were collected from ReMap. Motifs were extracted using HOMER package (http://homer.ucsd.edu/homer/). Raw ATAC-seq data were preprocessed using RGT-Hint (www.regulatory-genomics.org/hint) package with standard parameter (rgt-hint footprinting --atac-seq --paired-end) and matched to motifs (rgt-motifanalysis matching --organism=hg19). Differential transcription factor footprint was analyzed by Hint (rgt-hint differential --organism=hg19 --bc) on three independent pairs of biological replicates between preeclampsia and non-preeclampsia placenta.

Histone modification analysis

Histone modification peaks were identified by MACS2 using following parameters: H3K4me3/H3K4me1/H3K27ac: -g hs -nomodel -nolambda; H3K27me3: -g hs --broad --broad-cutoff 0.05 -nomodel -nolambda. Weak (summed RPKM < 20) or irreproducible peaks were removed for further analysis. Histone modification peaks overlapping known ATAC enrichment peak were used for further analysis. Read coverage of histone modification sequencing library were collected over the repeatable ATAC-and-histone-modification enrichment peak mentioned above with Sambamba, and normalized to RPKM (reads per kilo million for library), and Z-normalized.

Broad H3K27me3 regions were called by macs2 (-f BAMPE -g hs -q 0.01 --broad --broad-cutoff 0.1 --nolambda --SPMR --nomodel -B). The called regions were merged to neighboring regions within $+/-5$ kb to produce H3K27me3 domains. H3K27me3 CUT&Tag RPKM on each domain were computed. Mean and standard deviation of H3K27me3 RPKM were calculated for each region for PE and control samples, respectively. Differentially modified region were called as abs(mean(PE)-mean(control))/(std(PE)+std(control)) > 2. Enrichment of EZH2-bound regions for H3K27me3 gain or loss in PE were done with regionR::permutationTest. Overall profile of H3K27me3 on EZH2-bound regions was computed by deepTools.

DNA methylation data processing

Raw bisulfite-converted DNA methylation sequencing data, either downloaded from NCBI SRA or directly from in-house sequencing, were processed using fastp [156] (--trim-front2 20 -w 20) (https://github.com/OpenGene/fastp) and mapped to

GRCh37 + decoy reference genome using BWA-Meth (https://github.com/brentp/bwa-meth) using standard parameters. Mapped data were deduplicated and sorted using Sambamba (https://github.com/biod/sambamba) and Samblaster (https://github.com/GregoryFaust/samblaster). CpG methylation levels were extracted using Pile-O-Meth (https://github.com/dpryan79/MethylDackel) toolkit. For all libraries, conversion rate was quality-controlled by CHH methylation level > 99%. Basic statistics of in-house sequencing library were further quality-controlled by on-target rate and on-target coverage with bedtools (https://github.com/arq5x/bedtools), and duplication rate and mapping rate with Sambamba.

For mouse data of extraembryonic tissue (ExE) methylation [50], the sequencing reads were similarly preprocessed and mapped to mm9 reference. ExE-specific de novo methylation region from Smith et al. [50] were directly used. To compare methylation level on human homologous regions, these mouse methylation regions were lifted-over from mm9 to GRCh37.

For Chimpanzee data of sperm methylation, the processed CpG methylation level from Molaro et al. [157] was directly used, with lift-over from hg18 and pantro2 to GRCh37.

Differential methylation analysis

CpG methylation level (beta: defined as reads of C nucleotide over total read coverage on a single C or G base on CpG loci) was measured for each CpG loci across the genome as mentioned above using Pile-O-Meth. For each locus, beta from preeclampsia or non-preeclampsia (including normal, gestational hypertension, and gestational diabetes) pregnant female were summarized in R (3.6.2) using an in-house script. Differentially methylated loci (DML) were defined as follows: (1) $P < 0.01$ for t-test between preeclampsia- and non-preeclampsia individuals; (2) beta difference between preeclampsia and non-preeclampsia individuals > 0.1.

Differentially methylated region (DMR) analysis

Initial DMR candidate were made by merging within-100 bp-apart DML. The average beta of each initial DMR were calculated as mean beta of all CpG encompassed in the DMR. This average beta was subjected to t-test, and $P < 0.01$ regions were selected as candidate "seed" DMR. Segments of methylation difference level were computed using a circular binary segmentation approach on beta difference between preeclampsia and non-preeclampsia placenta with DNAcopy (https://github.com/veseshan/DNAcopy). K-means clustering was performed using R (3.6.2) on the methylation beta difference on each segment, and clusters of segments fully encompassed candidate "seed" DMR were selected as true DMR candidate (Additional file 6: Table S5).

Pseudotime analysis of methylation data

Single-cell or bulk methylation sequencing (bisulfite sequencing: WGBS, GSE81233 [42, 158]) data and metadata were collected from NCBI SRA. In-house data were described as mentioned before. Data mapping was done with Monocle3 (https://github.com/cole-trapnell-lab/monocle3) with standard practices, with the mean beta value as "expression value." In this and the following combined ATAC-seq analysis, we chose Monocle3 over other tools because of its simplicity in processing large number

of individually sequenced scATAC/scWGBS libraries and to combine them together with bulk sequencing data. Dispersion was estimated with negative binomial model. Mean beta values from PE-hypermethylated LTR12C regions and PE-hypomethylated de novo methylated loci in extraembryonic tissue are used for clustering and pseudotime analysis. Data were preprocessed using 10 dimensions with UMAP. To cluster samples from different techniques, we used "align_cds" function from Monocle3 using mutual nearest neighbour alignment method [159] and performed pseudotime trajectory analysis on these aligned data.

Pseudotime analysis of ATAC-seq data

Single-cell or bulk ATAC sequencing data (SRP163205 [45] and GSE101571 [76]) and metadata were collected from NCBI SRA. In-house data were described as mentioned before. The overall processing was similar to pseudotime analysis of methylation data mentioned above, with the only difference that ATAC-seq RPKM from all DMR regions were used as "expression level" in the analysis.

Overlapping region analysis

DMR and ATAC-seq peak regions were overlapped with known ultraconserved non-coding elements (UCNE) [79], human-accelerated regions (HAR) [78], and placental animal-specific accelerated regions (PAR) [77] and statistical significance of set enrichment were calculated with Fisher's exact test in R (3.6.2). For DMR region overlapping with known imprinted genes, imprinted gene list is downloaded from https://www.geneimprint.com/ and only experimentally validated imprinted genes were used. Permutation-based overlap size statistics is performed by regioneR (v1.28.0). Statistics of DMR region overlapping with all known chromVAR [80] region annotation sets is performed by LOLA (v1.26.0).

Sequencing of a paternally derived PEG10 allele

The paternally derived PEG10 allele chr7:94665871 T > A was initially discovered by comparing the single-cell RNA-seq data on fetal and maternal faces of placenta of a PE female, 2005139. Presence of the allele in placental trophoblast was confirmed by scATAC SNV analysis. However, the maternal cells have no scATAC reads covering on this region, probably because of no expression of PEG10 in the maternal cells. Oral swabs of the mother (2005139), father (2105466), and progeny (2105342, after birth) were taken and Sanger-sequenced using primers 5′CACATCCTCTCTGAAACGGCT3′ and 5′CCTTTCCACACTGCACCGAT3′. Primer validity was confirmed with in-house known heterozygous carrier and homozygous wild-type controls of the same allele. Low-pass WGS (LWGS) with standard double-stranded library preparation assay using in-house adaptors were also used to confirm haplotype transmission of the T allele in progeny was from mother. To do this, LWGS data were mapped using bwa-mem and SNV were analyzed using sentieon DNAscope. Haplotype analysis was done with Eagle [160].

Gong *et al. Genome Biology*   (2024) 25:117

Page 37 of 46

RNA velocity in arterial and trophoblast cells

RNA dynamical velocity analysis of single-cell 3′ RNA sequencing data, including inferring dynamical streamline, computing dynamic genes, latent time, and the probability of root cell and terminal cell, was performed with scVelo [66] using Reticulate [150] in R (3.6.2) with Python 3 and following standard protocol. PE and control trophoblast were performed RNA velocity separately, while PE and control arterial cells were performed as a whole.

Association of PE transcriptome with single-cell RNA profiles by Scissor

Raw data of bulk placenta RNA-seq was downloaded from GSE148241 [63, 64] and mapped by STAR [161] to give read count matrix covering genes. The parameters used in mapping and quantification were as follows: alignment was performed with STAR (version 2.5.3a). STAR reference genome was built with '--runMode genomeGenerate --runThreadN 16 --genomeDir b37/STAR-genome --genomeFastaFiles b37/human_g1k_v37_decoy.fasta --sjdbGTFfile b37/Homo_sapiens.GRCh37.82.gtf --sjdbOverhang 100'. Alignment was performed with '--chimSegmentMin 20 --chimScoreMin 5 --quanMode GeneCounts --twopassMode Basic --outSAMtype BAM SortedByCoordinate'. The quantified gene table was then passed to DEseq2 [162] and analyzed with the default parameter, only filtering for genes that are expressed. Expressed genes were filtered as 'count > = 10'. Differentially expressed genes were defined as the pvalue (raw *P*-value) < 0.05 and absolute value of log2FoldChange (original log2 fold change) > 0.5.

The correlation between bulk PE placental transcriptome and scRNA placental cell was inferred by Scissor [62] (https://sunduanchen.github.io/Scissor/vignettes/Scissor_Tutorial.html), setting alpha = 0.05, family = "binomial", cutoff = 0.2 (the default setting, Scissor selected cells should not exceed 20%). To avoid technical aberration caused by the imbalance of single-cell number in each cell type, scRNA placental cells were downsampled to 1000 cells per cell type. The Scissor-positive cell is defined as the average of correlations with all bulk samples is greater than 0 and the number of positive correlations is larger than the number of negative correlations, and the Scissor-negative cell is defined as the average of correlations with all bulk samples is less than 0 and the number of negative correlations is larger than the number of positive correlations. By Scissor::reliability.test, the *p*-value of Scissor identified association is 0.000 and AUC is 0.8738095.

Association of PE GWAS risk loci with single-cell chromatin accessibility profiles by SCAVENGE

For GWAS risk loci association, PE-associated GWAS loci from [102] were downloaded from GWAS catalog. The raw genotyping dataset is not publicly available and fine-mapping of risk loci is impossible. To take an approximation of posterior probability of association, we then took the log of adjusted *P*-value of these significantly associated loci and normalized them to a max of 1. A peak-x-cell matrix is extracted from the full scATAC dataset from ArchR and subjected to LSI, MNN analysis in SCAVENGE [65]. *Z*-scores of loci enrichment in single cells, and cell–cell similarity-based propagation of the *Z*-score, were performed with SCAVENGE. The seed cells were selected as the top 0.1% *Z*-score enriched cells. Cells with top 25% TRS were considered positively associated with trait.

Immunostaining

Placenta tissues (from 3 SPE, 3 control, and 1 SPE with sIUGR donors) were processed with standard paraffin section protocol to 10–30-µm-thick slides. Immunostaining with primary antibodies, including Anti-PEG10 (1:1500, Abcam, ab215035), Anti-CK7 (1:1500, Abcam, ab9021) and Anti-CD31 (1:1000, Abcam, ab9498), and DAPI followed the instruction of Opal 7 color manual Kit (AKOYA Biosciences, NEL811001KT). Fluorescent slides were scanned using the PhenoImager (Akoya Biosciences) using $\times 20$ objective with the following exposure times: DAPI MSI, 0.38 ms; Cy5 MSI (PEG10), 17.33 ms; Cy3 (CD31), 15.76 ms; FITC (CK7), 6.58 ms. Images were generated using inForm software (Akoya Biosciences).

Automatic quantification of fluorescent staining intensity

Image exportation was done by inForm Tissue Finder software (version: 2.6). After exportation, the images were imported to Fiji to split into different channels. Channel-specific images were imported into R (version: 4.1.1). Masking was performed to get endothelial and trophoblast cells, by identifying the contour with CD31 or CK7 staining. Quantification of PEG10 intensity inside of cell contours was subsequently performed. Masking and quantification were performed with computeFeatures.basic function from EBImage (package version 4.34.0).

Gene set enrichment analysis

Differentially expressed gene set between PE and control arterial cells were calculated by Seurat::FindAllMarkers, setting min.diff.pct = 0, logfc.threshold = 0, only. pos = F. Gene set enrichment analysis was performed with fgsea [163] R package (version:1.18.0), with differentially expressed genes in PE and msigdbr [164] database (verrsion: 7.5.1, category = "H" or "C2"). Significant pathway is defined as pval (raw *P*-value) < 0.05 and the absolute value of NES > 1.

Similarity between AVM and PE arterial endothelial cells

The brain vessel single-cell RNA-seq dataset was downloaded from NCBI GEO (GSE187875 [89, 100]) and used with its original cell type and pathology annotation. Seurat (v4.0) were used to extract differentially expressed genes (DEG) specifically upregulated in AVM or control arterial endothelial cells. Seurat::AddModuleScore was used to calculate the gene set activity in fetal and maternal endothelial cells of placenta.

Correlation of cargo expression profile (Additional file 1: Fig. S29)

scRNA-upregulated, scATAC-silent genes were firstly selected as mentioned in "Definition of VLP cargo genes." The mean expression levels of these genes in each cell type were extracted for PE and control samples, respectively. For each type of sample, Pearson's correlation was performed between all cell types.

Mitosis aging rate determination using EpiTrace

Mitosis age of single cells in the scATAC dataset were inferred by *EpiTrace* [69] (https://github.com/MagpiePKU/EpiTrace) following the standard protocol. Briefly,

chromatin accessibility on age-associated DML sites were measured for each cell and the resulting score is subjected to iterative HMM smoothing-approximation. The deduced cell mitotic age is then divided by gestation week of the sample, to give a single-cell mitotic aging rate which approximates cell proliferation rate. Differences between mitotic aging rate were measured by Wilcox test.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13059-024-03265-z.

Additional file 1: Fig. S1. Placental cell type composition in this study. Fig. S2. Placental cell types with mixed-origin. Fig. S3. Similarity between accessibility and gene expression in scATAC. Fig. S4. Scissor inferred phenotype associated cells. Fig. S5. Examples of marker gene expression in PE and control placenta bulk RNAseq dataset [172–179]. Fig. S6. Cell cycle scores in trophoblast clusters. Fig. S7. Expression of trophoblast cluster marker genes. Fig. S8. The frequency of trophoblast cluster and cell cycle phase. Fig. S9. Latent time distribution across gestational week in control and PE trophoblasts. Fig. S10. Latent time of trophoblasts, grouped by trimester. Fig. S11. scRNA analysis of trophoblast cells in external validation dataset. Fig. S12. Developmental trajectory switch in PE trophoblast. Fig. S13. RNA expression and transcription factor binding activity of master transcription factors (TF). Fig. S14. Transcription factor activities upstream of EZH2 in PE placenta. Fig. S15. Master transcription factor controlled velocity gene set expression in the trophoblast cells in external validation dataset. Fig. S16. Bulk ATAC seq assay on frozen placenta tissue. Fig. S17. Differential DNA methylation between control and PE. Fig. S18. Differential methylation in fetal and maternal face of placenta between PE and control. Fig. S19. Deficient ExE-specific *de novo* methylation in paternally imprinted loci in PE placenta. Fig. S20. DNA methylation levels on recently evolved, primate-specific retrotransposons, particularly the imprinted LTR12C, discriminate PE and control placenta. Fig. S21. PE DMR regions are enriched with PRC2 related binding loci. Fig. S22. Differential expression of imprinted genes in trophoblast. Fig. S23. Reduced H3K27me3 modification on EZH2-controlled genes in PE placenta. Fig. S24. Reduced H3K27me3 modification on paternally imprinted genes in PE placenta. Fig. S25. PE trophoblast overexpressed genes to stall its cell cycle progression. Fig. S26. Differential gene expression in each trophoblast lineage. Fig. S27. Differential epigenetic modification around *PEG10*. Fig. S28. Motif similarity of cargo genes. Fig. S29. Heatmap of similarity of cargo gene RNA expression between trophoblasts and endothelial cells. Fig. S30. Differential expression and cell cycle activity in endothelial cells. Fig. S31. Differential expressed genes in Wnt beta Catenin and TGF-beta pathways. Fig. S32. PEG10 qPCR and western blot. Fig. S33 Transcriptional dysregulations under PEG10 perturbation.

Additional file 2: Table S1. Characterization of donors in the study.

Additional file 3: Table S2. Single cell annotation.

Additional file 4: Table S3. Statistical details for trophoblast CytoTRACE analysis.

Additional file 5: Table S4. Differential peaks in bulk ATAC assay.

Additional file 6: Table S5. PE specific differential methylation regions.

Additional file 7: Table S6. Differentially expressed genes in trophoblast.

Additional file 8: Table S7. List of cargo genes.

Additional file 9: Table S8. QC metrics for sequencing assays in the study.

Additional file 10. Uncropped images of Western blots in Fig. S32.

Additional file 11. Protocol for accessing the data with controlled access on GSA.

Additional file 12. Review history.

Review history
The review history is available as Additional file 12.

Peer review information
Wenjing She was the primary editor of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Authors' contributions
Y.Z. and Y.W. conceived and designed the study. X.L.G., J.X.L., H.W., H.W.M., Y.W., J.X.Y., M.F., W.D., and Y.M.Q. collected pregnancy-related samples. Y.Z. and Y.X. arranged and collected clinical tumor and normal tissue samples. W.J., Q.C., H.H.G., and Y.Z. performed molecular biology experiments. G.W. and J.X.L. performed in vitro experiment. W.J. performed single-cell sequencing. Y.Z., W.J., X.J.L., and J.S.L. performed bioinformatic analysis. Y.Z., W.Y., Y.Y.Z., X.H.L., and A.H.Y. oversee the study. Y.Z. and W.J. wrote the paper with input from all authors.

Author's Twitter handle
@picapica668 (Yi Zhang).

Gong *et al. Genome Biology*   (2024) 25:117

Page 40 of 46

Availability of data and materials
The raw sequence data reported in this paper have been deposited in the Genome Sequence Archive (Genomics, Proteomics & Bioinformatics 2021) in National Genomics Data Center (Nucleic Acids Res 2022), China National Center for Bioinformation / Beijing Institute of Genomics, Chinese Academy of Sciences (accession number: HRA001423 [165]) that are publicly accessible at https://ngdc.cncb.ac.cn/gsa-human. The dataset has been reported to the China Human Genetic Resources Management Office (CHGR), Ministry of Science and Technology of The People's Republic of China. Data access is under regulation of the Chinese Biosafety Law and CHGR policy. Researchers wish to access the data should apply on the GSA website following the instructions provided on Additional file 11. The data access committee (DAC), Drs. Yuan Wei (weiyuanbysy@163.com) and Yi Zhang (zy@eulertechnology.com), would evaluate the request and help the researcher to access the data. The DAC would reply by email within a week of the request. The source codes for reproducing the figures were deposited under GPL 3.0 license at GitHub [166]. The source data accompanying the source codes were deposited under CC 4.0 license at Zenodo [167].
The study additionally used a series of public datasets from NCBI SRA sequencing read archive (https://www.ncbi.nlm.nih.gov/sra) and the Chinese Genome Sequencing Archive (https://bigd.big.ac.cn/) including H3K27ac ChIP-seq data of placenta from Roadmap Project: GSM1127147 [168, 169]; human embryo ATAC-seq: SRP163205 [45, 170] and GSE101571 [76, 171]; human embryo methylation sequencing: GSE81233 [42, 158]; human placenta single-cell RNA-seq: PRJNA492324 [54, 56], PRJEB28266 [55, 57], GSE150578 [58, 59], and GSE173193 [70]; human placenta bulk RNA-seq GSE148241 [63, 64]; human brain vascular single cells RNA-seq GSE187875 [89, 100].
The Encode transcription factor binding sites, ChromHMM tracks, CpG islands, repeat regions data, and lift-over chains were collected from UCSC Genome Browser (http://www.genome.ucsc.edu/). Chromatin interactions were downloaded from 4Dgenome database (http://4dgenome.int-med.uiowa.edu). Transcription factor binding information for bulk ATAC-seq analysis were downloaded from ReMap database (http://pedagogix-tagc.univ-mrs.fr/remap).

Declarations

Ethics approval and consent to participate
The study protocol was approved by Peking University Third Hospital Medical Science Research Ethics Committee (approval number: (2020)YiLun(310–02)), Peking University Institutional Review Board (approval number: IRB0001052-16009), Institutional Review Board of Guangdong Woman and Children's Hospital (approval number: YiLun [201701044]), Ethics Committee of the West China Second Hospital of Sichuan University (approval number: YiXueKeYan2020(064)) and Institutional Ethics Committee of Zhongnan Hospital of Wuhan University (approval number: 2015029 and 2020102). Informed consent was obtained from the donors and their guardians.

Competing interests
W.J., Q.C., H.H.G., X.J.L, J.S.L, and Y.Z. are current or former employees of Euler Technology, Beijing, China. Provisional patents were filed for the single-stranded NGS library preparation method Tequila 7N (WO 2020/073748) and the method for using cell-free DNA methylation pattern to predict placenta development status and pregnancy outcome (202010084924.X) by Euler Technology. These patents do not limit the use of the methods or reproducibility of the results. The other authors declare that they have no competing interests.

Author details
[1]Department of Obstetrics and Gynecology, Peking University Third Hospital, Beijing, China. [2]Medical Genetic Center, Guangdong Women and Children Hospital, Guangzhou, China. [3]Euler Technology, Beijing, China. [4]Department of Obstetrics and Gynecology, West China Second University Hospital of Sichuan University, Chengdu, China. [5]Department Key Laboratory of Birth Defects and Related Diseases of Women and Children (Sichuan University), Ministry of Education, Chengdu, China. [6]Department of Urology, Zhongnan Hospital of Wuhan University, Wuhan, China. [7]Department of Biological Repositories, Zhongnan Hospital of Wuhan University, Wuhan, China. [8]Human Genetic Resources Preservation Center of Hubei Province, Wuhan, China. [9]Laboratory of Precision Medicine, Zhongnan Hospital of Wuhan University, Wuhan, China. [10]Maternity Ward, Haidian Maternal and Child Health Hospital, Beijing, China. [11]Department of Obstetrics and Gynecology, Haidian Maternal and Child Health Hospital, Beijing, China. [12]Present Address: International Max Planck Research School for Genome Science, and University of Göttingen, Göttingen Center for Molecular Biosciences, Göttingen, Germany.

References
1. Abalos E, Cuesta C, Grosso AL, Chou D, Say L. Global and regional estimates of preeclampsia and eclampsia: a systematic review. Eur J Obstet Gynecol Reprod Biol. 2013;170:1–7.
2. Chappell LC, Cluver CA, Kingdom J, Tong S. Pre-eclampsia. Lancet. 2021;398:341–54.
3. Brosens I, Pijnenborg R, Vercruysse L, Romero R. The "Great Obstetrical Syndromes" are associated with disorders of deep placentation. Am J Obstet Gynecol. 2011;204:193–201.
4. Jones CJ, Fox H. An ultrastructural and ultrahistochemical study of the human placenta in maternal pre-eclampsia. Placenta. 1980;1:61–76.

5. Rodgers GM, Taylor RN, Roberts JM. Preeclampsia is associated with a serum factor cytotoxic to human endothelial cells. Am J Obstet Gynecol. 1988;159:908–14.
6. Brosens IA, Robertson WB, Dixon HG. The role of the spiral arteries in the pathogenesis of pre-eclampsia. J Pathol. 1970;101:Pvi.
7. Zhou Y, Damsky CH, Fisher SJ. Preeclampsia is associated with failure of human cytotrophoblasts to mimic a vascular adhesion phenotype. One cause of defective endovascular invasion in this syndrome? J Clin Invest. 1997;99:2152–64.
8. Cnattingius S, Reilly M, Pawitan Y, Lichtenstein P. Maternal and fetal genetic factors account for most of familial aggregation of preeclampsia: a population-based Swedish cohort study. Am J Med Genet A. 2004;130A:365–71.
9. Lie RT, Rasmussen S, Brunborg H, Gjessing HK, Lie-Nielsen E, Irgens LM. Fetal and maternal contributions to risk of pre-eclampsia: population based study. BMJ. 1998;316:1343–7.
10. Gray KJ, Kovacheva VP, Mirzakhani H, Bjonnes AC, Almoguera B, DeWan AT, Triche EW, Saftlas AF, Hoh J, Bodian DL, et al. Gene-centric analysis of preeclampsia identifies maternal association at PLEKHG1. Hypertension. 2018;72:408–16.
11. Hansen AT, Bernth Jensen JM, Hvas AM, Christiansen M. The genetic component of preeclampsia: a whole-exome sequencing study. PLoS One. 2018;13:e0197217.
12. McAllister KA, Grogg KM, Johnson DW, Gallione CJ, Baldwin MA, Jackson CE, Helmbold EA, Markel DS, McKinnon WC, Murrell J, et al. Endoglin, a TGF-beta binding protein of endothelial cells, is the gene for hereditary haemorrhagic telangiectasia type 1. Nat Genet. 1994;8:345–51.
13. McGinnis R, Steinthorsdottir V, Williams NO, Thorleifsson G, Shooter S, Hjartardottir S, Bumpstead S, Stefansdottir L, Hildyard L, Sigurdsson JK, et al. Variants in the fetal genome near FLT1 are associated with risk of preeclampsia. Nat Genet. 2017;49:1255–60.
14. Melton PE, Johnson MP, Gokhale-Agashe D, Rea AJ, Ariff A, Cadby G, Peralta JM, McNab TJ, Allcock RJ, Abraham LJ, et al. Whole-exome sequencing in multiplex preeclampsia families identifies novel candidate susceptibility genes. J Hypertens. 2019;37:997–1011.
15. Steinthorsdottir V, McGinnis R, Williams NO, Stefansdottir L, Thorleifsson G, Shooter S, Fadista J, Sigurdsson JK, Auro KM, Berezina G, et al. Genetic predisposition to hypertension is associated with preeclampsia in European and Central Asian women. Nat Commun. 2020;11:5976.
16. Zhao L, Bracken MB, DeWan AT. Genome-wide association study of pre-eclampsia detects novel maternal single nucleotide polymorphisms and copy-number variants in subsets of the Hyperglycemia and Adverse Pregnancy Outcome (HAPO) study cohort. Ann Hum Genet. 2013;77:277–87.
17. Haase N, Foster DJ, Cunningham MW, Bercher J, Nguyen T, Shulga-Morskaya S, Milstein S, Shaikh S, Rollins J, Golic M, et al. RNA interference therapeutics targeting angiotensinogen ameliorate preeclamptic phenotype in rodent models. J Clin Invest. 2020;130:2928–42.
18. Jeunemaitre X, Soubrier F, Kotelevtsev YV, Lifton RP, Williams CS, Charru A, Hunt SC, Hopkins PN, Williams RR, Lalouel JM, et al. Molecular basis of human hypertension: role of angiotensinogen. Cell. 1992;71:169–80.
19. Mistry HD, Kurlak LO, Gardner DS, Torffvit O, Hansen A, Broughton Pipkin F, Strevens H. Evidence of augmented intrarenal angiotensinogen associated with glomerular swelling in gestational hypertension and preeclampsia: clinical implications. J Am Heart Assoc. 2019;8:e012611.
20. Koga K, Osuga Y, Yoshino O, Hirota Y, Ruimeng X, Hirata T, Takeda S, Yano T, Tsutsumi O, Taketani Y. Elevated serum soluble vascular endothelial growth factor receptor 1 (sVEGFR-1) levels in women with preeclampsia. J Clin Endocrinol Metab. 2003;88:2348–51.
21. Levine RJ, Lam C, Qian C, Yu KF, Maynard SE, Sachs BP, Sibai BM, Epstein FH, Romero R, Thadhani R, et al. Soluble endoglin and other circulating antiangiogenic factors in preeclampsia. N Engl J Med. 2006;355:992–1005.
22. Maynard SE, Min JY, Merchan J, Lim KH, Li J, Mondal S, Libermann TA, Morgan JP, Sellke FW, Stillman IE, et al. Excess placental soluble fms-like tyrosine kinase 1 (sFlt1) may contribute to endothelial dysfunction, hypertension, and proteinuria in preeclampsia. J Clin Invest. 2003;111:649–58.
23. Rana S, Karumanchi SA, Levine RJ, Venkatesha S, Rauh-Hain JA, Tamez H, Thadhani R. Sequential changes in antiangiogenic factors in early pregnancy and risk of developing preeclampsia. Hypertension. 2007;50:137–42.
24. Venkatesha S, Toporsian M, Lam C, Hanai J, Mammoto T, Kim YM, Bdolah Y, Lim KH, Yuan HT, Libermann TA, et al. Soluble endoglin contributes to the pathogenesis of preeclampsia. Nat Med. 2006;12:642–9.
25. Ong CY, Liao AW, Spencer K, Munim S, Nicolaides KH. First trimester maternal serum free beta human chorionic gonadotrophin and pregnancy associated plasma protein A as predictors of pregnancy complications. BJOG. 2000;107:1265–70.
26. Chappell LC, Seed PT, Briley A, Kelly FJ, Hunt BJ, Charnock-Jones DS, Mallet AI, Poston L. A longitudinal study of biochemical variables in women at risk of preeclampsia. Am J Obstet Gynecol. 2002;187:127–36.
27. Taylor RN, Grimwood J, Taylor RS, McMaster MT, Fisher SJ, North RA. Longitudinal serum concentrations of placental growth factor: evidence for abnormal placental angiogenesis in pathologic pregnancies. Am J Obstet Gynecol. 2003;188:177–82.
28. Tidwell SC, Ho HN, Chiu WH, Torry RJ, Torry DS. Low maternal serum levels of placenta growth factor as an antecedent of clinical preeclampsia. Am J Obstet Gynecol. 2001;184:1267–72.
29. Thadhani R, Ecker JL, Kettyle E, Sandler L, Frigoletto FD. Pulse pressure and risk of preeclampsia: a prospective study. Obstet Gynecol. 2001;97:515–20.
30. Bosio PM, McKenna PJ, Conroy R, O'Herlihy C. Maternal central hemodynamics in hypertensive disorders of pregnancy. Obstet Gynecol. 1999;94:978–84.
31. Lafayette RA, Druzin M, Sibley R, Derby G, Malik T, Huie P, Polhemus C, Deen WM, Myers BD. Nature of glomerular dysfunction in pre-eclampsia. Kidney Int. 1998;54:1240–9.
32. Stillman IE, Karumanchi SA. The glomerular injury of preeclampsia. J Am Soc Nephrol. 2007;18:2281–4.
33. Barrientos G, Pussetto M, Rose M, Staff AC, Blois SM, Toblli JE. Defective trophoblast invasion underlies fetal growth restriction and preeclampsia-like symptoms in the stroke-prone spontaneously hypertensive rat. Mol Hum Reprod. 2017;23:509–19.

Gong *et al. Genome Biology*   (2024) 25:117

Page 42 of 46

34. Kumasawa K, Ikawa M, Kidoya H, Hasuwa H, Saito-Fujita T, Morioka Y, Takakura N, Kimura T, Okabe M. Pravastatin induces placental growth factor (PGF) and ameliorates preeclampsia in a mouse model. Proc Natl Acad Sci U S A. 2011;108:1451–5.

35. Naicker T, Khedun SM, Moodley J, Pijnenborg R. Quantitative analysis of trophoblast invasion in preeclampsia. Acta Obstet Gynecol Scand. 2003;82:722–9.

36. Bernstein BE, Meissner A, Lander ES. The mammalian epigenome. Cell. 2007;128:669–81.

37. Bird A. DNA methylation patterns and epigenetic memory. Genes Dev. 2002;16:6–21.

38. Ringrose L, Paro R. Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. Annu Rev Genet. 2004;38:413–43.

39. Xu Q, Xie W. Epigenome in early mammalian development: inheritance, reprogramming and establishment. Trends Cell Biol. 2018;28:237–53.

40. Xia W, Xu J, Yu G, Yao G, Xu K, Ma X, Zhang N, Liu B, Li T, Lin Z, et al. Resetting histone modifications during human parental-to-zygotic transition. Science. 2019;365:353–60.

41. Zhou F, Wang R, Yuan P, Ren Y, Mao Y, Li R, Lian Y, Li J, Wen L, Yan L, et al. Reconstituting the transcriptome and DNA methylome landscapes of human implantation. Nature. 2019;572:660–4.

42. Zhu P, Guo H, Ren Y, Hou Y, Dong J, Li R, Lian Y, Fan X, Hu B, Gao Y, et al. Single-cell DNA methylome sequencing of human preimplantation embryos. Nat Genet. 2018;50:12–9.

43. Ziller MJ, Gu H, Muller F, Donaghey J, Tsai LT, Kohlbacher O, De Jager PL, Rosen ED, Bennett DA, Bernstein BE, et al. Charting a dynamic DNA methylation landscape of the human genome. Nature. 2013;500:477–81.

44. De Iaco A, Planet E, Coluccio A, Verp S, Duc J, Trono D. DUX-family transcription factors regulate zygotic genome activation in placental mammals. Nat Genet. 2017;49:941–5.

45. Liu L, Leng L, Liu C, Lu C, Yuan Y, Wu L, Gong F, Zhang S, Wei X, Wang M, et al. An integrated chromatin accessibility and transcriptome landscape of human pre-implantation embryos. Nat Commun. 2019;10:364.

46. Lu F, Liu Y, Inoue A, Suzuki T, Zhao K, Zhang Y. Establishing chromatin regulatory landscape during mouse preimplantation development. Cell. 2016;165:1375–88.

47. Klymenko T, Papp B, Fischle W, Kocher T, Schelder M, Fritsch C, Wild B, Wilm M, Muller J. A Polycomb group protein complex with sequence-specific DNA-binding and selective methyl-lysine-binding activities. Genes Dev. 2006;20:1110–22.

48. Margueron R, Reinberg D. The polycomb complex PRC2 and its mark in life. Nature. 2011;469:343–9.

49. Zheng H, Huang B, Zhang B, Xiang Y, Du Z, Xu Q, Li Y, Wang Q, Ma J, Peng X, et al. Resetting epigenetic memory by reprogramming of histone modifications in mammals. Mol Cell. 2016;63:1066–79.

50. Smith ZD, Shi J, Gu H, Donaghey J, Clement K, Cacchiarelli D, Gnirke A, Michor F, Meissner A. Epigenetic restriction of extraembryonic lineages mirrors the somatic transition to cancer. Nature. 2017;549:543–7.

51. Brodowski L, Zindler T, von Hardenberg S, Schroder-Heurich B, von Kaisenberg CS, Frieling H, Hubel CA, Dork T, von Versen-Hoynck F. Preeclampsia-associated alteration of DNA methylation in fetal endothelial progenitor cells. Front Cell Dev Biol. 2019;7:32.

52. Cirkovic A, Garovic V, Milin Lazovic J, Milicevic O, Savic M, Rajovic N, Aleksic N, Weissgerber T, Stefanovic A, Stanisavljevic D, Milic N. Systematic review supports the role of DNA methylation in the pathophysiology of preeclampsia: a call for analytical and methodological standardization. Biol Sex Differ. 2020;11:36.

53. Cruz JO, Conceicao I, Tosatti JAG, Gomes KB, Luizon MR. Global DNA methylation in placental tissues from pregnant with preeclampsia: a systematic review and pathway analysis. Placenta. 2020;101:97–107.

54. 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.

55. Vento-Tormo R, Efremova M, Botting RA, Turco MY, Vento-Tormo M, Meyer KB, Park JE, Stephenson E, Polanski K, Goncalves A, et al. Single-cell reconstruction of the early maternal-fetal interface in humans. Nature. 2018;563:347–53.

56. Suryawanshi H, Morozov P, Straus A, Sahasrabudhe N, Max KEA, Garzia A, Kustagi M, Tuschl T, Williams Z. Placenta and decidua scRNA-seq. BioProject Accession: PRJNA492324; BioProject NCBI. 2018. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA492324/.

57. Vento-Tormo R, Efremova M, Botting RA, Turco MY, Vento-Tormo M, Meyer KB, Park J-E, Stephenson E, Polański K, Goncalves A, et al. Reconstructing the human first trimester fetal-maternal interface using single cell transcriptomics - 10x data. BioProject Accession: PRJEB28266; BioProject NCBI. 2018. https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJEB28266.

58. Yu L, Wei Y, Duan J, Schmitz DA, Sakurai M, Wang L, Wang K, Zhao S, Hon GC, Wu J. Blastocyst-like structures generated from human pluripotent stem cells. Nature. 2021;591:620–6.

59. Yu L, Wei Y, Duan J, Schmitz DA, Sakurai M, Wang L, Wang K, Zhao S, Hon GC, Wu J. Generation of blastocyst-like structures from human embryonic stem cells. GEO Accession: GSE150578; GEO NCBI. 2021. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE150578.

60. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, Baglaenko Y, Brenner M, Loh PR, Raychaudhuri S. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods. 2019;16:1289–96.

61. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol. 2018;36:411–20.

62. Sun D, Guan X, Moran AE, Wu LY, Qian DZ, Schedin P, Dai MS, Danilov AV, Alumkal JJ, Adey AC, et al. Identifying phenotype-associated subpopulations by integrating bulk and single-cell sequencing data. Nat Biotechnol. 2022;40:527–38.

63. Yang X, Yang J, Liang X, Chen Q, Jiang S, Liu H, Gao Y, Ren Z, Shi YW, Li S, et al. Landscape of dysregulated placental RNA editing associated with preeclampsia. Hypertension. 2020;75:1532–41.

64. Yang X, Yang J, Liang X, Chen Q, Jiang S, Liu H, Gao Y, Ren Z, Shi Y-W, Li S, et al. Placental editome profiling reveals widespread RNA-editing dysregulation in preeclampsia. GEO Accession: GSE148241; GEO NCBI. 2020. https://www.ncbi.nlm.nih.gov/gds/?term=GSE148241.

Gong *et al. Genome Biology*　(2024) 25:117

Page 43 of 46

65. Yu F, Cato LD, Weng C, Liggett LA, Jeon S, Xu K, Chiang CWK, Wiemels JL, Weissman JS, de Smith AJ, Sankaran VG. Variant to function mapping at single-cell resolution through network propagation. Nat Biotechnol. 2022;40:1644–53.

66. Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ. Generalizing RNA velocity to transient cell states through dynamical modeling. Nat Biotechnol. 2020;38:1408–14.

67. Tatiana B, Didier C, David R, Derek Y. mixtools: an R package for analyzing finite mixture models. J Stat Softw. 2009;32:1–29.

68. Gulati GS, Sikandar SS, Wesche DJ, Manjunath A, Bharadwaj A, Berger MJ, Ilagan F, Kuo AH, Hsieh RW, Cai S, et al. Single-cell transcriptional diversity is a hallmark of developmental potential. Science. 2020;367:405–11.

69. Xiao Y, et al. Tracking single-cell evolution using clock-like chromatin accessibility loci. Nat Biotechnol. 2024. https://doi.org/10.1038/s41587-024-02241-z.

70. Yu B, Yang Y, Zhou W, Zhang B. Single cell RNA-sequencing reveals placenta cellular heterogeneity in adverse pregnancy. GEO Accession: GSE173193; GEO NCBI. 2021. https://www.ncbi.nlm.nih.gov/gds/?term=GSE173193.

71. Angerer P, Haghverdi L, Buttner M, Theis FJ, Marr C, Buettner F. destiny: diffusion maps for large-scale single-cell data in R. Bioinformatics. 2016;32:1241–3.

72. Sinha S, Maity SN, Seldin MF, de Crombrugghe B. Chromosomal assignment and tissue expression of CBF-C/NFY-C, the third subunit of the mammalian CCAAT-binding factor. Genomics. 1996;37:260–3.

73. Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, Morey L, Van Eynde A, Bernard D, Vanderwinden JM, et al. The polycomb group protein EZH2 directly controls DNA methylation. Nature. 2006;439:871–4.

74. Atchison L, Ghias A, Wilkinson F, Bonini N, Atchison ML. Transcription factor YY1 functions as a PcG protein in vivo. EMBO J. 2003;22:1347–58.

75. Li H, Liefke R, Jiang J, Kurland JV, Tian W, Deng P, Zhang W, He Q, Patel DJ, Bulyk ML, et al. Polycomb-like proteins link the PRC2 complex to CpG islands. Nature. 2017;549:287–91.

76. Wu J, Xu J, Liu B, Yao G, Wang P, Lin Z, Huang B, Wang X, Li T, Shi S, et al. Chromatin analysis in human early development reveals epigenetic transition during ZGA. Nature. 2018;557:256–60.

77. Hecker N, Hiller M. A genome alignment of 120 mammals highlights ultraconserved element variability and placenta-associated enhancers. Gigascience. 2020;9:giz159.

78. Pollard KS, Salama SR, Lambert N, Lambot MA, Coppens S, Pedersen JS, Katzman S, King B, Onodera C, Siepel A, et al. An RNA gene expressed during cortical development evolved rapidly in humans. Nature. 2006;443:167–72.

79. Bejerano G, Pheasant M, Makunin I, Stephen S, Kent WJ, Mattick JS, Haussler D. Ultraconserved elements in the human genome. Science. 2004;304:1321–5.

80. Schep AN, Wu B, Buenrostro JD, Greenleaf WJ. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. Nat Methods. 2017;14:975–8.

81. Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH. Dnmt3L and the establishment of maternal genomic imprints. Science. 2001;294:2536–9.

82. Kaneda M, Okano M, Hata K, Sado T, Tsujimoto N, Li E, Sasaki H. Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature. 2004;429:900–3.

83. Lewis A, Mitsuya K, Umlauf D, Smith P, Dean W, Walter J, Higgins M, Feil R, Reik W. Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of DNA methylation. Nat Genet. 2004;36:1291–5.

84. Abed M, Verschueren E, Budayeva H, Liu P, Kirkpatrick DS, Reja R, Kummerfeld SK, Webster JD, Gierke S, Reichelt M, et al. The Gag protein PEG10 binds to RNA and regulates trophoblast stem cell lineage specification. PLoS One. 2019;14:e0214110.

85. Ono R, Nakamura K, Inoue K, Naruse M, Usami T, Wakisaka-Saito N, Hino T, Suzuki-Migishima R, Ogonuki N, Miki H, et al. Deletion of Peg10, an imprinted gene acquired from a retrotransposon, causes early embryonic lethality. Nat Genet. 2006;38:101–6.

86. Segel M, Lash B, Song J, Ladha A, Liu CC, Jin X, Mekhedov SL, Macrae RK, Koonin EV, Zhang F. Mammalian retrovirus-like protein PEG10 packages its own mRNA and can be pseudotyped for mRNA delivery. Science. 2021;373:882–9.

87. Hannula-Jouppi K, Muurinen M, Lipsanen-Nyman M, Reinius LE, Ezer S, Greco D, Kere J. Differentially methylated regions in maternal and paternal uniparental disomy for chromosome 7. Epigenetics. 2014;9:351–65.

88. Wei J, Lau SY, Blenkiron C, Chen Q, James JL, Kleffmann T, Wise M, Stone PR, Chamley LW. Trophoblastic debris modifies endothelial cell transcriptome in vitro: a mechanism by which fetal cells might control maternal responses to pregnancy. Sci Rep. 2016;6:30632.

89. Winkler EA, Kim CN, Ross JM, Garcia JH, Gil E, Oh I, Chen LQ, Wu D, Catapano JS, Raygor K, et al. A single-cell atlas of the normal and malformed human brain vasculature. Science. 2022;375:eabi7377.

90. Lux A, Beil C, Majety M, Barron S, Gallione CJ, Kuhn HM, Berg JN, Kioschis P, Marchuk DA, Hafner M. Human retroviral gag- and gag-pol-like proteins interact with the transforming growth factor-beta receptor activin receptor-like kinase 1. J Biol Chem. 2005;280:8482–93.

91. Koppes E, Himes KP, Chaillet JR. Partial loss of genomic imprinting reveals important roles for Kcnq1 and Peg10 imprinted domains in placental development. PLoS ONE. 2015;10:e0135202.

92. Shiura H, Ono R, Tachibana S, Kohda T, Kaneko-Ishino T, Ishino F. PEG10 viral aspartic protease domain is essential for the maintenance of fetal capillary structure in the mouse placenta. Development. 2021;148:dev199564.

93. Blobe GC, Schiemann WP, Lodish HF. Role of transforming growth factor beta in human disease. N Engl J Med. 2000;342:1350–8.

94. Li DY, Sorensen LK, Brooke BS, Urness LD, Davis EC, Taylor DG, Boak BB, Wendel DP. Defective angiogenesis in mice lacking endoglin. Science. 1999;284:1534–7.

95. Oh SP, Seki T, Goss KA, Imamura T, Yi Y, Donahoe PK, Li L, Miyazono K, ten Dijke P, Kim S, Li E. Activin receptor-like kinase 1 modulates transforming growth factor-beta 1 signaling in the regulation of angiogenesis. Proc Natl Acad Sci U S A. 2000;97:2626–31.

96. Sanford LP, Ormsby I, Gittenberger-de Groot AC, Sariola H, Friedman R, Boivin GP, Cardell EL, Doetschman T. TGFbeta2 knockout mice have multiple developmental defects that are non-overlapping with other TGFbeta knockout phenotypes. Development. 1997;124:2659–70.
97. Yu Q, Stamenkovic I. Cell surface-localized matrix metalloproteinase-9 proteolytically activates TGF-beta and promotes tumor invasion and angiogenesis. Genes Dev. 2000;14:163–76.
98. Park H, Furtado J, Poulet M, Chung M, Yun S, Lee S, Sessa WC, Franco CA, Schwartz MA, Eichmann A. Defective Flow-migration coupling causes arteriovenous malformations in hereditary hemorrhagic telangiectasia. Circulation. 2021;144:805–22.
99. Maharaj AS, Walshe TE, Saint-Geniez M, Venkatesha S, Maldonado AE, Himes NC, Matharu KS, Karumanchi SA, D'Amore PA. VEGF and TGF-beta are required for the maintenance of the choroid plexus and ependyma. J Exp Med. 2008;205:491–501.
100. RN D. Individual human cortical progenitors can produce excitatory and inhibitory neurons. GEO Accession: GSE187875; GEO NCBI. 2022. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE187875.
101. McCracken IR, Dobie R, Bennett M, Passi R, Beqqali A, Henderson NC, Mountford JC, Riley PR, Ponting CP, Smart N, et al. Mapping the developing human cardiac endothelium at single-cell resolution identifies MECOM as a regulator of arteriovenous gene expression. Cardiovasc Res. 2022;118:2960–72.
102. Yin X, Chan LS, Bose D, Jackson AU, VandeHaar P, Locke AE, Fuchsberger C, Stringham HM, Welch R, Yu K, et al. Genome-wide association studies of metabolites in Finnish men identify disease-relevant loci. Nat Commun. 2022;13:1644.
103. Tyrmi JS, Kaartokallio T, Lokki I, Jääskeläinen T, Kortelainen E, Ruotsalainen S, Karjalainen J, Ripatti S, Group FS, FinnGen: GWAS of preeclampsia and hypertensive disorders of pregnancy uncovers genes related to cardiometabolic, endothelial and placental function. medRxiv 2022:2022.2005. 2019.22275002.
104. James JL, Chamley LW, Clark AR. Feeding your baby in utero: how the uteroplacental circulation impacts pregnancy. Physiology (Bethesda). 2017;32:234–45.
105. Castellucci M, Scheper M, Scheffen I, Celona A, Kaufmann P. The development of the human placental villous tree. Anat Embryol (Berl). 1990;181:117–28.
106. Demir R, Kosanke G, Kohnen G, Kertschanska S, Kaufmann P. Classification of human placental stem villi: review of structural and functional aspects. Microsc Res Tech. 1997;38:29–41.
107. Hecht JL, Zsengeller ZK, Spiel M, Karumanchi SA, Rosen S. Revisiting decidual vasculopathy. Placenta. 2016;42:37–43.
108. Pijnenborg R, Dixon G, Robertson WB, Brosens I. Trophoblastic invasion of human decidua from 8 to 18 weeks of pregnancy. Placenta. 1980;1:3–19.
109. Thaler I, Manor D, Itskovitz J, Rottem S, Levit N, Timor-Tritsch I, Brandes JM. Changes in uterine blood flow during human pregnancy. Am J Obstet Gynecol. 1990;162:121–5.
110. Chua S, Wilkins T, Sargent I, Redman C. Trophoblast deportation in pre-eclamptic pregnancy. Br J Obstet Gynaecol. 1991;98:973–9.
111. Fisher SJ. Why is placentation abnormal in preeclampsia? Am J Obstet Gynecol. 2015;213:S115-122.
112. Huppertz B. Trophoblast differentiation, fetal growth restriction and preeclampsia. Pregnancy Hypertens. 2011;1:79–86.
113. Sasaki Y, Darmochwal-Kolarz D, Suzuki D, Sakai M, Ito M, Shima T, Shiozaki A, Rolinski J, Saito S. Proportion of peripheral blood and decidual CD4(+) CD25(bright) regulatory T cells in pre-eclampsia. Clin Exp Immunol. 2007;149:139–45.
114. Chuong EB. Retroviruses facilitate the rapid evolution of the mammalian placenta. BioEssays. 2013;35:853–61.
115. Gimenez J, Montgiraud C, Oriol G, Pichon JP, Ruel K, Tsatsaris V, Gerbaud P, Frendo JL, Evain-Brion D, Mallet F. Comparative methylation of ERVWE1/syncytin-1 and other human endogenous retrovirus LTRs in placenta tissues. DNA Res. 2009;16:195–211.
116. Matouskova M, Blazkova J, Pajer P, Pavlicek A, Hejnar J. CpG methylation suppresses transcriptional activity of human syncytin-1 in non-placental tissues. Exp Cell Res. 2006;312:1011–20.
117. Harris JR. Placental endogenous retrovirus (ERV): structural, functional, and evolutionary significance. BioEssays. 1998;20:307–16.
118. Johnson PM, Lyden TW, Mwenda JM. Endogenous retroviral expression in the human placenta. Am J Reprod Immunol. 1990;23:115–20.
119. Rote NS, Chakrabarti S, Stetzer BP. The role of human endogenous retroviruses in trophoblast differentiation and placental development. Placenta. 2004;25:673–83.
120. Abdulghani M, Jain A, Tuteja G. Genome-wide identification of enhancer elements in the placenta. Placenta. 2019;79:72–7.
121. Chuong EB, Rumi MA, Soares MJ, Baker JC. Endogenous retroviruses function as species-specific enhancer elements in the placenta. Nat Genet. 2013;45:325–9.
122. Deniz O, Ahmed M, Todd CD, Rio-Machin A, Dawson MA, Branco MR. Endogenous retroviruses are a source of enhancers with oncogenic potential in acute myeloid leukaemia. Nat Commun. 2020;11:3506.
123. Sun MA, Wolf G, Wang Y, Senft AD, Ralls S, Jin J, Dunn-Fletcher CE, Muglia LJ, Macfarlan TS. Endogenous retroviruses drive lineage-specific regulatory evolution across primate and rodent placentae. Mol Biol Evol. 2021;38:4992–5004.
124. Dupressoir A, Vernochet C, Harper F, Guegan J, Dessen P, Pierron G, Heidmann T. A pair of co-opted retroviral envelope syncytin genes is required for formation of the two-layered murine placental syncytiotrophoblast. Proc Natl Acad Sci U S A. 2011;108:E1164-1173.
125. Esnault C, Priet S, Ribet D, Vernochet C, Bruls T, Lavialle C, Weissenbach J, Heidmann T. A placenta-specific receptor for the fusogenic, endogenous retrovirus-derived, human syncytin-2. Proc Natl Acad Sci U S A. 2008;105:17532–7.
126. Roberts RM, Ezashi T, Schulz LC, Sugimoto J, Schust DJ, Khan T, Zhou J. Syncytins expressed in human placental trophoblast. Placenta. 2021;113:8–14.

Gong *et al. Genome Biology*   (2024) 25:117

Page 45 of 46

127. Soygur B, Sati L. The role of syncytins in human reproduction and reproductive organ cancers. Reproduction. 2016;152:R167-178.

128. Vargas A, Moreau J, Landry S, LeBellego F, Toufaily C, Rassart E, Lafond J, Barbeau B. Syncytin-2 plays an important role in the fusion of human trophoblast cells. J Mol Biol. 2009;392:301–18.

129. Wang YN, Ye Y, Zhou D, Guo ZW, Xiong Z, Gong XX, Jiang SW, Chen H. The role of syncytin in placental angiogenesis and fetal growth. Front Cell Dev Biol. 2022;10:852561.

130. Cohen M, Larsson E. Human endogenous retroviruses. BioEssays. 1988;9:191–6.

131. Cornelis G, Heidmann O, Bernard-Stoecklin S, Reynaud K, Veron G, Mulot B, Dupressoir A, Heidmann T. Ancestral capture of syncytin-Car1, a fusogenic endogenous retroviral envelope gene involved in placentation and conserved in Carnivora. Proc Natl Acad Sci U S A. 2012;109:E432-441.

132. Dunlap KA, Palmarini M, Varela M, Burghardt RC, Hayashi K, Farmer JL, Spencer TE. Endogenous retroviruses regulate periimplantation placental growth and differentiation. Proc Natl Acad Sci U S A. 2006;103:14390–5.

133. Dupressoir A, Vernochet C, Bawa O, Harper F, Pierron G, Opolon P, Heidmann T. Syncytin-A knockout mice demonstrate the critical role in placentation of a fusogenic, endogenous retrovirus-derived, envelope gene. Proc Natl Acad Sci U S A. 2009;106:12127–32.

134. Frendo JL, Olivier D, Cheynet V, Blond JL, Bouton O, Vidaud M, Rabreau M, Evain-Brion D, Mallet F. Direct involvement of HERV-W Env glycoprotein in human trophoblast cell fusion and differentiation. Mol Cell Biol. 2003;23:3566–74.

135. Harris JR. The evolution of placental mammals. FEBS Lett. 1991;295:3–4.

136. Heidmann O, Vernochet C, Dupressoir A, Heidmann T. Identification of an endogenous retroviral envelope gene with fusogenic activity and placenta-specific expression in the rabbit: a new "syncytin" in a third order of mammals. Retrovirology. 2009;6:107.

137. Mi S, Lee X, Li X, Veldman GM, Finnerty H, Racie L, LaVallie E, Tang XY, Edouard P, Howes S, et al. Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. Nature. 2000;403:785–9.

138. Villarreal LP. On viruses, sex, and motherhood. J Virol. 1997;71:859–65.

139. Haig D. Retroviruses and the placenta. Curr Biol. 2012;22:R609-613.

140. Schrader M, Jarrett BJM, Kilner RM. Parental care and sibling competition independently increase phenotypic variation among burying beetle siblings. Evolution. 2018;72:2546–52.

141. Snell-Rood EC, Van Dyken JD, Cruickshank T, Wade MJ, Moczek AP. Toward a population genetic framework of developmental evolution: the costs, limits, and consequences of phenotypic plasticity. BioEssays. 2010;32:71–81.

142. Brind'Amour J, Kobayashi H, Richard Albert J, Shirane K, Sakashita A, Kamio A, Bogutz A, Koike T, Karimi MM, Lefebvre L, et al. LTR retrotransposons transcribed in oocytes drive species-specific and heritable changes in DNA methylation. Nat Commun. 2018;9:3331.

143. Youngson NA, Kocialkowski S, Peel N, Ferguson-Smith AC. A small family of sushi-class retrotransposon-derived genes in mammals and their relation to genomic imprinting. J Mol Evol. 2005;61:481–90.

144. Gestational Hypertension and Preeclampsia. ACOG practice bulletin summary, number 222. Obstet Gynecol. 2020;135:1492–5.

145. Arnaoutova I, Kleinman HK. In vitro angiogenesis: endothelial cell tube formation on gelled basement membrane extract. Nat Protoc. 2010;5:628–35.

146. Ma S, Zhang B, LaFave LM, Earl AS, Chiang Z, Hu Y, Ding J, Brack A, Kartha VK, Tay T, et al. Chromatin potential identified by shared single-cell profiling of RNA and chromatin. Cell. 2020;183:1103-1116 e1120.

147. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol. 2016;34:525–7.

148. McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. Cell Syst. 2019;8:329-337 e324.

149. Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, Chak S, Naikawadi RP, Wolters PJ, Abate AR, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. Nat Immunol. 2019;20:163–72.

150. Ushey K, Allaire J, Tang Y. reticulate: interface to'Python', 2020. R package version, 1.

151. Herrmann C, Van de Sande B, Potier D, Aerts S. i-cisTarget: an integrative genomics method for the prediction of regulatory features and cis-regulatory modules. Nucleic Acids Res. 2012;40:e114.

152. Bravo Gonzalez-Blas C, Minnoye L, Papasokrati D, Aibar S, Hulselmans G, Christiaens V, Davie K, Wouters J, Aerts S. cisTopic: cis-regulatory topic modeling on single-cell ATAC-seq data. Nat Methods. 2019;16:397–400.

153. Giansanti V, Tang M, Cittaro D. Fast analysis of scATAC-seq data using a predefined set of genomic regions. F1000Res. 2020;9:199.

154. Granja JM, Corces MR, Pierce SE, Bagdatli ST, Choudhry H, Chang HY, Greenleaf WJ. ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. Nat Genet. 2021;53:403–11.

155. Reilly SK, Yin J, Ayoub AE, Emera D, Leng J, Cotney J, Sarro R, Rakic P, Noonan JP. Evolutionary genomics. Evolutionary changes in promoter and enhancer activity during human corticogenesis. Science. 2015;347:1155–9.

156. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34:i884–90.

157. Molaro A, Hodges E, Fang F, Song Q, McCombie WR, Hannon GJ, Smith AD. Sperm methylation profiles reveal features of epigenetic inheritance and evolution in primates. Cell. 2011;146:1029–41.

158. Zhu P, Guo H, Ren Y, Hou Y, Yan L, Qiao J, Tang F: Single cell DNA methylome sequencing of human preimplantation embryos. GEO Accession: GSE81233; GEO NCBI. 2017. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE81233.

159. Haghverdi L, Lun ATL, Morgan MD, Marioni JC. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. Nat Biotechnol. 2018;36:421–7.

160. Loh PR, Danecek P, Palamara PF, Fuchsberger C, A Reshef Y, K Finucane H, Schoenherr S, Forer L, McCarthy S, Abecasis GR, et al. Reference-based phasing using the Haplotype Reference Consortium panel. Nat Genet. 2016;48:1443–8.

Gong *et al. Genome Biology*   (2024) 25:117

Page 46 of 46

161. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15–21.

162. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550.

163. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A. Fast gene set enrichment analysis. BioRxiv. 2016:060012. https://bioconductor.org/packages/release/bioc/html/fgsea.html.

164. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102:15545–50.

165. Gong X, He W, Jin W, Ma H, Wang G, Li J, Xiao Y, Zhao Y, Chen Q, Guo H, et al. Dataset: disruption of maternal vascular remodeling by a fetal endoretrovirus-derived gene in preeclampsia. Accession: HRA001423; GSA Human in NGDC. 2024. https://ngdc.cncb.ac.cn/gsa-human/browse/HRA001423.

166. Gong X, He W, Jin W, Ma H, Wang G, Li J, Xiao Y, Zhao Y, Chen Q, Guo H, et al. Source code for preeclampsia single cell study paper. GitHub. 2023. https://github.com/MagpiePKU/Preeclampsia_Paper_2023.

167. Gong X, He W, Jin W, Ma H, Wang G, Li J, Xiao Y, Zhao Y, Chen Q, Guo H, et al. Source data for preeclampsia single cell study paper. Zenodo. 2023. https://zenodo.org/record/8079129.

168. Bernstein BE, Stamatoyannopoulos JA, Costello JF, Ren B, Milosavljevic A, Meissner A, Kellis M, Marra MA, Beaudet AL, Ecker JR, et al. The NIH roadmap epigenomics mapping consortium. Nat Biotechnol. 2010;28:1045–8.

169. Bernstein BE, Stamatoyannopoulos JA, Costello JF, Ren B, Milosavljevic A, Meissner A, Kellis M, Marra MA, Beaudet AL, Ecker JR, et al. UCSF-UBC Human Reference Epigenome Mapping Project. GEO Accession: GSE16368; GEO NCBI. 2009. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE16368.

170. Liu L, Leng L, Liu C, Lu C, Yuan Y, Wu L, Gong F, Zhang S, Wei X, Wang M, et al. An integrated chromatin accessibility and transcriptome landscape of human pre-implantation embryos. SRA Accession: SRP163205; SRA NCBI. 2018. https://trace.ncbi.nlm.nih.gov/Traces/?view=study&acc=SRP163205.

171. Wu J, Liu B, Xu J, Sun Y, Xie W. Analysis of chromatin landscapes in early human development reveals epigenetic transition during ZGA. GEO Accession: GSE101571; GEO NCBI. 2018. https://www.ncbi.nlm.nih.gov/gds/?term=GSE101571.

172. Jebbink J, Keijser R, Veenboer G, van der Post J, Ris-Stalpers C, Afink G. Expression of placental FLT1 transcript variants relates to both gestational hypertensive disease and fetal growth. Hypertension. 2011;58:70–6.

173. Rumer KK, Uyenishi J, Hoffman MC, Fisher BM, Winn VD. Siglec-6 expression is increased in placentas from pregnancies complicated by preterm preeclampsia. Reprod Sci. 2013;20:646–53.

174. Gormley M, Oliverio O, Kapidzic M, Ona K, Hall S, Fisher SJ. RNA profiling of laser microdissected human trophoblast subtypes at mid-gestation reveals a role for cannabinoid signaling in invasion. Development. 2021;148:20.

175. Liu Y, Fan X, Wang R, Lu X, Dang YL, Wang H, Lin HY, Zhu C, Ge H, Cross JC, Wang H. Single-cell RNA-seq reveals the diversity of trophoblast subtypes and patterns of differentiation in the human placenta. Cell Res. 2018;28:819–32.

176. Schupp JC, Adams TS, Cosme C Jr, Raredon MSB, Yuan Y, Omote N, Poli S, Chioccioli M, Rose KA, Manning EP, et al. Integrated single-cell atlas of endothelial cells of the human lung. Circulation. 2021;144:286–302.

177. Nikolova-Krstevski V, Yuan L, Le Bras A, Vijayaraj P, Kondo M, Gebauer I, Bhasin M, Carman CV, Oettgen P. ERG is required for the differentiation of embryonic stem cells along the endothelial lineage. BMC Dev Biol. 2009;9:1–14.

178. Medjkane S, Perez-Sanchez C, Gaggioli C, Sahai E, Treisman R. Myocardin-related transcription factors and SRF are required for cytoskeletal dynamics and experimental metastasis. Nat Cell Biol. 2009;11:257–68.

179. Katz FE, Tindle R, Sutherland DR, Greaves MF. Identification of a membrane glycoprotein associated with haemopoietic progenitor cells. Leuk Res. 1985;9:191–8.

Publisher's Note