Dysregulated trophoblast-specific gene expression mediated by retroviral regulatory sequences contributes to preeclampsia (PE)

Dissertation

zur Erlangung des akademischen Grades

Doctor of Philosophy (PhD)

im Fach Biologie

eingereicht an der

Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin

von

Master of Science

Rabia Anwar

Präsident der Humboldt-Universität zu Berlin

Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der der Lebenswissenschaftlichen Fakultät

Prof. Dr. Bernhard Grimm

Gutachter:

1. Prof. Dr. Ana Pombo
2. Prof. Dr. Kai Schmidt-Ott
3. Dr. Zsuzsanna Izsvák

Tag der mündlichen Prüfung: 30th September 2020

Content

Summary

Dysregulated trophoblast-specific gene expression mediated by retroviral regulatory sequences contributes to preeclampsia (PE)

Preeclampsia (PE) is a complication that occurs during pregnancy and affects almost 2-8% of all pregnancies and is often regarded as a human-specific disorder.[1,2] PE is one of the major causes of maternal and fetal death.[1] Failure of the trophoblast cells to invade into the maternal decidua results in the improper remodeling of spiral arteries leading to PE pathogenesis (Figure 1).[3] Clinically, it is diagnosed as a maternal syndrome, diagnosed by the new-onset of hypertension and proteinuria or other end-organ dysfunction after the 20th week of pregnancy. So far, the only effective treatment of the disorder is the removal of the placenta tissue and delivery of the infant.[4]

PE is a disorder of defective placentation and various factors (genetic and epigenetic) have been implicated in the abnormal placental development. Epigenetic modifications, such as low DNA methylation, as observed in the human placenta tissue increases the expression of human endogenous retroviruses (ERVs) essential for the placenta development.[5,6] Syncytin-1 (ERVWE1) derived from a human endogenous retrovirus HERV-W has an essential role in human placenta function and development including trophoblast syncytial formation and immunosuppression.[7,8] Low expression of Syncytin-1 due to hypermethylation in PE placentas has been detected.[9] There are several other examples of genes that are particularly expressed in the human placenta tissue due to the co-option of solitary endogenous retroviral (ERV) LTRs (Long terminal repeats) into the regulatory region that drive their placenta-specific gene regulation, either by contributing a unique enhancer or an alternative promoter.[10]

The aim of this study is to identify additional genes that are regulated by the human ERV-LTRs in the human placenta specifically, and are dysregulated in PE. To achieve this aim, the transcriptome of primary human trophoblast cells of 5 healthy and 5 early-onset PE placentas were analyzed by RNA sequencing (RNA-seq). RNA-seq analysis identified genes (n=335) with stronger expression in the trophoblast cells as compared to other human body tissues. Additionally, some of the genes (n=88) showed co-regulation of expression by the human ERV-LTRs in their vicinity (10-kb upstream of transcription start side (TSS) of the gene). Since my interest was to identify the new targets of PE pathogenesis, so I focused on genes (n=16) with dysregulated expression in women presented with PE.

Quantitative polymerase chain reaction (q-PCR) was performed on human PE placenta samples of Oslo-cohort-II to confirm the dysregulation of genes. *CYP11A1, EPS8L1, CSF2RB, SPINT1, ALDH3B2* and *DACT2* were identified to be significantly dysregulated in early-onset PE patients (Oslo-cohort-II). GFP-reporter assay confirmed the enhancer activity of the ERV-LTRs associated with the six candidate genes. To gain deeper insight into the functional analysis of the genes in PE pathogenesis, I focused on *EPS8L1*. *EPS8L1* is an epidermal growth factor receptor pathway substrate 8-related protein 1. I found it to be significantly upregulated in PE patients of Oslo-cohort-II, both at the level of mRNA and protein. The upregulation at the mRNA level was also seen in PE patients of Charite and Kiel cohort. Moreover, I also detected the EPS8L1 upregulation in the serum of PE patient of a high-risk pregnancy cohort from Manchester. These analyses made EPS8L1 a strong candidate to study further the function of this protein in PE pathogenesis. EPS8L1 expression was found to be consistent through-out pregnancy, by analyzing samples from early gestation (n=92) and at term (n=49) of pregnant women of Graz cohort. The analysis confirmed that the upregulation of EPS8L1 is PE specific and not related to the gestational-age changes.

To gain insight in the function, I established an *in-vitro EPS8L1* overexpressing trophoblast cell line by using *Sleeping Beauty (SB)* transposon system to mimic the *in-vivo* conditions of *EPS8L1* overexpression in PE patients. Global transcriptome changes by *EPS8L1* overexpression were analyzed by RNA sequencing. The differential gene expression analysis revealed dysregulation of important pathways involved in trophoblast function like invasion, proliferation, angiogenesis and maintenance of cell redox homeostasis upon overexpression of *EPS8L1*. Mass spectrometry analysis further confirmed that EPS8L1 interacts with different proteins involved in pathways that are important for placental development.

This study identified a new gene *EPS8L1,* regulated by primate-specific ERV-LTR in trophoblast cells that has a predominant role in the human placenta development and demonstrated that its dysregulation affected multiple pathways involved in trophoblast function like invasion, angiogenesis and maintenance of cell redox homeostasis. Furthermore, this study leads to the better understanding of the disease by explaining certain aspects of human-specific nature of PE.

Keywords: Preeclampsia (PE), Transposable elements (TEs), Endogenous Retroviruses (ERVs), Long Terminal Repeats (LTRs), Trophoblast-Specific Genes (TSGs)

Zusammenfassung

Eine dysregulierte Trophoblasten-spezifische Genexpression, die durch retrovirale regulatorische Sequenzen vermittelt wird, trägt zur Präeklampsie (PE) bei.

Präeklampsie (PE) ist eine Komplikation, die während der Schwangerschaft auftritt, fast 2-8% aller Schwangerschaften betrifft und human spezifisch ist.[1,2] PE ist eine der Hauptursachen für den Tod von Mutter und Kind.[1] Eine abnormale Plazentaentwicklung aufgrund einer verminderten Trophoblasteninvasion und einem gestörten Umbau der Spiralarterien trägt zur Pathogenese der PE bei (Figure 1).[3] Klinisch wird die PE durch Bluthochdruck und Proteinurie, auftretendnach der 20. Schwangerschaftswoche, diagnostiziert und kann durch eine Funktionsstörung von Organen begleitet werden. Bei besonders schweren Verläufen ist die frühzeitige Endbindung die letzte Möglichkeit das Überleben der Mutter zu gewährleisten.[4]

PE basiert auf einer gestörten Plazentation und eine Vielzahl an genetischen, sowie epigenetischen Faktoren wurde für eine abnormale Plazentaentwicklung beschrieben. Epigenetische Modifikationen, wie die DNA-Hypomethylierung verstärken die Expression von endogenen Retrovirus (ERVs) im menschlichen Plazentagewebe.[5,6] Syncytin-1 (ERVWE1), das aus einem humanen endogenen Retrovirus HERV-W stammt, spielt durch die Bildung von Trophoblasten-Syncytien und der Immunsuppression eine wichtige Rolle bei der Entwicklung und Funktion der menschlichen Plazenta.[7,8] Eine verminderte Expression von Syncytin-1 wird durch Hypermethylierung in der PE-Plazenta verursacht.[9] Weitere Gene werden aufgrund von co-optionen solitärer Retrovirus-LTRs in deren regulatorischen Regionen human spezifisch im Plazentagewebe exprimiert. Dies geschieht entweder durch die Bereitstellung eines neuen Enhancers oder eines alternativen Promotors.[10]

Das Ziel dieser Studie ist es, weitere Gene zu identifizieren, die durch ERVs in der menschlichen Plazenta spezifisch reguliert werden und in PE dysreguliert sind. Um dieses Ziel zu erreichen, wurde das Transkriptom von primären menschlichen Trophoblastenzellen von 5 gesunden und 5 früh einsetzenden PE-Plazenten mittels RNA-Sequenzierung analysiert. Es wurden 335 Gene identifiziert, welche eine höhere Expression in den Trophoblastenzellen im Vergleich zu anderen Geweben aufwiesen. Zusätzlich zeigten einige der Gene (n=88) eine Co-Regulation der Expression durch retrovirale LTRs (10-kb 5‘ des transcription start side (TSS) des Gens). Hauptinteresse lag hierbei auf den Genen, welche ebenfalls eine Dysregulation in der PE aufwiesen (n = 16).

Quantitative Polymerasekettenreaktion (q-PCR) wurde an humanen PE-Plazentaproben der Oslo-Kohorte durchgeführt, um die Dysregulation von Genen zu bestätigen. Es wurde festgestellt, dass die Gene *CYP11A1, EPS8L1, CSF2RB, SPINT1, ALDH3B2* und *DACT2* bei Frauen mit früh einsetzender PE signifikant dysreguliert sind. Ein GFP-Reporter-Assay bestätigte die Enhancer-Aktivität von 6 Genen welche assoziiert waren mit Retrovirus-LTRs.

Im folgenden konzentrierte ich mich auf das Gen *EPS8L1 (Epidermal growth factor receptor kinase substrate 8-like protein 1)*. Seine Funktion ist nicht bekannt. Ich konnte zeigen, dass es in Plazenta Proben von Frauen mit PE sowohl auf mRNA-, als auch auf Proteinebene signifikant hochreguliert ist. Diese Hochregulation wurde in zwei weiteren Kohorten verifiziert. Darüber hinaus konnte ich auch eine Hochregulation von EPS8L1 im Serum von Frauen mit PE detektieren. Die EPS8L1-Expression erwies sich während der gesamten Schwangerschaft als konsistent, wie anhand von Plazenten aus der frühen Schwangerschaft (n = 92) gezeigt werden konnte. Die Analyse bestätigte, dass die Hochregulation von EPS8L1 PE-spezifisch ist und nicht mit den Veränderungen des Gestationsalters zusammenhängt.

Eine EPS8L1-überexprimierende Trophoblasten-Zelllinie wurde mittels des *Sleeping Beauty* (SB) -Transposonsystem etabliert. Die globalen Transkriptomveränderungen durch die *EPS8L1*-Überexpression wurden mittels RNA-Sequenzierung analysiert und ergab eine Dysregulation wichtiger Signalwegen, wie Invasion, Proliferation, Angiogenese und Aufrechterhaltung der Zellredoxhomöostase. Eine Massenspektrometrie-Analyse bestätigte ferner, dass EPS8L1 mit verschiedenen Proteinen interagiert, welche für die Plazentaentwicklung wichtig sind.

Diese Studie identifizierte *EPS8L1*, das durch primaten-spezifisches ERV-LTR (MLT1G1) in Trophoblastenzellen reguliert wird, als einen wichtigen Faktor in der Entwicklung der menschlichen Plazenta. EPS8L1 ist in der PE Plazenta dysreguliert und involviert in mehrere Signalwege und die Funktionalität von Trophoblasten wie Invasion, Angiogenese und Redoxhomöostase. Hierdurch führt diese Arbeit zu einem besseren Verständnis der PE und deren human-spezifischer Natur.

1. Background and Introduction

1.1. Preeclampsia (PE)

Preeclampsia (PE) is one of the obstetrical complications characterized as the new emergence of hypertension (>140/90 mmHg) that develops during the second half of gestation, followed by excessive proteins in the urine (>300 mg/l protein in a 24 hr urine collection) or any of the features of end-organ injury, including low blood platelet count (<100,000), compromised liver function, kidney failure (increased serum creatinine >1.1 mg/dL), excessive fluid in the lungs or the start of cerebral or visual disturbances.[11] PE is one of the foremost cause of maternofetal mortality and morbidity.[1] Placenta tissue is the primary source of PE development.[12] The symptoms disappear when the placenta is delivered. Defects in the deep invasion of the trophoblasts into the maternal blood vessels is the key factor in PE development (Figure 1). The existence of PE has only been observed in humans and not in other mammals or non-human primates.[2] To understand the pathophysiology of PE, it is important to know the normal/ healthy placenta development and the cell composition of the placenta tissue, which is explained further in detail.

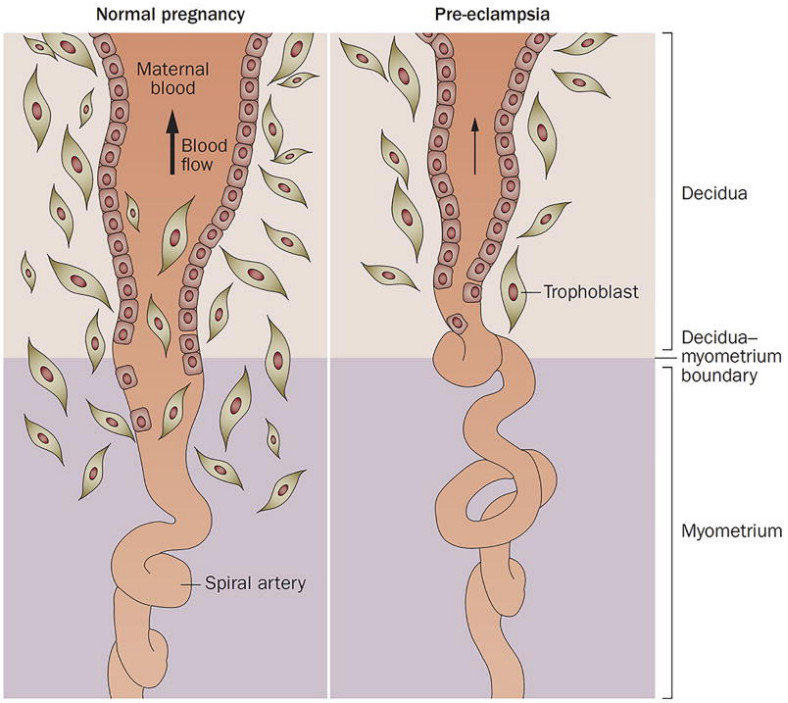

Figure 1. Non-transformed spiral arteries in PE pregnancy

In PE placentas, Extravillous trophoblasts (EVTs) fail to invade and transform the myometrial segment of the spiral arteries, resulting in narrow arteries and decreased utero-placental blood flow. In contrast, in a healthy placenta, the trophoblasts increase the vasculature of spiral

arteries by invading into the one-third of the myometrium, resulting in proper placental perfusion. Adapted from[13]

1.2. Healthy pregnancy

Pregnancy is the time from fertilization to child birth, lasting for about 38 weeks in humans.[14] The time span of the pregnancy varies between different species. A healthy placenta is a pre-requisite for a healthy pregnancy outcome. Placenta functions as an essential feto-maternal unit and controls the growth and development of the fetus. Placentas are quite diverse among mammals based on their morphology, cell type and endocrine function. For example the placentas of humans, rodents, and ruminants (sheep and cattle) are all very different. Humans have hemochorial placentation; the fetal trophoblasts are in contact with the maternal blood similar to the mouse/rat placenta but it is highly invasive.[15] The human placenta when delivered has a disc-like shape and a weight of about 470 grams. It normally has a diameter of 22 cm and a central thickness of around 2–2.5 cm. The fetus is connected with the placenta by an umbilical cord of around 55–60 cm in length, which has two arteries and a vein.[16] A healthy placenta expresses around 70% of the total protein coding genes in human cells.[17] Some of these genes have higher expression (350 genes) and a few (100 genes) were exclusively detected in the human placenta tissue. Placenta-specific genes are implicated in the human placenta evolution and sustenance of pregnancy. Abnormal development of placenta is the predominant feature of major disorders of pregnancy such as preeclampsia (PE), intrauterine growth restriction (IUGR), recurrent miscarriage and preterm birth (PTB).[18]

1.2.1. Placenta development in a healthy pregnancy

After fertilization (6-10 days), the blastocyst, also known as pre-implantation embryo, attaches to the maternal endometrium, initiating the process of placentation. This phase of placental development is called as prelacunar.[19] The pre-implantation embryo consists of an inner cell mass (ICM) that is surrounded by the trophectoderm layer. The embryo and the umbilical cord are generated from the ICM and the trophectoderm leads to the establishment of the fetal membranes and the placenta tissue. Following implantation, the trophoblast cells in the trophectoderm layer differentiates into syncytiotrophoblasts, that invades through the surface epithelium and transforms the endometrium into decidua.[20] Fluid filled spaces also known as Lacunae appear, after 8 days of conception, within the syncytiotrophoblasts layer. The cells around the lacunae (trabeculae) infiltrate into the wall of the decidua leading to the establishment of villous trees of the placenta tissue.[21] This is called the lacunar stage.

The cytotrophoblast cells underlying the syncytiotrophoblasts, after 12 days of conception, proliferate exponentially to form the projections in the primary syncytium ultimately establishing the direct contact with the maternal tissue and transforming the projections into primary villi. Soon after the day 17th, mesenchymal cells follow the cytotrophoblast and penetrate into trabeculae to form secondary villi. Trophoblastic cell columns (CCC) filled with cytotrophoblast are established at the end of trabeculae and the mesodermal cells are unable to reach the maternal side of the trabeculae. First fetal capillaries appear; after 18 days of fertilization that converts the corresponding villi into tertiary villi. The villous tree grows rapidly all through the gestation but the proportion of newly established primary and secondary villi structures in a placenta at term is lower as compared to the fully functional tertiary villi. Cytotrophoblast cells of the cell column (CCC) located at the decidua-placenta interface invades into decidua as extra villous trophoblasts (EVTs). The maternal blood is supplied to the developing placenta until the 12th week of pregnancy and the placenta tissue continues to grow throughout the gestation.

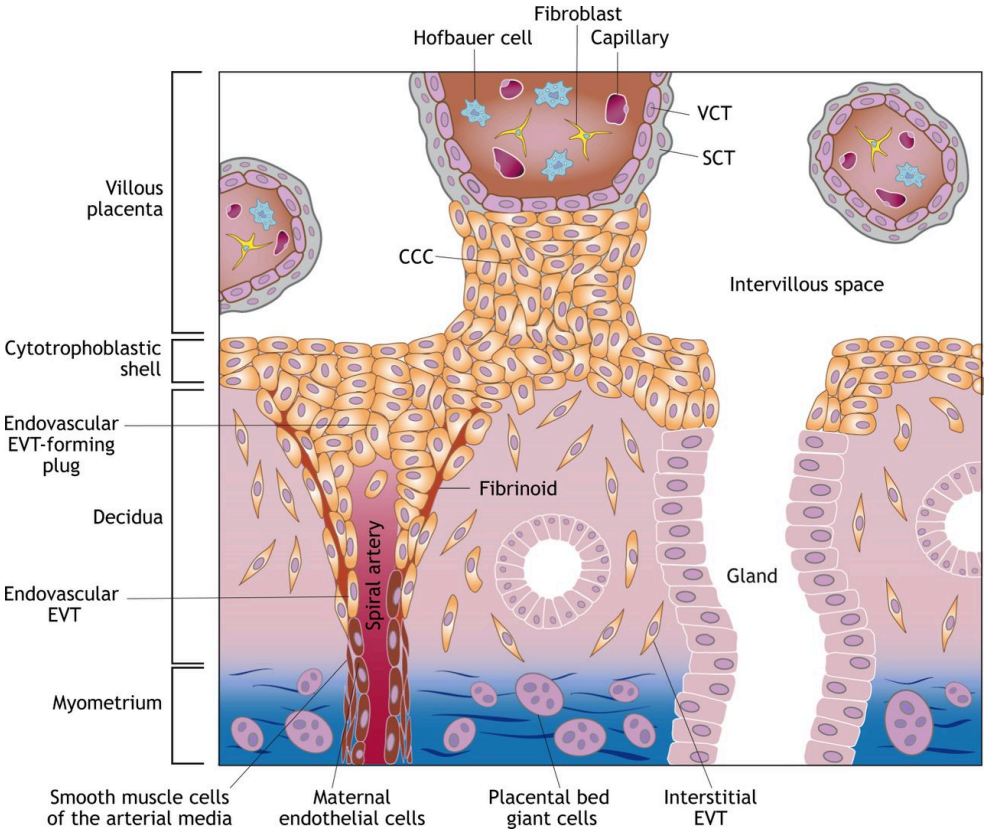

Figure 2. First trimester villi in a healthy pregnancy
Human placental villi attached to the maternal decidua and the major trophoblast subtypes are shown. Extravillous trophoblasts (EVT) including endovascular and interstitial EVTs, syncytiotrophoblasts (SCT) and villous cytotrophoblast (VCT) are shown. Adapted from[22]

1.2.2. Placenta tissue composition

The placenta is a heterogeneous tissue comprising of numerous cell types that are involved in highly diversified functions ranging from adhesion, migration, invasion, remodeling of blood vessels, production of pregnancy associated hormones, metabolism and interchange of gases and nutrients between the mother and the baby.[23] The functional unit of human placenta is the trophoblast. In 1889, the word 'trophoblast' was first termed by Ambrosius Arnold Willem Hubrecht. He described them as cells that are involved in the transportation of the nutrients and maintain a protective shield between the mother and the developing embryo.[24] Trophoblast cells are further divided into three different subtypes: these include the syncytiotrophoblast (SCT), the villous cytotrophoblast (VCT) and the extravillous trophoblast (EVT) (Figure 2).

1.2.2.1. Syncytiotrophoblast (SCT)

Syncytiotrophoblasts (SCTs) are the cells that form the outer lining of the placental villi and are in close proximity with the maternal blood and the uterine gland (Figure 2). SCTs have more than one nucleus and are terminally differentiated. All the gaseous and nutrient interchange between the mother and the fetus is carried out by these cells. The SCTs have an endocrine function, they secrete hormones and proteins into the maternal circulation that are crucial for the maintenance of a healthy pregnancy. Moreover, they also act as an immunological shield and protect the allogeneic fetus from the response generated by maternal immune system. Since SCTs does not express any human leukocyte antigen (HLA) molecules, they escape the response from maternal immune cells.[20,25]

1.2.2.2. Villous cytotrophoblast (VCT)

The cells directly underneath the multinucleated syncytial layer are the mononucleated villous cytotrophoblast (VCT) (Figure 2). Due to their extremely proliferative and mitotic capacity, they are considered as trophoblast stem cells.[26] VCT differentiate into either fusion lineage and lead to SCT formation or they differentiate into invasive interstitial cytotrophoblast cells.[27] As pregnancy progresses, the number of VCT declines and at the time of delivery they are only about 25% of the entire amount of the VCT in the placenta tissue.

1.2.2.3. Extravillous trophoblast (EVT)

The extravillous trophoblasts (EVT) are derived from the cytotrophoblast cell columns (CCCs) located at the ends of the anchoring villi (AVs). The invasive EVTs are additionally divided into interstitial, endovascular and endoglandular EVTs based on their invasion in

either into the maternal decidua or the glands of the uterus. The interstitial trophoblasts (iEVT) drift through the maternal uterine stroma in the direction of spiral arteries and veins and invade more than half of the myometrium.[28] The endovascular trophoblast (eEVT) moves internally in the arteries and replaces the native endothelial and smooth muscle cells to establish the trophoblast plugs.[28,29,30] Trophoblast plugs then block the maternal blood entering in the spaces of the villi of placenta until the time point where proper circulation (haemochorial placentation) is established, which is usually towards the end of the first trimester.[31,32,33] This trophoblast mediated conversion of highly resistant vessel into a low resistant is an essential adaptation for proper fetal perfusion and a healthy pregnancy.[34] Endoglandular trophoblasts invade specifically into the maternal glands of the uterus and restore the epithelium of the glands to establish the histotrophic nutrition for the fetus under developement.[35,36] The process of invasion is at peak during the end of 12th week of pregnancy.[37] In pregnancies affected by PE, the maternal blood vessels are not sufficiently transformed due to the shallow invasion of the endovascular trophoblasts (eEVT) leading to abnormal placenta development and decreased blood flow between the mother and the fetus (Figure 1). Therefore, the significance of the role of different types of extra-villous trophoblasts (EVTs) in human placenta development during the early stages of pregnancy cannot be disregarded. Other than the different types of trophoblast cells, the placenta also contains cells that includes fibroblasts, immune (Hofbauer) and vascular cells (Figure 2).

1.3. Incidence of PE

The incidence of PE varies between different ethnic backgrounds and geographical locations.[38,39] In under developed countries due to the inadequate facilities for maternal and neonatal intensive care, maternal death rate is as high as 15% and the infant mortality rate is three times higher as compared to 0% to 1.8% in developed countries.[40] PE is one of the most common cause of preterm birth (PTB) (15-20%) and fetal growth restriction (FGR) (12-25%).[41,42]

1.4. Classification of PE

Depending upon the gestational age at diagnosis, the disease can be categorized as early-onset (< 34 weeks) and late-onset PE (≥ 34 weeks). Early-onset PE is relatively uncommon (12% of all PE) contrary to late-onset PE which is 88% of all PE. However, early-onset PE has a greater chance of maternal and fetal complexity as compared to late-onset PE. Late-onset PE often leads to a syndrome (HELLP) characterized by haemolysis, elevated liver transaminases and low platelets, and can also lead to eclampsia and maternal death.[43,44]

1.5. Risk factors of PE

Multiple risk factors have been linked with PE. Among them, women having a family history of PE are at a higher risk of developing PE. Besides this, there are some other known risk factors that can contribute to PE. For example women having age <20 or >35 years, having pre-gestational diabetes, pre-existing chronic hypertension/ renal disease, previous history of intrauterine growth restriction (IUGR), placental abruption or stillbirth, pregnancy with more than one baby, increased BMI before pregnancy, antiphospholipid antibody syndrome, systemic lupus erythematosus and assisted reproduction technology (ART).[45,46]

1.6. Pathophysiology of PE

PE has a complex pathophysiology and is a consequence of disproportion of factors produced by the placenta and the maternal adaptation to them. Maternal-fetal genetic incompatibility, exposure to paternal antigens and immunological factors lead to improper invasion of the trophoblasts resulting in the failure of remodeling of arteries in the myometrium, hence causing PE (Figure 3).[13] However, still the etiology of PE is not known clearly and it has been observed that placental ischemia/ hypoxia is a key factor in the development of this disorder. The ischemic placenta under stress, releases inflammatory cytokines, anti-AT_1 autoantibodies, trophoblast-derived micro particles, angiogenic and antiangiogenic factors into the maternal blood, thus generating intravascular inflammatory reaction and resulting in the clinical presentation of the disorder (Figure 3).[47,48]

Remodeling of the spiral arteries of the uterus supports fetal growth by achieving increase in the uterine blood flow during the development of placenta. The trophoblast cells are the key players in this process; they invade the arteries of not only the decidua region, but also of the myometrial segment of the spiral arteries. They increase the vasculature of the arteries by destroying the arterial media and endothelium. This results in lower resistance in the blood vessels and allows proper flow of the blood to the fetus resulting in a healthy pregnancy.[49] In case of PE (and eclampsia), trophoblast cells fail to migrate and reform the myometrial segment of the spiral arteries resulting in defects in placentation and ultimately leading to utero-placental ischaemia.[50] Since trophoblast cells are responsible for this key developmental process, defects in trophoblast could be one of the reasons for the improper placenta development as seen in case of PE.

In a healthy pregnancy at the site of trophoblast invasion, the maternal immune response against the paternal antigens of the fetus is not generated, possibly, because of the activity of the regulatory T (Treg) cells and the natural killer (uNK) cells in the decidua.[51] Uterine

natural killer (uNK) cells play a very important role in mediating the process of invasion of the trophoblasts during pregnancy. The invading extravillous trophoblasts (EVTs) expresses classical class I molecule HLA-C and non-classical HLA-E and HLA-G, that interacts with the receptors present on the uterine natural killer (uNK) cells such as CD94/ NKG2, killer-cell immunoglobulin-like receptors (KIR) and members of the immunoglobulin-like transcript (ILT) family, to mediate the process of trophoblast invasion at the site of implantation.[51,52] Angiogenic factors as well as chemokines are secreted by the uNK cells to facilitate the trophoblast invasion. In PE, the interaction between the invading trophoblast cells and the uterine NK cells is disturbed leading to intravascular inflammatory response.[51,52]

Elevated levels of pro-inflammatory/ Th1 cytokines, for example IL-1, IL-2 and interferon- -γ and decreased levels of Th2 cytokines IL-10 and IL-5 have been demonstrated in PE pathogenesis.[53] Additionally, activation of NF-kappa B, endoplasmic reticulum (ER) and oxidative stress in the placenta tissue generates inflammatory response by inducing the production of inflammatory cytokines (TNF-alpha, IL-6 and IL-17) resulting in the endothelial cell damage, therefore contributing to PE pathology.[54] Defective immune responses and genetic disposition increases the sensitivity against the effects of angiotensin II.[55] PE patients have been reported with higher levels of autoantibodies against type-1 angiotensin II receptor, which are known to activate the angiotension II receptor (AT_1) in the endothelial cells.[56] The anti-AT_1 autoantibodies when administered in pregnant rats, induced hypertension and proteinuria which is the clinical presentation of PE.[57]

Human placental development is very robust. After 21 days of fertilization, a protective shell of trophoblast cells encloses the fetus and establishes an interface with the maternal side. This phase of fetal development is supported by the secretions from the uterine glands (histotrophic nutrition).[58] During the early phases of implantation, the proliferation of the trophoblast is facilitated by the low oxygen tension in the gestational sac. This condition is conducive for trophoblasts to attach the blastocyst to the maternal tissues.[59] Lacunae are created within the trophoblasts which fuse to establish the intervillous space. A shift from histotrophic nutrition to the establishment of haemochorial placentation occurs when spiral arteries open at the spaces between the villi. Initial phase of placenta development happens under a state of relative hypoxia.[60] Oxygen tension is ultimately increased due to the first rush of blood into the intervillous space which generates oxidative stress and supports differentiation of trophoblasts from a proliferative to an invasive phenotype. This shift of trophoblast cells facilitates the deep influx into the decidua resulting in the physiological

transformation of the spiral arteries.[60] Under hypoxic conditions, the natural process of protein folding is interrupted by endoplasmic reticulum, indicated as the unfolded protein response (UPR).[61] The UPR affects cell proliferation and can undergo apoptosis of the trophoblast cells. The apoptosis of the trophoblasts produces micro and nanoparticles that are discharged into the maternal circulation, ultimately activating the immune response.[62,63] Many studies have reported the connection of ER distress and the stimulation of the UPR in PE and IUGR.[61,64]

Besides ER stress, oxidative stress is also linked with pathogenesis of PE as mentioned before. It originates when the generation of reactive oxygen species (ROS) overcomes the underlying antioxidant defense systems.[65] Excessive release of pro-inflammatory cytokines and chemokines have been detected due to the higher oxidative stress.[66] In the placentas of PE patients, due to the defected spiral arteries remodeling, the oxidative stress is induced because of sporadic hypoxia and reoxygenation.[65] The strength of uterine ischemia is affected by the severity of improper placenta development and requirement of the fetus for blood supply. Higher ROS affects protein carboxylation, lipid peroxidation and DNA oxidation as noticed before in placentas of PE patients.[65] Many studies have indicated that the antioxidant processes in placenta tissue are compromised in PE patients as compared with healthy controls.[67]

For a healthy placentation, the process of angiogenesis is very crucial.[68] Improper blood vessel development is well known to contribute to the establishment of PE. In PE patients, antiangiogenic factors have been detected to have higher expression in contrast to the healthy controls, which include soluble vascular endothelial growth factor receptor 1 (VEGFR-1) and soluble endoglin (Figure 3).[69] Soluble VEGF receptor binds to the circulating angiogenic factors (VEGF and PlGF) in the maternal blood and hinders their biological activity, resulting in imperfect angiogeneisis.[70] Soluble endoglin affects the migration and proliferation of endothelial cells by acting as a co-receptors of transforming growth factor β1 and β3 (TGF-β1 and TGF-β3).[71] Defects in the human *ENG* is known to cause disease of vascular impairment, characterized as haemorrhagic telangiectasia.[72] In-vitro experiments confirmed that the effect of s-endoglin on endothelial tube formation is similar to the sVEGFR-1.[71] The abnormal angiogenic and antiangiogenic balance is not limited to PE but has been observed in other obstetric disorders like IUGR, PTB, spontaneous abortion and fetal death (Figure 3).[73,74,75,76,77] However the clinical outcome (fetal death, PE with IUGR, IUGR alone and late-PE) is depended on the maternal response and the severity of the antiangiogenic state.

Decreased utero-placental blood flow, hypoxia, oxidative stress, inflammatory cytokines, release of trophoblast derived particles, anti-AT_1 autoantibodies, intravascular inflammation, leukocyte activation and endothelial cell dysfunction results in the imbalance, therefore affecting multiple organs resulting in the clinical presentation of PE.

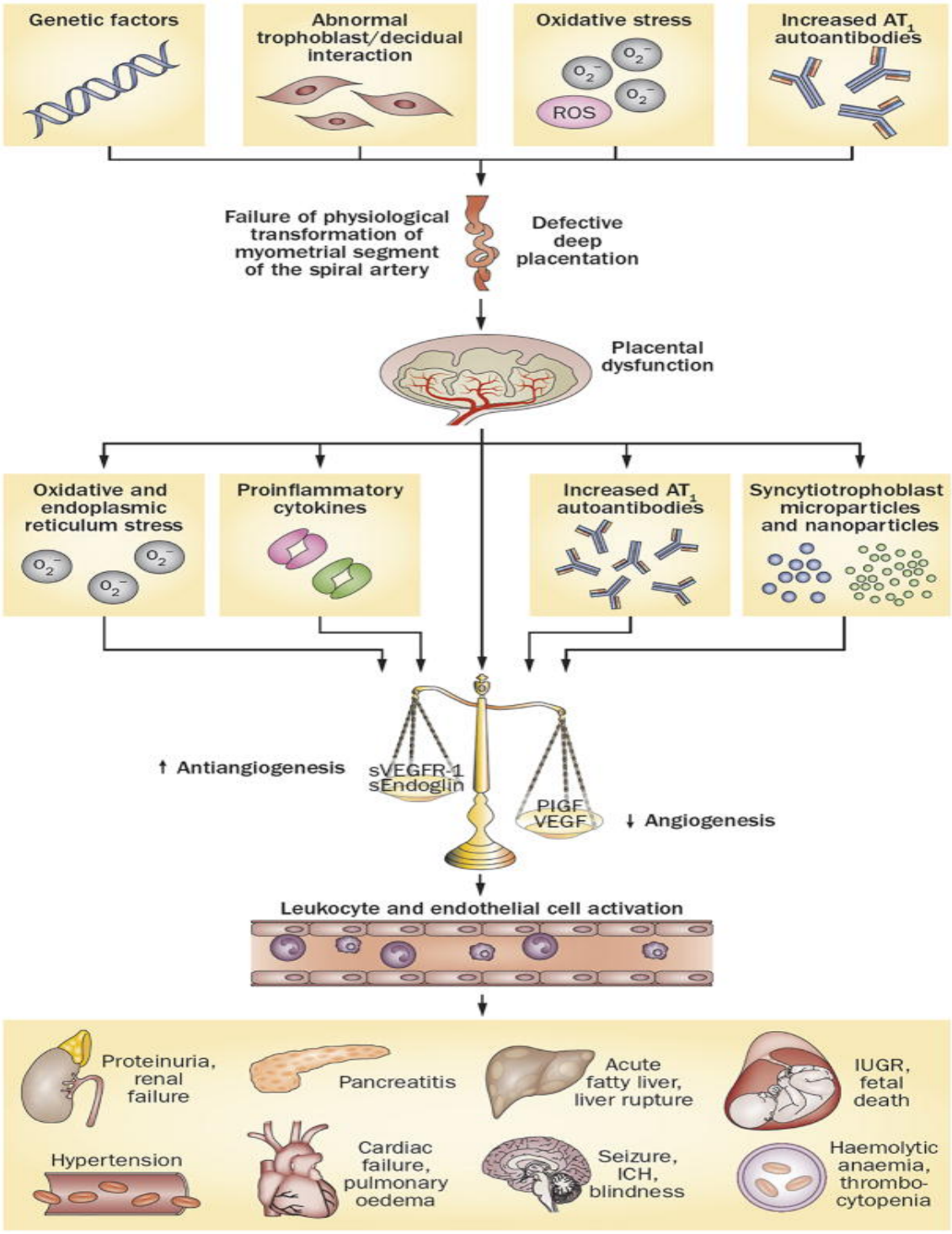

Figure 3. Pathophysiology and clinical features of PE

Risk factors related to pathogenesis and their contribution in the multiorgan features of PE are indicated. (AT_1: anti-angiotensin II type I receptor antibodies, ER: endoplasmic reticulum, ICH: intracerebral haemorrhage, IUGR: intrauterine growth restriction, PlGF: placental growth factor, ROS: reactive oxygen species, VEGF: vascular endothelial growth factor and sVEGFR-1, soluble vascular endothelial growth factor receptor 1). Adapted from [13]

1.7. Epigenetics and PE

DNA methylation and histone modifications are two well studied epigenetic modifications that are known to regulate the expression of the genes without altering DNA sequences (Figure 4).[78] Epigenetics has been recognized as an important mechanism, it plays a vital role in human placenta development and alterations in epigenetic mechanisms have been implicated in PE pathogenesis.[79] DNA methylation regulates various other biological mechanisms, together with lineage specification, inactivation of chromosome X, imprinted genes, stability of genome and control of the retrotransposon activation.[78] In human placenta, the transcription of various genes is influenced by changes in the DNA methylation due to external factors including exposure to toxic chemicals (smoking), nutrition, psychological state and assisted reproductive treatments.[80,81]

Mammalian placentas have unique epigenetic profile due to the less DNA methylation as compared to somatic tissues. Human placenta has 14–25% lower amount of global CpG DNA methylation.[82-85] Epigenetic modifications related to placenta has been speculated to support unique functions of placenta [5,6] and alterations in DNA methylation process can result in abnormal placental morphology and outcomes of pregnancy.[86] Decreased expression of Syncytin-1 due to increased DNA methylation has been observed in PE patients.[9] Syncytin-1 has a vital role in human placental evolution, growth and for a successful pregnancy. [7,8]

Non-imprinted genes were found to be affected by the epigenetic alterations, such as *SERPINB.* Higher expression of *SERPINB* was detected in PE placentas due to low promoter CpG methylation.[87] Similarly, impaired DNA methylation in the promoter regions of Vascular Endothelial Growth Factor A (*VEGF)*, Vascular Endothelial Growth Factor Receptor 1 and 2 *(FLT-1* and *KDR)* in PE patients were also identified.[88,89,90] However, it is not understandable yet that whether the observed differences are the effect or the reason behind the development of PE.

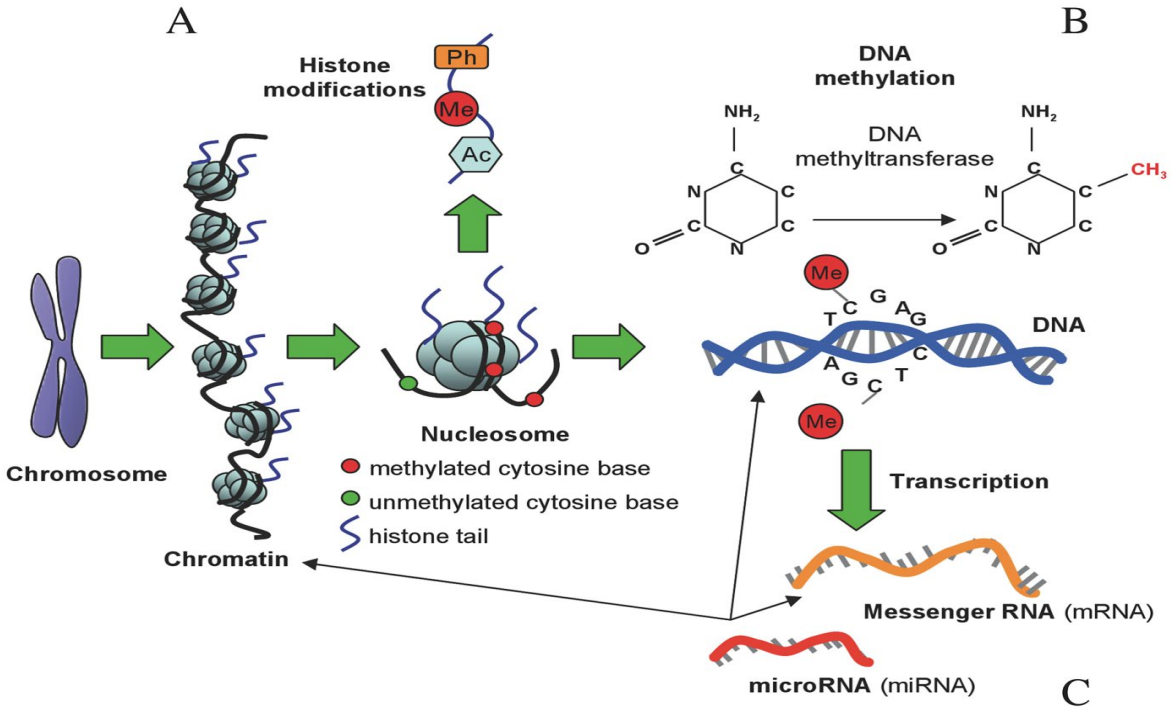

Figure 4. Epigenetic mechanisms of gene regulation

A) Histone modifications such as acetylation (Ac), phosphorylation (Ph) and methylation (Me) are shown. B) DNA methylation is carried out by DNA methyltransferase enzymes (DNMTs) that are involved in adding the methyl group at cytosine base of CpG sites, therefore, repressing the gene expression as well as the activity of TEs in the genome. C) Non-coding RNA (ncRNA) such as microRNAs (miRNA) regulates the protein translation by blocking the transcribed mRNAs. Adapted from[78]

1.8. Epigenetics and activation of transposable elements (TEs)

In mammals including humans, TEs are globally methylated in the somatic tissues and therefore are not able to be transcriptionally active.[91,92,93] In placenta tissue, DNA methylation is very low in certain regions due to which TEs, especially, endogenous retroviral (ERV) derived sequences such as ERV-LTRs are exponentially expressed as compared to other human tissues.[82,83,84,85,94,5,6,10] In human placenta, LTRs derived from HERV-E, function as alternative promoter of many cellular genes and it was found to be less methylated as compared to the blood cells where they do not exhibit the promoter activity.[5,38]

Correspondingly, another type of transposable elements, LINE1 elements (non-LTR retrotransposons), showed lower (almost 43%) DNA methylation in placenta as compared to the blood cells.[95] Placenta-specific DNA methylation changes were identified to be linked with the regulation of the expression of Syncytin-1 throughout pregnancy.[6,96] It is exponentially expressed at the beginning of placentation, to mediate the cell-cell fusion

activity and the expression declines at term, when the cell fusion activity is no longer required. Expression of some of the essential placenta-specific genes (*INSL4*, *EDNRB, PTN, MID1* and *IL2RB)* derived from human endogenous retroviral sequences (ERV-LTRs) is also a consequence of low methylation in the human placenta.[10,6] The function of ERV-LTRs in driving the placenta-specific gene expression is explained later in detail.

1.9. Transposable elements (TEs)

Almost half or more of the human DNA is composed of TEs and during primate evolution TEs have been co-opted to carry out a predominant role in rewiring the human genome.[97] TEs were found in late 1940s by Barbara McClintock in maize and she introduced them as "controlling elements".[98] Originally, TEs were considered junk DNA for a very long time but after that many studies have reported that almost 20% of the conserved DNA regulatory regions such as promoters, enhancers and transcription factor binding sites are derived from TEs.[99,100,101,102,103] Normally, their transcriptional ability is repressed in the cells, but their co-option has been demonstrated to mediate events such as recombination, splicing, exonification, and diverse means of gene regulation.[104]

1.10. Tissue-specific regulation by TEs

TEs are known to contribute to the tissue and lineage specific effects besides their critical role in regulation of adjacent genes, splicing and exonification.[104] They have several characteristics that impart the tissue-specific effect.

1. They provide specific transcription factor binding sites due to which the expression is confined to a particular tissue.[105 ,99]
2. They have the ability to reproduce inside the genome and regulate multiple genes.
3. They expand the possibility of distinct lineage variability because their effect is limited to the genomes they have infected.

Distinct tissue and developmental phase specific consequences have been predominantly disclosed in cancers, where specific TEs are often detected to be expressed tremendously.[106,107,108,109,110,111,112] Orthologous genes are regulated in homologous tissues (Prolactin (*PRL* gene)) due to various autonomous co-options of distinct TEs across species.[113] The envelope gene of different endoviruses (Syncytin gene) has been recruited to perform the similar activity in the placenta tissue of different lineages.[114-116,117] TEs are known to rewire the transcriptional control of gene expression in a contextual manner.

1.11. Types of TEs

TEs are divided into two important classes

1. Retrotransposons: that mainly function through reverse transcription.
2. DNA transposons: code their own transposase that mediate their insertion and excision.

However, in the human genome only retrotransposons are active and have played an important role in the genomic evolution.[118,119]

1.11.1. Retrotransposons

Almost half (42%) of the human genome consists of retrotransposons, while DNA transposons are less than 3%.[120] Retrotransposons have the ability to initially transcribe into an RNA, and then convert back into similar DNA sequences by reverse transcription. They can amplify easily by a "copy-and-paste" system, due to which they exist all over in the eukaryotic genomes. Based on the presence or non-existence of long terminal repeats (LTRs), type of open reading frames (ORF), coding sequences and duplication of target sites, retrotransposons are further divided into two classes,

1. LTR-retrotransposons (LTR: long terminal repeats) and
2. Non-LTR retrotransposons (long interspersed elements (LINEs) and short interspersed elements (SINEs)

1.11.1.1. LTR retrotransposons (Long Terminal Repeat)

Long terminal repeats (LTR) retrotransposons consist of sequences that range in size between 100 bp to over 5 kb. They are similar to a retrovirus but they have been classified separately. LTR endogenous retroviruses (ERVs) evolved from ancient descendants of exogenous retrovirus that infected germ line cells and were conserved in the germline. Endogenous retroviruses (ERVs) possess three viral genes; group-specific antigen (gag), polymerase (pol) and envelope (env) surrounded by two long terminal repeats (LTRs) that contain promoter elements, enhancer and polyadenylation sites (Figure 5).[121] Transposed LTRs in the human genome account for 8% and the mouse genome contains about 10% of the ERV-derived sequences.[122] Occasionally, due to the genetic rearrangements, solo ERV-LTRs are generated. A solo ERV-LTR co-opted near a cellular gene, can augment the transcription from the native gene promoter in a tissue-specific way.

1.12. Role of LTR retrotransposons in human placenta development

Human placenta development is very unique and novel. Humans have extremely invasive placenta as compared to the placentas of mouse and rat.[123] The invasive phenotype can be attributed to the primitive infection of the retroviruses (ERV) resulting in the expression of genes that are precisely elevated in the human placenta and are involved in the process of invasion during pregnancy.[124] Functional evolution of the human placentation owing to the insertion of human endogenous retrovirus (HERV) has been a major event.[125,126,127]

Phylogenetic analysis indicated that during primate evolution endogenous retroviruses (ERVs) integrated into the vertebrate's germ cells around 25-45 million years ago after the divergence of New and Old World monkeys.[124] The higher expression of few HERV families was found specifically in the human placenta tissue, that includes human endogenous retrovirus K, W, F, R and H (HERV-K, HERV-W, HERV-F, HERV-R and HERV-H).[125,128,126,127,129,130,131] LTRs derived from these ERVs are also highly transcribed in the placenta tissue as compared to other body tissues and there are many studies that manifest the role of ERV-LTRs in the human placentation (Figure 5).[132,124,133,109,134,135,136,137]

Endogenous retroviruses (ERVs) have contributed to the diversification of the human placenta and might explain the differences in the species in terms of the gestation, the total number of fetuses in one pregnancy, the size of the fetus and the time of the gestation.[138] Certain genes are expressed in the human placenta that are derived from the genes of endogenous retroviruses (ERVs). Besides this, there are a few genes that have placenta-specific enhancers and promoters derived from ERV sequences. All of them are known to have essential roles in the human placentation. ERV-derived genes and ERV-derived placenta-specific promoters and enhancers that are known so far in the literature are discussed in detail below.

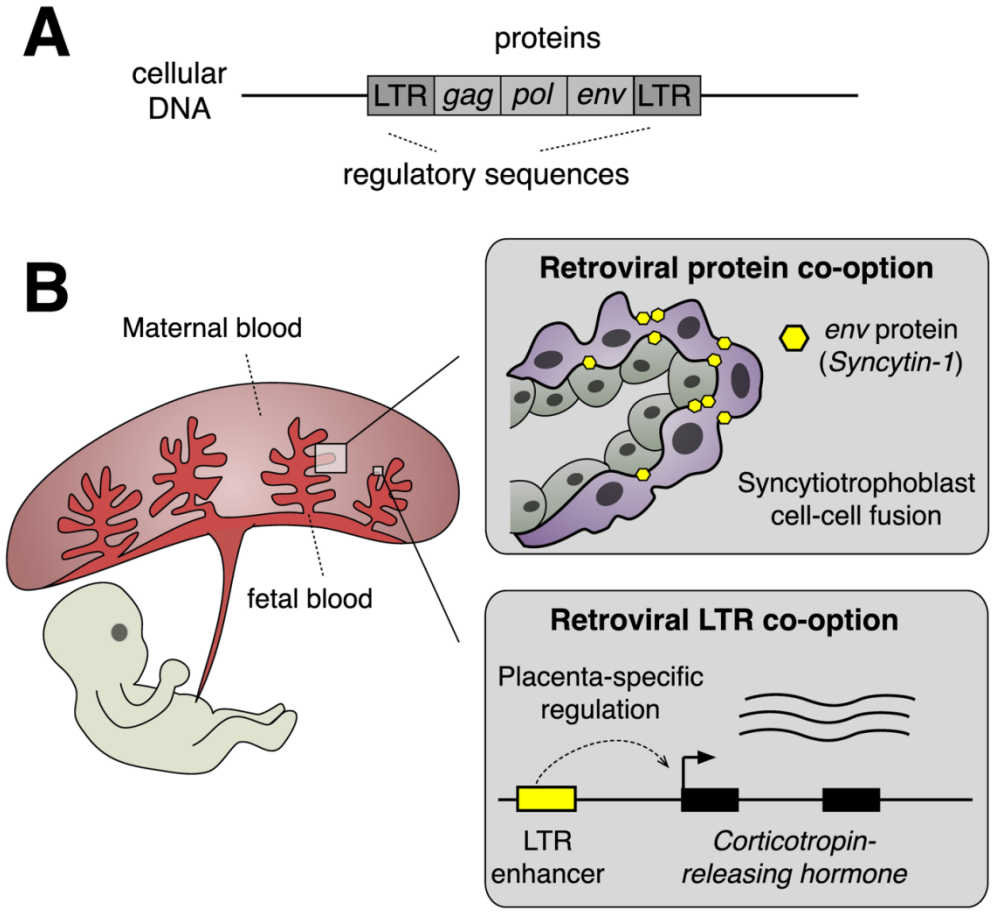

Figure 5. Role of endogenous retroviruses (ERVs) in human placenta
A) Genomic structure of an endogenous retrovirus (ERVs) containing retroviral genes flanked by LTRs. B) Examples of Syncytin 1 as a retroviral derived protein[125] and co-option of the solo ERV-LTR as an enhancer for *CRH* [139] in the human placenta. (LTR: long terminal repeats). Adapted from[140]

1.12.1. Placenta-specific genes derived from endogenous retroviruses (ERVs)

The number of genes confined to the placenta tissue that are not derived from ERVs are limited.[141] Trophoblast-restricted protein (Tpbp) a and b in mouse[142] and the similar spongiotrophoblast specific (SSP) protein in rats[143] are exclusively expressed in the placenta tissue. However, these genes have been found only in rodents. In humans, placenta-specific protein 1 *(PLAC-1)* was identified to have enriched expression in the placenta tissue. However, recently its expression has been found in human testis as well.[144] Orthologous genes of *PLAC1* have been discovered in cow, rat and mouse placenta tissue. Biological function of *PLAC1* has not been defined yet, but it is thought to be linked with trophoblast differentiation and fibroblast growth factor (FGF7) signaling in the placenta tissue.[145,146]

Throughout human evolution most of the endogenous retroviruses (ERVs) have gone through a lot of modifications that have restricted their transcription. However, numerous endogenous retroviruses (ERVs) have intact open reading frames (ORFs) of their genes and are expressed.[147,148] Among them, the best characterized genes are Syncytins and ERV-3 that have been acquired from the envelope gene of the endogenous retroviruses (ERVs). The envelope gene plays a critical role in the human placenta development, such as executing the process of cell fusion to form the syncytiotrophoblasts, regulating the immune response generated against the fetus, and restricting the chances of getting infected by exogenous viruses.[149,147] ERV-3, despite its expression in the syncytiotrophoblasts, is not only considered as a fusogenic protein. Relatively, ERV-3 has been demonstrated in the differentiation of trophoblasts and regulation of human chorionic gonadotropin (hCG) expression in the placenta tissue.[147]

The most expressed placenta-specific genes are human Syncytin genes, Syncytin-1 and Syncytin-2, derived from envelope genes of endogenous retroviruses HERV-W and HERV-FRD. Due to their fusogenic ability, they induce the establishment of the multinucleated syncytiotrophoblasts in the placenta tissue.[134,147,150] Syncytin-1 is expressed in the syncytiotrophoblasts throughout the pregnancy (Figure 5).[125,116,148,151] However, Syncytin-2 is different from Syncytin-1, although they share similarity in their sequences and the ability to fuse the cells. Syncytin-2 additionally contains a protein domain that suppresses the maternal immune system and protects the embryo.[116] Syncytin-2 is not localized to syncytiotrophoblasts, however its expression was also found in a subset of villous cytotrophoblast cells.[152,153] In contrast to Syncytin-1, the levels of Syncytin-2 declines significantly at term.[152] The differences indicate their diverse role in the human placenta tissue. Similar to the human Syncytins, mice have autonomously obtained retroviral envelope proteins from a distinct endogenous retrovirus (ERV) family; known as syncytin A and B.[115] They are exclusively expressed in the placenta tissue and function as a fusion protein.[115] In the mouse placenta, they are primarily localized in the labyrinth layer containing syncytiotrophoblasts.[115] The example of the mouse and the human syncytin genes indicate the function of ERV-derived genes in the development and evolution of the placenta and possibly explain the differences within the species such as between humans and mice placentation.

1.12.2. Placenta-specific promoters and enhancers co-opted from human endogenous retroviral LTRs (ERV-LTRs)

In humans, certain genes are broadly expressed in most of the tissues of the body. However, in placenta, their expression has been found to be elevated or sometimes a placenta-specific isoform has been detected, mainly due to the co-option of an ERV-LTR element near the regulatory region of the genes (Figure 5).[132,154] As mentioned before, the human ERV-LTRs are structurally enriched in the transcription factor (TF) binding sites, they utilize these LTR sequences for their own replication by using the host cellular factors.[102] Co-option of a single ERV-LTR as a promoter, enhancer or both has been reported to drive the tissue-restricted expression by providing a novel regulatory sequence.[155]

The human *CYP19* has been evolved to utilize several promoters in different tissues.[156] Placenta-specific promoter of human *CYP19* has been characterized as MER21A, which is derived from an ERV-LTR.[157] The exponential expression of *CYP19* in human placenta is as a result of the co-option of MER21A. It encodes for aromatase, a steroidogenic enzyme that converts androgens to estrogens and regulates the growth of the maternal uterus and the placenta and also plays a vital role in parturition.[158,159] Different species (bovine, ovine and humans) have independently acquired their placenta-specific promoters for *CYP19*.[160,161] Conservation of *CYP19* indicates its crucial role for pregnancy in all the mammals. Since, it is required throughout gestation so it seems that this requirement is fulfilled by the co-option of an ERV-LTR that was acquired during evolution.

Some other genes involved in the human placenta development were identified to have increased placenta-specific transcription mediated by a unique tissue-specific promoter/ enhancer obtained from an LTR of HERV-E retroviral family. For example, the promoter/ enhancers of Opitz syndrome associated gene Midline1 (*MID1*), endothelin receptor B gene *(EDNRB)* and human growth factor pleiotrophin gene *(PTN)*. The reason behind the placenta-specific transcription of *PTN* is mainly due to the generation of an exclusive binding site for Sp1 (a transcription factor for estrogen receptor) due to the insertion of an ERV-LTR. *PTN* is expressed in different sub-types of trophoblasts and is known to regulate trophoblast life cycle and placental angiogenesis.[162] Human ERV-LTR associated with *EDNRB* contains binding sites for Thing1/Hand1 that is required for the early trophoblast differentiation. Due to the available binding site on the human ERV-LTR it drives the placenta restricted expression of the *EDNRB*. Endothelin besides their role in vasodilation or vasoconstriction, mediate trophoblast proliferation and invasion through the Endothelin B receptor in humans.[163] The

sequences of the human ERV-E LTRs associated with *EDNRB*, *PTN* and *MID1* genes are highly similar, but the promoter/enhancer activity varies in terms of the tissue specificity and intensity of the gene expression. However, they provide unique transcription factor (TF) binding sites and their selective DNA methylation in placenta (as compared to other tissues) have contributed to the placenta-restricted expression of the genes. [5,164,132,165,166]

Another gene, Leptin is immensely expressed in the human placenta tissue and MER11 (ERV-LTR) has been identified as its placenta-specific enhancer.[131] Leptin prepares the uterus for parturition and establishes the lactating state of pregnancy. In human trophoblasts, it promotes cell proliferation and survival.[167] Orthologous genes of human Leptin, *EDNRB*, and Pleiotrophin have been identified in the placentas of rodents but their expression is not mediated by an ERV-LTR.[138] The sustenance of their expression in placenta is suggestive that these genes have preserved their function in the development of placenta tissue.[168,169,170]

Placenta-specific expression of the early placenta insulin-like peptide 4 (*INSL4)* gene, in humans, is carried out by a primate-specific ERV-LTR of a HERV-derived sequence. Since the co-opted human ERV-LTR is near the TSS of *INSL4*, it gives additional features to the native gene promoter. The primate-specific ERV-LTR has also detected in the promoter of *INSL4* homologue gene in the Rhesus Monkey.[124] *INSL4* has a robust expression during the initial phase of placentation and in differentiated syncytiotrophoblasts suggesting its role in placental morphogenesis.[171,172] Similarly, the trophoblast-restricted expression of the Interleukin-2 Receptor beta subunit (*IL2RB*) has been reported to be mediated by a placenta-specific promoter, derived from an ERV-LTR of the ancient domesticated THE1D retroviral family. *IL2RB* is required to establish a functional cytokine receptor for interleukin-2 (IL-2) or interleukin-15 (IL-15) signaling that is mediated at the maternal-fetal interface during placentation.[173,174,175]

Some of the above discussed placenta-specific genes have been found to be dysregulated in PE patients. For example, dysregulated expression of *INSL4*, *PTN*, Leptin, and *CYP19* have been reported before in PE patients.[176,177,178,179,180] Currently, in humans a limited number of genes are known to be dysregulated in PE and have a placenta-restricted gene regulation mediated by an ERV-derived LTRs. Many more studies are required to identify such genes. Additionally, this would also help to understand the human-specific aspect of normal and abnormal placenta development as seen in PE patients. This study was focused on identifying new target genes that are highly expressed in the human trophoblasts, regulated by ERV-LTRs and dysregulated in PE patients.

1.13. Objectives of the study

PE has only been observed in humans and not in non-human primates. The underlying mechanisms of human-specific factors related to PE development are not well studied. In this study, I focus on the identification of new PE associated genes that are regulated by human endogenous retroviral LTRs (ERV-LTRs) and are predominantly expressed in the trophoblast cells. This strategy has not only identified already known genes but also new candidate genes contributing to PE pathogenesis.

The aim of the thesis was further divided into the following hypotheses:

1. Human ERV-LTRs contribute/ regulate the trophoblast-specific gene expression.

It is known that the primate-specific ERV-LTRs are expressed in the human placenta due to low DNA methylation and their activation can lead to the possibility of regulation of genes located in their vicinity.[5] In the project, one of the focus areas was identification of the role of human ERV-LTRs in trophoblast-specific gene regulation and understanding of human-specific nature of PE development. RNA sequencing was performed on the primary human trophoblast cells of 5 healthy and 5 early-onset PE placentas. Genes that were enriched in the trophoblast cells as compared to other 38 human body tissues were identified. The expression analysis of TEs belonging to different families was done in human tissues. TEs enriched in placenta tissue were then selected for further analysis. RNA-seq analysis identified genes that have stronger expression in trophoblasts and showed co-regulation of expression with the human ERV-LTRs located in their vicinity (10-kb upstream of transcription start side (TSS) of the gene. The focus was on the human ERV-LTRs that were found to be transcriptionally active in the placenta tissue. The aim was to confirm the trophoblast-specific gene expression by q-PCR on pregnancy-related tissue panel samples and to investigate the expression of candidate genes in trophoblast sub-types by analysis of online available RNA-seq data of single-cells obtained from human placenta tissue. The potential activity of the human ERV-LTRs associated with the trophoblast-specific gene was predicted by ChIP-seq data available for active histone marks and transcription factors important for trophoblasts differentiation, followed by an *in-vitro* GFP reporter assay.

2. The second working hypothesis was that the genes that are exclusively expressed in trophoblast cells or have trophoblast-specific gene regulation due to the associated ERV-LTR might have a predominant role in the development of human placenta tissue and their dysregulation might lead to pregnancy-related disorders such as PE.

Here the aim was to validate the dysregulated expression of trophoblast-specific genes by q-PCR on human PE patient samples of Oslo-cohort-II, consisting of 28 healthy, 24 early-onset PE and 22 late-onset PE patients. To gain deeper insight into the functional analysis of the novel candidate genes in PE pathogenesis, *EPS8L1* was further investigated. Moreover, the trophoblast-specific localization of EPS8L1, its expression in human placental villi, EPS8L1-protein interactors in trophoblast cell lines and the role of *EPS8L1* overexpression *(in-vitro)* in SGHPL-4 cells (EVT-like) was investigated by functional assays.

Collectively, the strategy was successful in identification of additional genes, regulated by human ERV-LTRs in trophoblast cells. As an example, the function of a previously uncharacterised gene *EPS8L1* with an important role in human placenta development and how its dysregulation is related to the development of PE was demonstrated.

2. Materials and methods

2.1. Materials

2.1.1. mRNA isolation, quantitative polymerase chain reaction (q-PCR)

QIAzol Lysis Reagent	(QIAGEN)
RNeasy Mini Kit	(QIAGEN)
Direct-zol^{TM} RNA MiniPrep	(Zymo Research)
High Capacity cDNA Reverse Transcription Kit	(Applied Biosystems)
TaqMan Fast Universal	(Applied Biosystems)
Fast SYBR Green Master Mix	(Applied Biosystems)
Aglient RNA 6000 Nano Kit	
TruSeq Stranded mRNA LT Set A kit	(Illumina)

2.1.2. Primers and sgRNAs

The primers and oligonucleotides for quantitative polymerase chain reaction (q-PCR) were designed using Primer Express ®ABI PRISM software. Some of the primer sequences were taken from public resource for q-PCR primers, primer bank MGH-PGA. Primers for cloning were designed either using the NEBuilder Assembly Tool or manually and were checked according to the general guidelines for primer design and best practices. The sgRNAs for *EPS8L1* knock-out (KO) target sequence were designed using CRISPOR guide design tool.

Table 1. Primers and sgRNA

	Gene	Primer	Sequence (5' > 3')
q-PCR	*ALDH3B2*	Forward	ATGAAGGATGAACCACGGTCC
		Reverse	GTTCCAGGGTGCGATGATGA
	CSF2RB	Forward	CTCGTCAACGTGACCCTCAT
		Reverse	CGACAAAACTCTGGCAGGGA
	CYP11A1	Forward	CTTCACCCCATCTCCGTGAC
		Reverse	GTCTTTGCTCAGCCATCGG
	PHYHIPL	Forward	GTGCAGACTGCCTCAAAACA
		Reverse	AAAACTTAAGCATGCGTCCTG
	SPINT1	Forward	AAGGTACAACCCCAGGAACC
		Reverse	CCTCTGGGTGGTCTGAGCTA
	DACT2	Forward	CGGTCGGTTGATGAGACTACT
		Reverse	CAGGGCTCTGTCAAGATCACC
	SLC22A11	Forward	TATTAAGGGCAAACCAGACCAAG
		Reverse	CCAGCCCATAGTAGGAGATCAA
	EPS8L1	Forward	GAGAGCTTTGTATCGAGGCTG
		Reverse	GCCCGAAAAGGAAGTGCAAC
	PLEKHA8	Forward	AGCCTCGATGGTTCCTTCTCT
		Reverse	TCAGGTCCATGCGTGTATTATCT
	RAD51	Forward	CAACCCATTTCACGGTTAGAGC
		Reverse	TTCTTTGGCGCATAGGCAACA
	KIF23	Forward	TGGTTCCTACATTCAGAAATGAGA
		Reverse	CGTTCTGATCAGGTTGAAAGAGTA
	NRK	Forward	CATTGGCCTTGGTACTTATGGC
		Reverse	GTCTTACGAGCGTTCATCACTT
	C1QTNF6	Forward	GAAAGGGTCTTTGTGAACCTTGA
		Reverse	CTGCGCGTACAGGATGACAG

	EPS8	Forward	TGAATGGCTACGGATCATCACC
		Reverse	CACTGTCCCGTGCATAATTCT
	MYC	Forward	GTCAAGAGGCGAACACACAAC
		Reverse	TTGGACGGACAGGATGTATGC
	CCND1	Forward	GCTGCGAAGTGGAAACCATC
		Reverse	CCTCCTTCTGCACACATTTGAA
	MMP9	Forward	AGACCTGGGCAGATTCCAAAC
		Reverse	CGGCAAGTCTTCCGAGTAGT
Housekeeping gene	*18S*	Forward	ACATCCAAGGAAGGCAGCAG
		Reverse	TTTTCGTCACTACCTCCCCG
	GAPDH	Forward	AGCCACATCGCTCAGACAC
		Reverse	GCCCAATACGACCAAATCC
HA-EPS8L1 construct	hEPS8L1-EcoRV	Forward	acgtcccagactacgctgatatcATGAGCACCGCCACAGGC
	hEPS8L1-NotI	Reverse	aataaacaagttaacaacgcggccgcTCAAATGACCTCCATTTCCACCTCG
non HA-EPS8L1 construct	hEPS8L1-SgrAI	Forward	CCCGGAGCGCCGGCGGCTGTC
	hEPS8L1-EcoRV	Reverse	CGGGATATCGGTGAATTCTTT
LTR-construct (GIBSON assembly)	CYP11A1-MLT1F2-1	Forward	catggctttagaagcttgatGGGTGGCTGGGTGGCTGA
		Reverse	cactagtgaattccatggatAGTCCTGAAACAGCCACTGAAAAGTTTTG
	CYP11A1-MLT1F2-2	Forward	catggctttagaagcttgatTGTATTAGCTGTTTACTGCTG
		Reverse	cactagtgaattccatggatAGTCACCCAGAAAA

			GCTG
	CYP11A1-MLT1J	Forward	catggctttagaagcttgatATGTGGTTAAGTTCCCAC
		Reverse	cactagtgaattccatggatACAGAAAATCCAAATCACAG
	ALDH3B2-MLT1F2	Forward	catggctttagaagcttgatAATGCTGCAGCCAAGCTG
		Reverse	cactagtgaattccatggatACTAATACAATGGCAAAAAGTAACAGG
	SPINT1-MLT1F2	Forward	catggctttagaagcttgatATTCCAAAAGTTACCAGCTTG
		Reverse	cactagtgaattccatggatTGTGGTAGGTTGCTTCTAAATG
	CSF2RB-MLT1C	Forward	catggctttagaagcttgatTGTTATGAGGTTGGCTTG
		Reverse	cactagtgaattccatggatTGTGTCAATTTCCTAGGG
	DACT2-MLT2B4	Forward	catggctttagaagcttgatTGTGATGATTAATTTTATGTGTCAAC
		Reverse	cactagtgaattccatggatAAGAATCCCTCTTGGTGAAG
	DACT2-MER41D	Forward	catggctttagaagcttgatTTAGGGAATCAGGAGCCC
		Reverse	cactagtgaattccatggatTGTTACCAGAAAAGCAGTC
	EPS8L1-MLT1G	Forward	catggctttagaagcttgatAAGGAGCTCCAGTGGTCC
		Reverse	cactagtgaattccatggatTTCCAGTTACTATGGCCATATAACAAATTATTC
Sequencing primers	hEPS8L1-F1	Forward	ACAATTACCGCTCGGGCCGC
	hEPS8L1-F2	Forward	GAGGCCGAGTACACCGACGT

	hEPS8L1-F3	Forward	ATCTGAGCCTCAGCTGGAGT
	HA Seq-F3	Forward	ATGGTAATCGTGCGAGAG
gRNA EPS8L1 KO	EPS8L1gRNA1-F	Forward	CACCgACAATTACCGCTCGGGCCGC
	EPS8L1gRNA1-R	Reverse	AAACGCGGCCCGAGCGGTAATTGTc
	EPS8L1gRNA2-F	Forward	CACCgCACAATTACCGCTCGGGCCG
	EPS8L1gRNA2-R	Reverse	AAACCGGCCCGAGCGGTAATTGTGc
	EPS8L1gRNA3-F	Forward	CACCgCAATTACCGCTCGGGCCGCG
	EPS8L1gRNA3-R	Reverse	AAACCGCGGCCCGAGCGGTAATTGc

Table 2. Antibodies used in the study

Technique	**Antigen**	**Host**	**Conjugate**	**Dilution**	**Manufacturer**
Western blot	hEPS8L1	Rabbit	-	1:250	Atlas Antibodies
	Actin	Mouse	-	1:5000	Dianova
	Rabbit IgG	Goat	HRP	1:5000	Thermo Scientific
	Mouse IgG	Goat	HRP	1:5000	Thermo Scientific
Immunohistochemistry	hEPS8L1	Rabbit	-	1:300	Atlas Antibodies
	HRP Polymer	-	-	-	Lab Vision
Immunostainings	hEPS8L1	Rabbit	-	1:100	Atlas

					Antibodies
	Actin	Mouse	-	1:1000	
	Anti-Rabbit	Goat	AlexaFluor 488	1:200	Life Technologies
	Anti-Mouse	Donkey	Alexafluor -647	1:200	Life Technologies

2.1.3. Western blot

- RIPA buffer:
  - 0.05 M Tris-HCl pH 7.4
  - 0.150 M NaCl
  - 0.001 M EDTA
  - 1% Triton-X100
  - 1% Na-Deoxycholate
  - 0.1% SDS
  - Complete Protease Inhibitor Cocktail Tablets (Roche Diagnostics)
- 6X SDS loading buffer:
  - 3 ml Glycerol
  - 1.5 ml B-Mercaptoethanol
  - 9 ml 10% SDS
  - 1 M Tris-HCl (pH 6.8)
  - Bromophenol Blue (pinch)
  - Adjust to 10 ml water
- TEMED
- APS
- Triton X-100 (Sigma)
- Skim milk powder (Fluka)
- Page Ruler™ Plus Prestained protein ladder (Thermo Scientific)
- TGX stain free FAST cast acrylamide gel kit (Bio-Rad)
- Trans-Blot Turbo midi-size transfer stacks (Bio-Rad)
- Trans-Blot Turbo midi-size PVDF membrane (Bio-Rad)
- 1X Transfer buffer:

- 5X TransBlot Turbo transfer buffer 200 ml
- 600 ml of Mili-Q H_20
- 200 ml of ethanol

- 10X TBS:
  - 0.5 M Tris
  - 1.5 M NaCl
  - 1 L Mili-Q H20
  - pH 7.5
- 1X TBS-T:
  - 1/10 TBS
  - 0.1% Tween 20
- Mild stripping buffer:
  - 0.2 M Glycine
  - 0.1% SDS
  - 1% Tween-20
  - pH 2.2

2.1.4. Flow cytometry

- Fix/ Perm buffer (BD Biosciences)
- Perm/ Wash buffer 10X (BD Biosciences)
- 1 % PBS/ FCS (v/v)

2.1.5. Immunostaining and immunohistochemistry

- 0.1 M Citrate buffer (pH 7.0)
- Ultra V block (Lab Vision)
- 0.05 % PBS-T (v/v)
- DAPI (Sigma)
- ProLong Gold Antifade Mounting Medium (Life Technologies)
- Aminoethylcarbazole 45 Chromogen (Lab Vision)
- Kaiser's Glycerol Gelatine (Merck)

2.1.6. Isolation of primary human trophoblasts

- HBSS/ HEPES (25 mM HEPES):
  - 100 ml 10X HBSS (Hanks Balanced Salt w/o Phenol red, with Ca, Mg) (Gibco)
  - 5.958 g HEPES (Gibco)
  - 900 ml H_2O dest.
  - Set pH on 7.4 with NaOH
- DMEM/ HEPES (25mM HEPES):
  - 500 ml DMEM (Gibco)
  - 2.9788 g HEPES
  - Set up pH on 7.4 with NaOH and sterile filtrate
- Buffer 1a:
  - 285 mg Trypsin (Sigma)
  - 20 mg DNase I (Roche)
  - 225 ml warm HBSS/ HEPES
- Buffer 1b:
  - 10 mg DNase I
  - 25 ml warm HBSS/ HEPES
- Buffer 2:
  - 190 mg Trypsin
  - 10 mg DNase I
  - 150 ml warm HBSS/ HEPES
- Buffer 3:
  - 150 mg Trypsin
  - 10 mg DNase I
  - 110 ml warm HBSS/ HEPES
- Buffer 4:
  - 10 mg DNase I

- 100 ml DMEM/HEPES
- Dilution of New Born Calf Serum (Biochrome) (45 ml NBCS and 7.5 ml H_2O dest.)
- HBSS/HEPES for Percoll (10X HBSS w/o Ca, Mg and 250 mM HEPES)
- 90 % Percoll (5 ml HBSS/ HEPES for Percoll and 45 ml Percoll) (GE Healthcare)

Table 3. Percoll gradient for trophoblast isolation

	g/ml	90% Percoll	HBSS/HEPES
70%	1,084	3,107 ml	0,893 ml
60%	1,071	2,667 ml	1,333 ml
50%	1,0595	2,227 ml	1,773 ml
40%	1,0485	1,773 ml	2,227 ml
30%	1,0375	1,333 ml	2,667 ml
20%	1,0270	0,893 ml	3,107 ml

- Monoclonal Mouse Anti-Human HLA-ABC Antigen (Dako Cytomation)
- Dynabeads Pan Mouse IgG (Dynal Biotech)
- SmBM Basal medium + SmGM-2 Single Quot Kit Suppl & Growth Factors (Lonza)

2.1.7. Cell transfection

jetPRIME DNA Transfection Reagent (PolyPlus)

Lipofectamine™ 3000 Transfection Reagent (Invitrogen)

NEON™ Transfection System (Life Technologies)

2.1.8. Mass spectrometry

- Lysis buffer:
  - 50 mM TRIS-HCL (pH:8.0)
  - 100 mM NaCL
  - 10 uM EDTA
  - 5% Glycerol
  - 1% NP-40
  - Benzonase and Protease inhibitor were added fresh
- Washing buffer (10X lysis buffer):
  - 5 mM TRIS-HCL pH 8.0
  - 10 mM NaCL

- 1uM EDTA
- 0.5% Glycerol
- 0.5% NP-40

- Pre-elution buffer:
- 200 mM KCL
- 10 mM TRIS-HCL pH 8.0

- Ezview Red Anti-HA Affinity Gel (Sigma Aldrich)

2.1.9. Reagents and kits for assays

- DCFH (2`7`-Dichlorofluorescin) (Sigma)
- Human EPS8-like 1 ELISA Kit (MBS9317820) (MyBioSource)
- Recombinant Human Epidermal Growth Factor (hEGF) (Peprotech)

2.1.10. Chemicals

Chemicals were ordered from the following companies: Karl Roth GmbG & Co. KG, Jena Bioscience, Merck KGaA, Sigma-Aldrich GmbH, Amersham-Pharmacia, Invitrogen, Qiagen, GE Healthcare, BD Bioscience and Thermos Fischer Scientific.

2.1.11. Software

- 7500 Fast System SDS Software (Applied Biosystems)
- Prism Version 7 (GraphPad)
- ImageJ/ Fiji (GNU General Public License)
- Primer Express 3.0 (Applied Biosystems)
- Flow Jo (TreeStar Inc.)
- BD FACS Diva (BD Biosciences)
- Case Viewer (3DHISTECH)
- NEBuilder Assembly tool v2.2.6 (New England BioLabs)
- CRISPOR (Tefor infrastructure)
- Image Analysis Software (Wimasis)

2.1.12. Hardware

- NanoDrop Spectrophotometer ND-1000 (PeqLab)
- 7500 Fast Real-Time PCR System (Applied Biosystems)
- FACS Caliber system (BD Biosciences)
- LSM710 point-scanning single photon confocal microscope (Leica)

- FACS Aria II (BD Biosciences)
- FACS Aria III (BD Biosciences)
- ChemiDoc MP V3 Western workflow (Bio-Rad)
- Neon electroporation transfection system (Life Technologies)
- Microplate reader GENios plus (Tecan)
- EVOS Imaging System (Thermos Fischer Scientific)
- Aligent 2100 Bio Analyzer machine (Agilent)

2.2. Methods

2.2.1. Patient cohorts

The following cohorts were used in this study.

1. **Oslo cohort:** Microarray data analysis of human placental samples has been published before comprising of PE patient samples from the bio-bank collection at Oslo University Hospital, Norway, authorized by the Regional Committee of Medical Research Ethics in Eastern Norway.[181,182] Results were verified in a second cohort that consists of placental tissues collected following C-section from PE patients (n=50) and controls (n=28) with normal blood pressure and uncomplicated pregnancies. The PE group was further divided into early-onset PE (< 34 gestational weeks, n = 26) and late-onset PE (> 34 gestational weeks, n = 24). The control group included healthy, normotensive women undergoing cesarean section due to breech presentation or other reasons. For validation of the differentially expressed genes in PE (Table 9), I used samples of this cohort consisting of control (n = 28), early-onset PE (n = 24), late-onset PE (n = 22) and IUGR+ PE (n= 6). All the quantitative PCRs were performed on the same set of samples, however some samples of this cohort were used up already and were not enough to be tested. Patient characteristics are shown in Table 4. The expression level of EPS8L1 in the placenta tissue isolated from the patients was correlated with the sFlt1/ PIGF ratio which is the prognostic marker for PE, and also with the gestation age (GA) of the PE patients and controls. The serum levels of PIGF and sFLT1 of the patients were detected from the second PE cohort on Elecsys (Roche Diagnostics) at HELIOS Klinikum GmbH.

Table 4. Clinical characteristics Oslo cohort

Clinical characteristics of PE patients and healthy controls. Data are presented as mean ± standard deviation, BMI: body mass index.

Characteristics at delivery	Normotensive Controls (n = 28)	Preeclampsia (n = 50)	P value (Mann-Whithey U)
Maternal age (years)	31.2 ± 4.2	31.6 ± 5.6	0.63
BMI (kg/m^2)	28.6 ± 3.4	31.5 ± 5.2	0.01
Gestational weeks at delivery	39.0 ± 0 .9	33.7 ± 3.7	< 0.001
Blood pressure Systolic (mm Hg)	119 ± 11.6	165 ± 16.2	< 0.001
Blood pressure Diastolic (mm Hg)	72.1 ± 10.9	101.1 ± 6.7	< 0.001
Baby weight (g)	3492 ± 402	2170 ± 1005	< 0.001

2. **Charite and Kiel cohort**: The samples from Charite and Kiel cohort were used additionally to validate the dysregulation of EPS8L1 in PE patients by using q-PCR (Figure 12). The samples from two cohorts were merged and the clinical characteristics are shown in Table 5. All PE samples are gestational age matched with their controls.

Table 5. Clinical characteristics Charite and Kiel cohort
Clinical characteristics of the PE, IUGR and control cases from the study population. Early is defined as delivery < 34 gestational weeks. Late is defined as delivery ≥ 34 gestational week. Data are presented as mean ± standard deviation, BMI: body mass index. *P<0.05; **P<0.001 vs. control.

Characteristics at delivery	Control early (n=36)	Control late (n=65)	IUGR early (n=28)	IUGR late (n=13)	PE early (n=14)	PE late (n=23)
Maternal age (years)	31.4 ± 5.7	32.5 ± 5.6	27.7 ± 6.4	27.6 ± 5.8*	33.0 ± 5.3	31.2 ± 6.2
BMI (kg/m^2)	25.0 ± 6.1	23.6 ± 4.6	24.0 ± 4.8	26.2 ± 7.0	26.0 ± 5.1	27.5 ± 7.0*
Gestational days	203.3 ± 22.1	264.7 ± 13.0	208.6 ± 19.0	264.3 ± 11.8	206.1 ± 21.3	254.5 ± 12.0
Blood pressure Systolic (mm Hg)	115 ± 9.7	116.6 ± 10.8	122.9 ± 13.7**	104.5 ± 15.5*	152.6 ± 12.4**	153.5 ± 14.5**
Blood pressure Diastolic (mm Hg)	65.0 ± 8.0	69.7 ± 8.1	70.7 ± 11.7**	65.2 ± 10.3	95.9 ± 8.7**	94.8 ± 8.8**

3. **Manchester cohort (High risk pregnancy cohort):** ELISA was performed on a total of 24 serum samples (Controls n=12, early-onset PE n=12) from this cohort. All controls had a history of hypertension and some of the control cases had diabetes as well. Early-onset PE samples included those who developed PE during pregnancy. Samples included for this study were from 24th - 28th weeks of gestation. The age of the patients included in this analysis was more than 18 years, belong to different ethnic backgrounds and had no smoking history.

4. **Berlin Menstrual Study cohort:** Serum sample of 6 non-pregnant healthy females were used from this cohort for the EPS8L1 ELISA. The women in this cohort had no self-reported medical history and the age of all the women was between 25-30 years of age. The samples were collected at MDC, Buch, Berlin. Women taking any medications (birth control pills) were not included in this study.

5. **Graz cohort**: A total of 141 placental samples, 92 samples from early gestation and 49 samples at term were analyzed by q-PCR for the EPS8L1 expression throughout the gestation (Figure 11). Women in this cohort had no self-reported medical history, no smoking habit and the age of all the women was more than 18 years of age.

2.2.2. Primary human trophoblast isolation

Human placental tissue samples from healthy and PE patients were collected at HELIOS Klinikum in Buch, Berlin as described earlier.[183] Primary human trophoblasts cells were isolated from placental tissue of 10 pregnant women that included healthy (n= 5) and PE (n=

5) primary trophoblast cells. The Regional Committee of the Medical Faculty of Charité Berlin approved the sampling procedure.

In this study, all the placentas were processed within 2 hours after the delivery. The whole placental tissue was washed with pre-chilled 0.9% NaCl buffer. The same buffer was used for all further tissue washing steps until the digestion procedure. The decidua parietalis and basalis were removed from the placental tissue as well as the umbilical cord. The whole process of dissection and cutting of the placenta tissue was carried out on ice. The tissue was cut into smaller pieces and washed twice with the buffer using sieve and 500 ml cylinder. The tissue was cut into smaller pieces until pale pink homogenous tissue was obtained. Small veins, capillaries, fatty and fibrotic tissue were removed and the tissue was further mashed and washed and then was transferred to two 50 ml falcon tubes (around 100g of tissue in total).

Mashed tissue from both the falcons was transferred to Erlenmeyer flask along with Buffer 1a and the mashed tissue was continuously mixed by using magnetic stirrer at 37°C (For all buffers and reagents see Materials section). Buffer 1b was added after 10 minutes of incubation. After 10 minutes of continuous mixing, Erlenmeyer cylinder was tilted and 140 ml of supernatant was aspirated. The supernatant was filtered through the gauze twice by changing cylinders and gauzes. After that Buffer 2 was added to the undigested tissue and was continuously stirred. After the 20 minutes of incubation at 37°C, the supernatant was collected and divided into four 50 ml falcon tubes (around 35 ml each) and 3, 3 ml of New Born Calf Serum (NBCS) was added to the bottom of each falcon. Falcons were centrifuged (2500 rpm, 4°C) for 10 minutes, supernatant was removed and each of the pellets was resuspended with Buffer 4 to 35 ml of overall volume from 4 falcons. The whole volume was transferred into one falcon and put into 37°C water bath with shaking. After 20 minutes of digestion with Buffer 2, whole procedure was repeated and Buffer 3 was used instead of Buffer 2 and repeated once again to get at the end three falcons with cell suspensions. Percoll (90%) 10 ml was added to each of the 3 falcons and centrifuged (1500 rpm, RT) for 10 minutes. Supernatant was discarded and the cell pellets were resuspended with cold DMEM/ HEPES (15 ml of overall volume). The total cell suspension of 15ml was placed onto the percoll gradient. Gradient was made by pouring 4 ml of different percoll solutions on top of each other, highest being 70% and the lowest 20% (Table 3). The gradient was made by pouring the percoll on the side of the 50 ml falcon which was placed tilt, using 10 ml pipette

at room temperature and after adding the cell suspension carefully, the mixture was centrifuged (3500 rpm, w/o breaks) for 20 minutes.

After centrifugation, cells from the 3rd and 4th rings (counting from the bottom) were carefully aspirated by using syringe and transferred into 30 ml of cold sterile DMEM/HEPES. Cells were centrifuged (1000 rpm) for 10 minutes to remove percoll. Cell pellet was resuspended in 30 ml of cold DMEM/ HEPES and all further steps were performed in sterile conditions. Viable cells were counted by staining the cells with Trypan Blue staining dye. Cells were then centrifuged (1000 rpm) for 5 minutes, after removing the supernatant, cell pellet was resuspended in 10 ml of cold HBSS/B (BSA 0.1%) buffer. The selection of the trophoblast cells was made by adding (140 µl) mouse anti-human HLA-ABC (monoclonal) primary antibody and it was incubated for 30 minutes at 4°C (on a shaker, w/o light). After that 40 ml of HSBB/B was added and centrifuged for 10 minutes at 1000 rpm. In the meantime, beads (Dynabeads Pan Mouse IgG) were prepared, for that 500 µl of beads was transferred into 15 ml falcon. Afterwards, the falcon was placed onto the magnet and buffer of the beads was discarded and 2 ml of HBSS/B was added. The mixture was stirred gently and was placed into magnet and the medium was discarded. Again 500 µl of HBSS/B was added to beads and mixed. The solution was added to resuspended cells after centrifugation in 8 ml of HSBB/B and incubated for 30 minutes at 4°C on a shaker. After the incubation time, the falcon with beads and cell solution was placed into the magnet for 3 minutes and cell suspension was carefully transferred to a new falcon tube. The suspension was then centrifuged at 1000 rpm for 5 minutes and pellet was resuspended in 30 ml of SmGM. The cells in the suspension were counted and the viability was also assessed. To confirm the trophoblasts cell purity, the cells were stained with cytokeratin-7 and positive cells were checked by flow cytometry. The primary human trophoblast isolation protocol gave us around 92% of pure primary trophoblast cells every time. Cells were transferred into the flask containing complete SmGM medium and were kept overnight in the incubator at 37°C. After 12-14 hours in the incubator, isolated primary trophoblasts were collected and stored at -80°C for further processing to isolate RNA, DNA and proteins.

2.2.3. Primary human trophoblasts mRNA isolation and sequencing

Trizol lysis reagent was used to isolate total RNA from primary human trophoblast cells isolated from 5 healthy controls and 5 early-onset PE patients. The procedure was according to the manufacturer`s protocol, by using Direct-zol ™ RNA MiniPrep kit including on-

column digestion with DNase I (Zymo Research). NanoDrop spectrophotometer ND-1000 was used to determine the concentration of RNA and the quality of RNA was determined by using Aglient RNA 6000 Nano Kit and Aligent 2100 Bio Analyzer machine. The Illumina TruSeq Stranded mRNA LT Set A kit (cat. no. RS-122-2101) was used to prepare library for RNA sequencing by using 550 ng of total RNA. Sample-specific indices were used to index all ten samples, which allowed for the sequencing of the five samples in two pools. The sequencing was performed as 100 bp first strand specific paired-end reads on an Illumina HiSeq 2000 platform of the BIMSB Genomics Platform of Max Delbruck Center for Molecular Medicine (Berlin, Germany). The clustering of the index-coded samples was performed on a cBot Cluster Generation System using PE TruSeq Cluster Kit v3-cBot-HS (Illumina).

Sample-specific barcoded sequencing reads were de-multiplexed from multiplexed flow cells and the resulting BCL file was converted to the FASTQ format files using CASAVA 1.8.2. FastQC was used to analyze the quality of the raw reads (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). The reads having quality score below 30 were removed. Two nucleotides were truncated from either of the ends from sequencing reads, since their quality scores were inconsistent with the rest of nucleotides. This resulted in at least 70 million reads per sample (Table 6). Next, reads were mapped over the reference genome (Human hg19/GRCh37) and transcriptome model (hg19.refseq.gtf), downloaded from UCSC tables (http://hgdownload.cse.ucsc.edu/goldenPath/hg19/bigZips/) using *Top Hat v2.0.8*, *sam tools 0.1.17.0* and *Bowtie 2.0.5.0* applying parameters as: "*tophat2 -p 8 –r 150--mate-std-dev 140 -library-type fr-firststrand*". On average 75% of the total reads could be uniquely mapped to annotated gene models (Table 6). Approximately 10% of the reads were uniquely mapped on the repeated fraction of the genome. *Cufflinks v2.08* was used to measure the transcript assembly as FPKM (Fragments Per Kilobase of transcript per Million mapped reads) for each individual sample.

For calculation of differentially expressed genes (DEGs), we calculated counts per million (CPM) using *FeatureCounts*.[184] We then implemented the algorithms from "*DESeq2*", which performed the quantification and statistical inference of systematic changes between conditions, as compared to within-condition variability.[185] The package "*DESeq2*" provides methods for testing differential expression by the use of negative binomial generalized linear models and the estimates of dispersion and logarithmic fold changes incorporate prior data-driven distributions. In addition, for the outlier samples DEGs were calculated using a single-

replicate model. The read counts were calculated with *featureCounts* from a *subread* package (http://subread.sourceforge.net/), FPKM was calculated using *bamutils* (http://ngsutils.org/modules/bamutils/count/). Next, the *RandomVariable1* (*Var1=nlx*) formula was used. Where *x*, or Random Variable2, is the expression level of this gene, while *n* reflects the sequencing depth and *l* is the gene length. The two random variables used are from the published model of the *"GFOLD"* algorithm. This calculates the normalization constant and variance to extract fold changes from un-replicated RNA-sequencing data. The bioinformatics analyses of the RNA sequencing data were done in collaboration with Manvendra Singh, from the laboratory of Dr. Zsuzsanna Izsvák.

Table 6. Human trophoblasts RNA sequencing results
The number of total reads in each sample and number of uniquely mapped reads over genes.

Sample ID	Total number of reads	Uniquely mapped reads over genes
Sample001	75909926	56873255
Sample002	74095582	55833441
Sample003	75118812	57760827
Sample004	71567772	51161663
Sample005	71067420	53587259
Sample006	79580464	60383151
Sample007	76227398	56777791
Sample008	71580658	52147477
Sample009	83151188	62222722
Sample010	79090178	59843244

2.2.4. Trophoblast-specific genes (TSGs) and transposable elements (TEs) expression analysis

RNA-sequencing data obtained from 10 samples of human primary trophoblast cells was compared to the online available data on brain, heart, liver, kidney and lungs to generate a list

of trophoblast-specific genes. The bioinformatics analyses were done in collaboration with Manvendra Singh, from the laboratory of Dr. Zsuzsanna Izsvák. The data, except the trophoblast RNA-sequencing data, was obtained from *Illumina* Human Body Map 2.0 (E-MTAB-513) with 73-83 million 50 bp paired-end and 75 bp single-end reads from five normal non-placental human tissues (brain, heart, liver, kidney and lungs). RNA-seq data from human placenta was obtained from a published study which was performed in similar layout as the rest of the data sets analyzed.[186]

All data was processed in similar fashion using the pipeline given above (Section: 2.2.3). At least 70-80% of the paired-end reads were uniquely mapped reads in each sample and were chosen for further analysis. Expression levels were calculated in Transcripts per million (TPM). Batch effect and other variations were removed by surrogate variance analysis. All samples were merged into one data frame by their genes. Low expressing genes, housekeeping genes and genes that didn't show fluctuation (Mean/ (SD > 10)) were removed prior to analysis. Tissue specific genes were computed by 'tau' algorithms by setting cutoff to 90%. [187] Similar set of analyses were performed for analyzing the expression of TEs in the above mentioned tissues. Average expression of any given TE family or locus was then calculated in CPM or RPKM. In this instance, we considered multi-mapping reads only if they were mapping exclusively within a TE family to estimate the TE family level expression in a given tissue. One alignment per read was employed to calculate the counts per million (counts normalized per million of total reads mappable on the human genome). The expression level of repeat families was calculated as Log2 (CPM+1) prior to comparison. Tau algorithms were then run to calculate tissue-specific TEs.

In order to further evaluate the ability of TEs to modulate gene expression, we paired the expression level of TEs with the expression of the genes across 10 trophoblast samples.
Correlation analysis (pairwise spearman ranked) was calculated for TEs and gene expression across the samples. This resulted in a matrix of correlational values between ~8000 genes and 18000 TEs. We melted the matrix and obtained 1.44e+ 08 pairs. Finally, we classified them as high (r > 0.70) or low (< 0.70). To get the statistically significant pairs, a *RandomForests (RF)* models were coupled with a linear regression was used. To perform these tests a random genomic background, the negative control, and a training set, the reference, were generated.

a) Training *RandomForest* for predicting TEs as distal promoters/ enhancers

For this comparison a training dataset was generated together with a genome-wide map of distal DHS-to-promoter connectivity. The DHS data was acquired from an online source and

came from 79 cell lines.[188] The genomic coordinates of distal DHS were extracted and intersected with the TE loci. Finally, we paired the obtained distal DHS-TEs with promoters based on their linkage value obtained from pairwise ranked correlations. We then classified them as high (> 0.70) or low (< 0.70). We used this classification to train cell line-specific *RandomForests (RF).*

b) *RandomForest* on paired data

After generating gene-TE pairs across the analyzed tissues by the given classification, we used *RandomForest,* [189] *gradient boosting machine* (GBM) and support vector regression (SVR) algorithms to predict the gene expression levels linked to upstream TE.

The above prediction algorithms were implemented in the R-packages of "*RandomForest"*, "*gbm*" and "*e1071*". For the RF model, we set the parameter of *n.trees = 500* (number of built trees) and selected the *best mtry* (number of variables randomly sampled as candidates at each split). For the SVR model, we selected the non-linear radial basis kernel. In the GBM model, we used the parameters [interaction. *depth = 5,,n.trees = 5000*, *shrinkage = 0.01*, *n.minobsinnode = 10* and used R_{squared}] to select the optimal model with the largest value.

We chose AUC > 0.70 as significant pairs. To investigate positional genomic relationships between TEs and trophoblast specific genes (TSGs), we filtered the TSGs. The list of significant pairs created troubleshooting as we found that several TSGs were served by TEs as their transcriptional start site (TSS) or transcriptional end site (TES). To overcome this threshold, the metric relationship of TE and TSS of protein coding genes was explored. The TE locus 10 kb upstream of the transcription start site of TSGs was arbitrarily defined to predict if a given TE is serving as the distal promoter/ enhancer for its neighboring gene. We fetched all endogenous retroviral LTR elements (ERVs) upstream of the TSGs whose pairing was significant. In search of drivers, the endogenous retroviral LTR elements (ERVs) which were over lapping with TSS, or were located within 1 kb upstream of TSS, were removed.

Trophoblast-specific gene expression of the candidate genes was additionally analyzed across tissues catalogued by the GTEx consortium (PMID: 25954002). It consists of 32 tissues across 846 individuals. Trophoblast-specific upregulation was found to be consistent with the shortlisted candidate genes as compared to the rest of the human tissues.

2.2.5. Quantitative polymerase chain reaction (q-PCR)

Total RNA was extracted according to the manufacturer's protocol from placenta tissue and trophoblast cells using QIAzol lysis reagent and Qiagen RNeasy Mini Kit (including on-

column DNAase I digestion) (Qiagen). The quality and the concentration of the total RNA was measured by NanoDrop-1000 spectrophotometer (Peq Lab). The mRNA was converted to cDNA by using High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). The real-time polymerase chain reaction (q-PCR) was performed on this cDNA to determine the expression of genes (Table 9) by using ABI 7500 Fast Sequence Detection System (Applied Biosystems). Data was analyzed by 7500 Fast System Software (Applied Biosystems). Primers and probes were designed with Primer Express 3.0 (Applied Biosystems) synthetized by Biotez (Germany) (Table 1). The expression of all of the genes was normalized to housekeeping genes (18S and GAPDH).

2.2.6. Single cell RNA-seq (scRNA-seq) data analysis

Human placental single cell RNA sequencing data (GEO accession: GSE89497; PMID: 30042384) generated on SMART-seq2 platform was reanalyzed to estimate the expression of candidate genes in the trophoblast subtypes.[190] The single cell RNA-seq analyses was done in collaboration with Manvendra Singh, from the laboratory of Dr. Zsuzsanna Izsvák. Reads were mapped to the human genome (hg19) using STAR (https://github.com/alexdobin/STAR) with the defined settings i.e. *--alignIntronMin 20 --alignIntronMax 1000000 --chimSegmentMin 15 --chimJunctionOverhangMin 15 --outFilterMultimapNmax 20* and only uniquely mapped reads were considered for the calculation of expression. Counts were obtained using *FeatureCounts* (http://subread.sourceforge.net/) at the gene level with Refseq annotations. Gene expression levels were calculated at Transcript Per Million (TPM) from counts over the entire gene (defined as any transcript located between the TSS and TES). This was done using our in-house R script (available on request). If a gene had a log2 TPM value above 1 it was regarded as expressed. We clustered the cells using the default parameters of the "*Seurat*" package from R. This applies to the most variable genes to get top 10 principle components, and the most discriminating genes cluster the cells on tSNE or UMAP. The major (cell/gene) clusters identified corresponded to Cytotrophoblasts, Syncytiotrophoblasts, Extravillous trophoblasts, macrophages and mesenchymal stromal cells. The cell type for each cluster was defined according to the known marker genes of trophoblasts and other cell types (Figure 21).

2.2.7. Chromatin immunoprecipitation DNA-sequencing (ChIP-seq) data analysis

Publically available ChIP-seq data of chromatin marks for active and inactive regions (H3K9Ac, H3K27Ac, H3K4Me1 and H3K27Me3 (GEO accession: *GSE127288* and

GSE118289)) during syncytiotrophoblast differentiation in primary culture and in human placental tissue was used. Further, ChIP-seq data on from the earliest trophoblast differentiation transcription factors (GATA2, GATA3, TFAP2A and TFAP2C (GEO accession: *GSE105081;* PMID: 29078328) was also used.[191,192,193] The analyses was done in collaboration with Manvendra Singh, from the laboratory of Dr. Zsuzsanna Izsvák.

The ChIP-seq datasets were downloaded in the raw FASTQ format. ChIP-seq reads were aligned to the hg19 human reference genome using *Bowtie2* (http://bowtie-bio.sourceforge.net/bowtie2/manual.shtml) in the *--very-sensitive-local* mode. All unmapped reads, reads with MAPQ < 10 and PCR duplicates were removed using *Picard* and *SAMtools* (http://www.htslib.org/). All ChIP-seq peaks were called by *MACS2* (https://github.com/taoliu/MACS) with the parameters in the *narrow mode* for TFs and *broad mode* for histone modifications, while keeping FDR < 1%. Blacklisted regions were excluded from called peaks (https://www.encodeproject.org/annotations/ENCSR636HFF). The peak sets were intersected with repeat elements from hg19 repeat-masked coordinates using *bedtools* (https://bedtools.readthedocs.io/en/latest/) *intersectBed* with 50% overlap. To enable visualization through IGV over Refseq genes (hg19), the raw signals from the ChIP-seq were obtained from *MACS2*, using the parameters: *-g hs -q 0.01* -B. The conservation track was visualized through UCSC genome browser under net/chain alignment of the given non-human primates (NHPs) and were merged beneath the IGV tracks.

2.2.8. Cell culture

Two trophoblast model cell lines were used in this study, choriocarcinoma line (BeWo) and transformed trophoblast lines, first trimester human extravillous trophoblast cell line (SGHPL-4). BeWo cells were maintained in DMEM/ F12-GlutaMAX supplemented with 10% (v/v) FBS and 1% Antibiotic-Antimycotic (AA). SGHPL-4 cells were cultured in Ham`s F10 medium supplemented with 10% (v/v) FBS and 1% AA. Cells were cultured at 37°C in 5% CO_2.

2.2.9. Genomic DNA extraction

Frozen placenta tissues were pulverized with liquid nitrogen. According to the manufacturer's instructions, genomic DNA from placentas was extracted using the DNeasy Blood and Tissue Kit (Qiagen, cat. 69504). Genomic DNA was stored at -20°C. Genomic DNA was used for amplification and cloning of the candidate endogenous ERV-LTR sequences.

2.2.10. ERV-LTR constructs

Sequences of ERV-LTRs of all the candidate genes were downloaded from repeat masker from UCSC human genome browser version hg19. ERV-LTRs were amplified from genomic DNA of human placenta tissue by using sequence specific primers designed by using NEBuilder High Fidelity DNA assembly cloning tool (v1.12.15). Generated primers had overhangs with backbone of the plasmid. PT2-GFP reporter construct was used to clone the ERV-LTR upstream of the CAGs promoter by using Eco-RV site. A total of 9 constructs were generated and as a negative control a construct without an ERV-LTR was used for the assay. As an internal control an m-cherry construct having the same CAGs promoter was used. To calculate number of pmols of each fragment for optimal assembly, based on fragment length and weight, the following formula was used:

pmols = (weight in ng) x 1,000 / (base pairs x 650 daltons)

50 ng of 5000 bp dsDNA is about 0.015 pmols.

50 ng of 500 bp dsDNA is about 0.15 pmols

PCR was done on placental genomic DNA to amplify the ERV-LTR sequences by using Hybrid DNA polymerase (EURx) and vector backbone was digested with the Eco-RV restriction enzyme. HiFi DNA assembly reaction was setup by taking a ratio of 1:2 = vector: insert and total amount of the DNA was around 0.2 pmols. HiFi DNA master mix was used and the reaction mix was incubated at 50°C for 15 minutes. After incubation, ligation reaction was transformed into the Mix & Go (DH5α) cells. The colonies were picked after 24 hours and were sent for sequencing. The positive clones were then further sequenced to make sure there were no mutations in ERV-LTR sequence.

2.2.11. GFP reporter assay

Wild-type BeWo and SGHPL-4 cells were cultured for 24 hours in a 12-well plate. Retroviral LTR reporter constructs were transfected by using Lipofectamine 3000 Transfection Reagent according to the manufacturer's protocol. M- Cherry construct was transfected in each well as an internal control along with ERV-LTR containing constructs. As a control, one well of plate was transfected with a GFP construct without any sequence, one with m-cherry construct only, one well with both the GFP and m-cherry constructs and one well remained un-transfected. After 48 hours of transfection in the trophoblast cells, both GFP and m-Cherry signals were measured independently for each transfected ERV-LTR construct by flow

cytometry. FlowJo_V-10_CL software was used to analyze FACS data and the population of cells co-expressing both GFP and RFP was selected for analysis. The median florescence intensity (MFI) was calculated for the GFP signal and was normalized to the MFI of the m-Cherry signal of each sample. The experiment was repeated 3 times with 3 technical replicates in each experiment.

2.2.12. Western blot

Equal numbers of cells were lysed using 6X SDS loading buffer and were cooked at 96°C for 10 minutes. Denaturized proteins were separated on 10% SDS-PAGE gel. Proteins were electrophoresed in stacking gel with 90V until they reached resolving gel and then with 180V until the end. To transfer proteins from gel to PVDF membrane, trans-blot turbo transfer system was used. Membranes were equilibrated in 100% ethanol for 1 minute and proteins were transferred onto the membranes using trans-blot turbo transfer system 25V for 10 minutes. Membranes were blocked with 5% w/v non-fat dry milk in TBS-T for 1 hour. After blocking, blots were incubated with needed concentration of antibody and incubated overnight at 4°C with gentle shaking. Secondary antibody conjugated with horseradish peroxidase was used and blots were incubated for minimum 1 hour, then membranes were developed with Amersham ECL prime western blotting detection reagent (GE Healthcare). Mild stripping buffer was used to strip the membranes (Abcam protocol), after that they were blocked again and incubated with loading control antibody. The antibodies used in the study and concentrations are mentioned in Table 2. ImageJ software was used to quantify the protein bands of EPS8L1 isoforms from western blot films. The quantification data shows the relative amounts as ratio of each isoform band relative to loading control (Actin) (Figure 12C).

2.2.13. Enzyme linked immunosorbent assay (ELISA) of EPS8L1

Serum samples from Manchester cohort (high risk pregnancy cohort) were used to analyze EPS8L1 expression in PE patients as compared to healthy controls. Controls enrolled in this cohort had a prior history of hypertension and diabetes and had not developed PE later on during their pregnancy. Controls were gestational age matched with the PE patient and the gestational age (GA) was between 24^{th} to 28^{th} weeks of gestation. Serum EPS8L1 levels were analyzed in 12 PE, 12 healthy and 6 non-pregnant control serum samples according to the manufacturer's instructions (MyBioSource, MBS9317820, USA). Non-pregnant serum

samples were collected in Berlin (Menstrual cycle study cohort). Each sample was tested twice for EPS8L1 levels and mean of the two measurements was used for statistical analysis.

2.2.14. Immunohistochemistry (IHC) staining of primary villi and term placenta tissue (human)

Paraffin embedded placenta tissue sections from healthy and PE patients were deparaffinized. Antigen retrieval was done by heat-induced epitope retrieval (HIER) method in 0,01M citrate buffer pH 6. Peroxidase activity was blocked with hydrogen peroxide (Lab Vision/ Thermo Scientific) for 10 minutes. Washing was performed by using tris-buffered saline with 0.05% Tween-20 (v/v) (TBS-T). To block non-specific immunoglobulin binding, ultra V Block (Lab Vision) was used for 7 minutes. Primary antibody (rabbit anti-EPS8L1 antibody (Atlas Antibody)) was diluted (1:250) in antibody diluent (Dako) and was incubated for 45 minutes. To identify non-specific tissue staining, antibody diluent without the primary (anti-EPS8L1) antibody was used as a negative control. Slides were washed 3X with TBS-T. After that, the secondary antibody conjugated with HRP polymer (ready-to-use; Lab Vision) was used and incubated for 20 minutes. Slides were incubated in dark with Aminoethylcarbazole 45 chromogen (ready-to-use; Lab Vision) for 10 minutes and peroxidase activity was detected. Tissue sections were washed again with TBS-T and were counterstained with Haematoxylin for nuclei staining and mounted with Kaiser's glycerine gelatine (Merck). Stained placenta tissue slides were analyzed by Dr. Martin Gauster from the Institute for Cell biology, Histology and Embryology of Medical University of Graz, Austria.

2.2.15. Immunostaining of EPS8L1 in SGHPL-4 cells

Equal number of (0.1 x 10^6 cells/ well) of WT-SGHPL-4 and OE-*EPS8L1*_SGHPL-4 cells were seeded on coverslips in a 12-well cell culture plates and incubated in 10% (v/v) FCS Ham´s F10 medium, 1% AA overnight at 37°C and 5% CO_2. Cells were washed twice with PBS, then were fixed using 4% paraformaldehyde (Sigma) supplemented with Hoechst 33, 342 (1: 250, Invitrogen) in PBS for 15 minutes and permeabilized with 0.1% Triton X-100 in PBS for 2 minutes. Cells were blocked by using 5% BSA in DPBS for 1 hour. Coverslips were then incubated with primary antibodies for actin and rabbit anti-EPS8L1 antibody 1:100 (5% BSA prepared in PBS) overnight at 4°C then washed 3X with PBS and incubated with secondary antibodies Alexa Fluor 488 anti-rabbit (goat) and Alexa Fluor 647 anti-mouse (donkey) 1:200 for additional 1 hour. After extensive washes with PBS, nuclei were stained with DAPI solution (Vectashield with DAPI) for 10 minutes at RT in dark. After additional

washing, the samples were mounted using ProLong® Gold Antifade Reagent (Invitrogen) and images were taken using a Leica LSM710 point-scanning single photon confocal microscope.

2.2.16. Knock-out (KO) of *EPS8L1*

To study the effect of KO of *EPS8L1* in trophoblast cells, CRISPR-Cas9 system was used and the protocol was followed as published by Broad Institute.[194] The sgRNAs targeting *EPS8L1* were designed *in silico* by using CRISPOR Design Tool (http://tools.genome-engineering.org). CRISPOR Design Tool identified suitable target sites to design oligos and primers for preparing sgRNA. The top three recommended oligos (phosphorylated) were ordered from BioTez. Top (100uM) and bottom strands (100uM) of phosphorylated oligos were mixed in ddH_2O for annealing at RT for 30 minutes. The sgRNA oligo inserts were cloned into a plasmid backbone containing Cas9 and mVenus (gRNA Cas9 mVenus). Ligation reaction for each sgRNA was setup as mentioned in Table 7 and was incubated for 1 hour (12 cycles of 37°C for 5 minutes and 21°C for 5 minutes). After incubation, the ligation mix was transformed into Mix & Go (DH5α) cells. Colonies were picked and after plasmid purification, the constructs were sent for sequencing. Three different guide RNAs were designed to access their ability to mediate targeted *EPS8L1* cleavage. Guide RNAs containing constructs were transfected into WT-BeWo cells along with a control plasmid having no insert. Transfection was performed by Lipofectamine 3000 (Invitrogen) transfection reagent according to the manufacturer's instructions. After 48 hours, FACS was used to sort the GFP positive cells. After FACS sorting, the cells were cultured together in a dish for 5-7 days. Western blot was performed on total protein isolated from cells to confirm the KO of *EPS8L1*.

Table 7. CRISPR reaction setup

Components	Amount(ul)
Backbone (100ng) (Cas9 mVenus)	1
Diluted oligo insert	2
Cutsmart Buffer 10X	2
DTT,10mM	1
ATP, 10mM	1
FastDigest *Bb*sI	1
T4 ligase	0.5
ddH_2O	Up to 20ul

2.2.17. Overexpression (OE) of *EPS8L1*

To generate a stable overexpression of *EPS8L1* in a trophoblast cell line, pT2B-puro *Sleeping Beauty* construct having the cassette (pCAGGS-HA-N-DLX5) was used. The expression construct contains an N-terminal Human influenza hemagglutinin (HA)-tagged human *EPS8L1* open reading frame (ORF), driven by the ubiquitous CAGs promoter/ enhancer (chicken β-actin promoter/ enhancer coupled with the cytomegalovirus (CMV) immediate-early enhancer). Puromycin cassette controlled by the SV40 (Simian virus 40) promoter was used for selecting cells containing the stably integrated transposon (Figure 16A). Puromycin selection was carried out for more than 12 days to strictly select for the stably expressing EPS8L1 cells. Human *EPS8L1* coding sequence (Sequence ID: RefNM_133180.2) was amplified by PCR from placental cDNA with gene specific primers generated by NEBuilder HiFi DNA Assembly tool, harboring restriction sites for cloning into pT2B-puro *Sleeping Beauty* vector. PCR reaction was carried out with Kappa polymerase (Roche) on 50 ng of cDNA template from human placenta with 0.2mM dNTPs, 0.5µM primers, GC buffer (10 µl) in a final volume of 50 µl. Cycling conditions were: 95°C for 3 minutes, followed by 35 cycles of 95°C for 20 seconds, 55°C for 20 seconds and 72°C for 3 minutes and final extension of 10 minutes at 72°C. The PCR product of ~ 2172 bp was extracted from 1% agarose gel; the vector backbone was digested with Eco-RV and Not-I restriction enzymes and was gel purified by using Qiagen gel purification kit. HiFi ligation reaction was prepared by using the gel purified plasmid backbone (0.05 pmol) and *EPS8L1* amplified transcript (0.1pm) and mixed with HiFi mix (10 µl) and water (6 µl) and then incubated at 50°C for 30 minutes. After incubation, the ligation mix was transformed to bacteria (DH5α) to clone the plasmid. Colonies were picked from the plate and sent for sequencing. The positive clone was then further sequenced to make sure there is no mutation in the coding gene sequence.

2.2.18. Mass spectrometry (MS) of EPS8L1

The protein interactors of EPS8L1 were determined by immunoprecipitation (IP) of HA-tagged EPS8L1. HA tagged EPS8L1 was overexpressed transiently in WT-SGHPL-4 cells and WT-BeWo cells and the same construct without the HA tag was used as a control. Transfection was performed by Lipofectamine 3000 (Invitrogen) transfection reagent according to the manufacturer's instructions. Three independent transfections were performed to have three biological replicates. After 48 hours, the cells were collected via cell scraper

without using trypsin. Cells were washed 2X with PBS and were centrifuged at 4000 rpm. For the last washing step 200 µl of PBS was added, from which 50 µl was used for the western blot analysis to confirm the transfection. Cell pellet was dissolved in 600 µl of lysis buffer (including benzonase). Cells were mixed gently and incubated on a shaker at 4°C for 40 minutes. Cells were centrifuged at 14000 rpm for 10 minutes and the lysate was transferred into the new Eppendorf tube.

In the meantime, beads (EZview TM Red Anti-HA Affinity Gel) were prepared. For 12 samples, approximately 60 µl of beads per sample were used and washed with 5 ml of washing buffer twice (lysis buffer, 10X diluted). Centrifugation was carried out at 3000 rpm for 5 minutes. Washing buffer was added to make the final solution 2 ml along with beads. The washed red beads (160 µl) were added into the cleared cell lysate and incubated for 2 hours at 4°C in the shaker. Supernatant was separated from the beads by centrifugation at 3000 rpm. Beads were washed 4 times with 600 µl of washing buffer each time. Centrifugation was performed at 3000 rpm for 1 minute. After that, the beads were washed twice with pre-elution buffer (500 µl) and transferred to 1.5 ml new Eppendorf tubes. For protein elution, 300 µl of glycine (pH 2. 5) was added to the beads and incubated for 5 minutes at room temperature with continuous rotation. Tubes were centrifuged at 3000 rpm for 10 seconds. Supernatant was transferred to the column with the filter and centrifuged for 10 seconds. 15 µl of protein mix solution was used for western blot and the remaining 285 µl solution was precipitated by adding 70 µl of 2.5 NA, Acetate (pH 5.0) in each sample tube and each tube was filled with chilled ethanol (99%). Samples were stored overnight at 4°C. Tubes were centrifuged at 20,000 x g for 45 minutes at 4°C, ethanol was removed completely and the pellets were air-dried. Protein pellets were sent to the proteomics facility at MDC for mass spectrometry analysis.

2.2.19. *EPS8L1* overexpression (OE) and transcriptome analysis

Wild-type SGHPL-4 cells were transfected with pT2B-puro *Sleeping Beauty* construct containing the *EPS8L1* overexpression cassette to establish the stable overexpression *in vitro* model. When the cells reached 70% confluency after the 13th passage, they were washed twice with PBS without Ca^{2+} and Mg^{2+}, and harvested with trypsin. The cells were collected in Ham´s F10 medium, 10% (v/v) FCS counted and centrifuged. After two subsequent PBS washing steps, total 1.4x10^5 SGHPL-4 cells were resuspended in 20 µl Resuspension Buffer R (Neon 10 µl, Invitrogen). The buffer also included a 10:1 ratio of the vector carrying the

EPS8L1 overexpression cassette (1 µg DNA), and the plasmid pcGlobin2-SB100X. With the latter containing the hyperactive *Sleeping Beauty (SB)* transposase (100 ng DNA) which facilitates the genomic integration of the vector. They were electroporated with the Neon Transfection System (Life Technologies), using the parameters: pulse voltage: 1260, pulse width: 20ms and number of pulses were 2. After the electroporation the cells were cultured in 6-well plates after containing 2 ml Ham´s F10 medium with 10% (v/v) FBS. Two days post-transfection puromycin was introduced (3 µg/ml) and used for 15 days to select for cells with successful integrations. *EPS8L1* overexpression was confirmed by q-PCR and Western blot. Its expression and localization in the cells was determined by immunofluorescent staining.

Transcriptome sequencing of *EPS8L1*-overexpressing SGHPL-4 cells (OE-EPS8L1_SGHPL-4 cells) was performed by BGI on the BGISEQ-500 platform. Total RNA was isolated from the wild-type (WT) and OE-*EPS8L1*_SGHPL-4 cells using the RNeasy mini kit (Qiagen). Three independent transfections were performed and used as biological replicates. The minimum concentration of total RNA was 1000 ng of all the six samples that were sent to the company. All the samples passed the quality check RIN ≥ 7.0 28S/ 18S ≥ 1.0 and were processed for sequencing. After the sequencing run, the raw reads were filtered by BGI. Data filtering included removing adaptor sequences, contamination and low-quality reads from raw reads. The statistics of data generated are shown in the Table 8A. Column description for data production is described in a separate Table 8B.

Table 8. **OE-*EPS8L1* and WT-SGHPL-4 cells RNA-sequencing data statistics**

A)

Sample name	Clean reads	Clean bases	Read length (bp)	Q 20 (%)	GC (%)
BGI-1	63,310,194	6,331,019,400	100	97.64 %	49.98 %
BGI-11	60,672,626	6,067,262,600	100	97.45 %	50.10 %
BGI-3	71,589,886	7,158,988,600	100	97.72 %	50.49 %
BGI-4	67,364,740	6,736,474,000	100	97.54 %	50.69 %
BGI-6	67,534,274	6,753,427,400	100	97.62 %	50.82 %
BGI-9	72,579,966	7,257,996,600	100	97.86 %	50.41 %

B)

Header	Description
Sample name	Sample name that identify each sample
Read length	The length of the total reads (bp)
Clean reads	Total clean reads number
Clean bases	Total clean bases number
Q20 %	The number of nucleotide with quality higher than 20/nucleotide (clean read1, read2)
GC %	GC number / nucleotide (clean read 1, read 2)

Original image data was transferred into sequence data via base calling, which is defined as the raw data or raw reads and was saved as FASTQ file. The bioinformatics analyses of the RNA sequencing data were done in collaboration with Manvendra Singh, from the laboratory of Dr. Zsuzsanna Izsvák. The methodology to calculate the differentially expressed genes is similar to the pipeline described in the data analysis of RNA isolated from human primary trophoblasts (Section = 2.2.3).

2.2.20. Intersection of PE patients and OE-*EPS8L1*_SGHPL-4 cells datasets

The transcriptome of the overexpressing *EPS8L1*-cells and PE patients was compared and analyzed. Two PE microarray datasets generated by Zadora et al. (2017) (PMID: 28904069) were used for this analysis. The data was analysed in collaboration with Manvendra Singh from the laboratory of Dr. Zsuzsanna Izsvák. A matrix of the expression levels for the unique genes in each sample was generated and the two datasets from different platforms were merged by their unique gene names for 24 PE and 22 control samples. The batch effect arising from two different platforms was corrected by normalizing surrogate variances from the *"Combat"* package from the R Bioconductor. The corrected batch effect was confirmed by a Principal Component Analysis (PCA). Each gene expression value was further assigned by their relative abundance value, which is the ratio of the expression level of a gene in each sample and the mean expression values of the gene across all the samples. Lastly, the *EPS8L1* gene was fetched from the obtained relative abundance matrix. A correlation analysis (Pairwise Spearman Rank) was performed on Log2 fold-change to identify the commonly dysregulated genes between the differentially expressed genes (DEGs) in the PEs vs controls,

and *EPS8L1*-OE vs controls. The DEGs were obtained by setting the criterion of the adjusted *P*-value at < 0.05.

2.2.21. Invasion assay

Transwell invasion assay was performed on OE-*EPS8L1*_SGHPL-4 cells and WT-SGHPL-4 cells. The transwell inserts were pre-chilled and then coated with 50 µl of growth factor reduced matrigel (0. 25 mg/ml) diluted in Hams F10 w/o FBS. The plate containing the inserts was then incubated for 24 hours in a cell culture incubator at 37°C. OE-EPS8L1_SGHPL-4 cells and WT-SGHPL-4 were serum starved by incubating them for 24 hours in Ham's F10 media containing only 0.5% FBS and were incubated for 24 hours at 37°C.After 24 hours, transwell inserts were prepared in two steps. In the first step, equal number of serum starved WT-SGHPL-4 cells and OE-*EPS8L1*_SGHPL-4 cells were plated onto the matrigel coated wells. Serum starved cells were rinsed with 5 ml DPBS (without Ca^{2+} and Mg^{2+}), trypsinized by 1 ml trypsin and incubated at 37°C for 3-5 minutes and then 3-4 ml of Ham's F10 media with 0% FBS was added to the cells. Cells were counted by hemocytometer and centrifuged at 300 x g for 5 minutes. After removing the supernatant, cells were resuspended in 0% FBS Ham's F10 to obtain the cell suspension concentration of 5 x 10^5 cells /ml. From this suspension culture, 200 µl of cells (5 X 10^4) per well were plated onto the matrigel coated inserts. In the second step, chemoattractant was prepared. EGF (10 ng/ml) was used as a positive control and was prepared in 10% FBS in Ham's F10 and as a negative control 0% FBS in Ham's F10 was used. Rest of the wells were filled with 750 µl of 10% FBS/Ham's F10 per well.

Cells were incubated at 37°C for 12-14 hours for invasion assay. After 12-14 hours of incubation, invaded cells were fixed and stained. Media was removed from the top and the bottom of the insert; the inserts were washed twice with PBS. To fix the cells 3.7% PFA was added and the cells were incubated for 2 minutes at room temperature. Cells were washed again 2X with PBS to completely remove the PFA. For permeabilization, 100% methanol was added and the cells were incubated for 20 minutes at room temperature. The transwell inserts were washed twice and the cells both invaded and non-invaded were stained with 0.2% crystal violet. The plate containing the cells was covered with aluminum foil and then incubated at room temperature for 15-20 minutes. Cells were washed again after incubation. Non-invaded cells were scraped off by using cotton swabs. Images were taken under the light

microscope (bright field). Images were analyzed and invaded cells were counted by Fiji/ ImageJ software. Experiment was repeated at least three times.

2.2.22. Tube formation assay (Angiogenesis assay)

For the tube formation assay, *u*-slide angiogenesis (Ibitreat) were coated with 10 µl/ well growth factor reduced matrigel. A total of 8000 cells/well (50 µl/well) both OE-*EPS8L1*_SGHPL-4 cells and WT-SGHPL-4 cells (control) were coated on to the matrigel and were incubated for 6 hours at 37°C. EGF was used to stimulate the tube formation (angiogenesis) and 0% FCS was used as a negative control. Images were taken after 6 hours and were analyzed by using the Wimasis Image Analysis software. Total number of tubes, tube length, loops and branching points were quantified by the software between OE-*EPS8L1*_SGHPL-4 cells and WT-SGHPL-4 cells. Experiment was repeated at least three times.

2.2.23. DCFH-DA assay (ROS assay)

A total of 5 X 10^4 OE-*EPS8L1*_SGHPL-4 cells and WT-SGHPL-4 cells (control) per well were seeded on a 24-well plate in 1 ml of Ham´s F10 medium with 10% (v/v) FCS, 1% AA and were incubated for 24 hours in 37°C, 5% CO_2. Cells were labelled with 0.05 mM DCFH-DA for 45 minutes. DCFH-DA diffuses into the cell, becomes deacetylated by cellular esterases to non-fluorescent DCFH, which is then oxidized to fluorescent DCF by ROS (Figure 20A). GFP fluorescence intensity is proportional to the ROS inside the cell. After incubation, cells were washed 3 times with PBS, trypsinized and GFP signal was analyzed on FACS caliber system using Cell Quest Software. Data is presented as a median and mean fluorescent intensity (MFI) of the GFP signal. Experiment was repeated at least three times.

2.2.24. Statistics

Statistical analysis was performed with GraphPad Prism (GraphPad Software) and R Statistical Programming Language. Data are presented as either mean ± SEM (for normally distributed data) or median with inter-quartile range (for non-normally distributed data). Normal distribution was assessed by D'Agostino & Pearson omnibus normality test, Shapiro-Wilk normality test and Kolmogorov-Smirnov tests. Data sets were compared using the Unpaired t-test, Mann–Whitney test, Wilcoxon-rank sum test, ANOVA and Kruskal-Wallis test as appropriate. Post hoc testing included Dunnet's test and Dunn's test for multigroup comparisons. Statistical correlation analysis between nonparametric variables was performed

using Spearman´s rank correlation. Techniques for each analysis are specified in the figure legends. Two-tailed testing with normal-based 95% confidence interval was performed for each analysis and $P<0.05$ was considered statistically significant.

3. Results

3.1. Search strategy to identify trophoblast-specific genes associated with ERV-LTRs

To identify trophoblast-specific genes that are regulated by ERV-LTRs and are also dysregulated in PE patients, RNA sequencing (RNA-seq) data of human primary trophoblasts (done in our lab) was compared with the online available RNA-seq data on human brain, heart, liver, kidney, lungs and placenta tissue (*Illumina* Body Map and GTEx). The comparative analysis (as mentioned in methods section 2.2.4) generated a list of 335 genes that showed higher expression in the placenta tissue and trophoblast cells as compared to other tissues. To identify which trophoblast-specific genes (candidate genes= 335) are regulated by ERV-LTRs, first, the transposable element (TE) families that have enriched expression in the human placenta as compared to other human tissues including liver, lung, brain, kidney, heart and placenta were identified (Figure 6). The analysis identified certain TEs families to be enriched in the human placenta tissue, specifically. From the three clusters of TEs families that showed enriched expression in placenta tissue as shown in the heat map (Figure 6), I focused on those TE families that are also primate-specific, including endogenous retroviral elements (ERVs) (encircled in purple, Figure 6). In order to identify the role of TE-derived regulatory sequences as tissue-specific enhancers, a workflow was designed to identify those TE-derived sequences (primate-specific) that potentially performs a regulatory role in gene expression in human trophoblast cells. To find the TE-gene pairs (e.g. enhancer-target), the expression of the genes that showed higher expression in trophoblasts (candidate genes=335) was compared with the expression of the TEs located within 10-kb region upstream of the transcription start site (TSS) of the genes (as mentioned in methods section 2.2.4). This approach identified 88 genes that showed co-regulated expression with TEs inserted in the gene vicinity, suggesting that these TEs might be involved in regulating the neighboring trophoblast-specific genes.

In addition, I asked whether the TE-derived dysregulation of genes might contribute to the pathology of certain pregnancy-associated diseases, e.g. PE. Thus, from the 88 genes, I focused on those genes that were dysregulated in PE patients. In order to identify the dysregulated genes in PE patients, microarray data of the published Oslo-Cohort-I (24 PE patients compared with 22 healthy controls) and RNA-seq data of primary human trophoblasts (early-onset PE patient's (n=5) and healthy controls (n=5)) were analyzed and from the list of 88 genes, 16 were found to be dysregulated in PE patients (Table 9). Interestingly, the 16 shortlisted

candidate genes were found to be associated with primate-specific endogenous retroviral LTR elements (ERV-LTRs) that were also found to be enriched in the placenta tissue and the trophoblast cells (Table 9). For further analysis, I focused on these 16 candidate genes as mentioned in the table 9. First I identified the real dysregulated candidate genes in PE patients and for that I analyzed the expression of the candidate genes in a bigger sample cohort that consist of 24 controls and 50 PE patients by q-PCR (Figure 7). Next question was to validate their trophoblast-specific gene expression and for that I established a pregnancy related tissue panel and analyzed the expression of the dysregulated candidate genes (Figure 8). In order to validate the potential activity of primate-specific ERV-LTRs associated with the genes, I performed GFP-reporter assay in the trophoblast cell line (Figure 9). Since my interest was to characterize the function of dysregulated ERV-LTR regulated trophoblast-specific genes in PE pathogenesis, so in this project, I focused on one candidate gene *(EPS8L1)* and identified its role in PE pathogenesis which is explained further in detail.

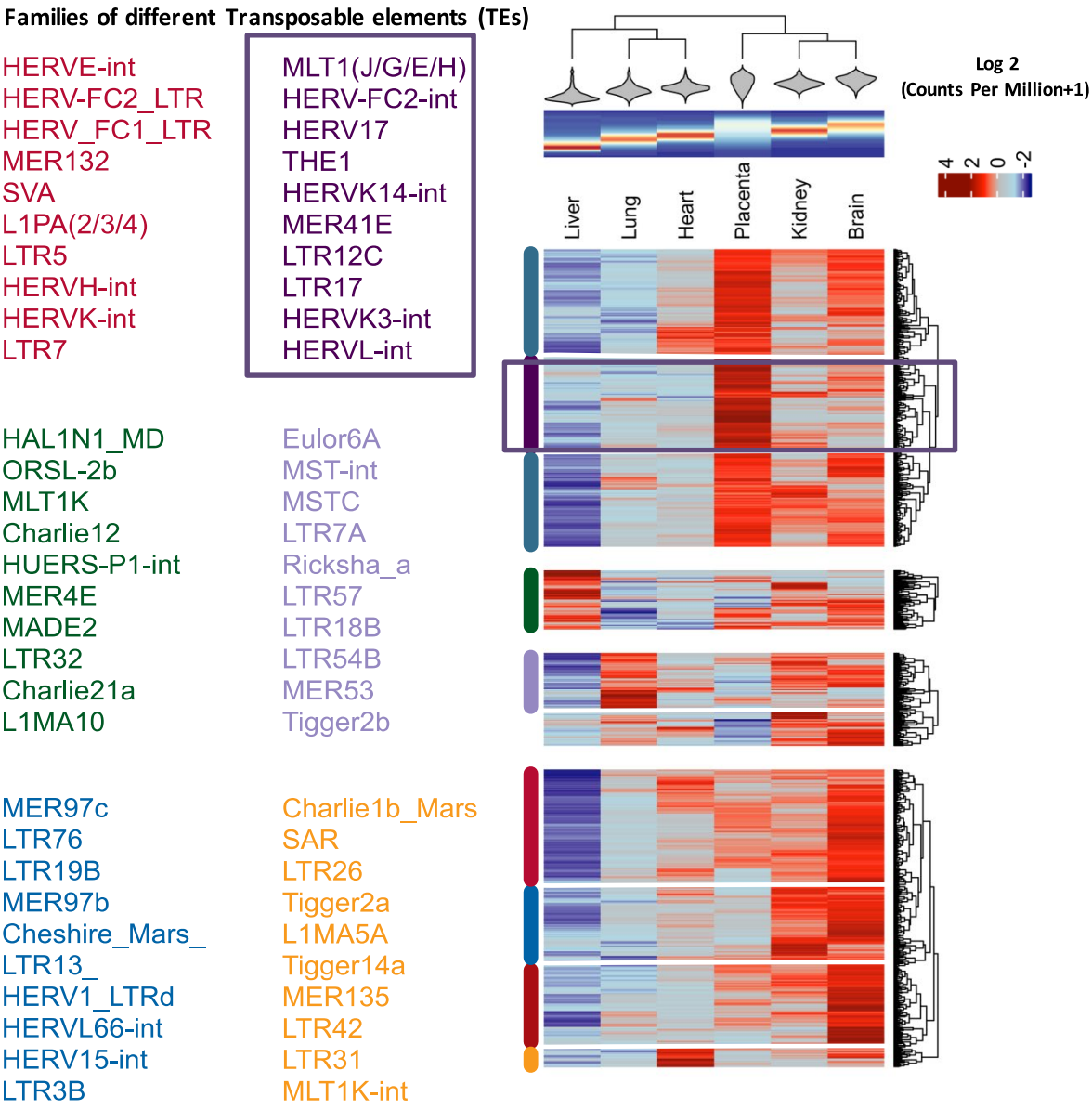

Figure 6. Expression of transposable elements (TEs) in human tissues
Heat map displaying the scaled expression (Log 2 (counts per million+1) of 954 expressed TE families across the 6 human tissues (Illumina Body Map 2.0; E-MTAB-513). Each row represents a single TE family and columns represent tissues. Group of TE families are shown as 10 distinct k-means clusters on the basis of their dynamic expression. Given clusters are manually annotated by 10 TE families which show the variable expression across the tissues but enriched in the respective tissues. Color codes are used to match the clusters with their annotations. Encircled TE families (purple box) are those which are strongly expressed in placenta tissue and are primate-specific as compared with other human tissues.

Table 9. Short-listed trophoblast-specific genes with their associated ERV-LTRs
Genes that are already known to be placenta-specific and have been reported to be dysregulated in PE patients were excluded for further analyses (underlined, N/A means not applicable).

No.	Genes	Retroviral Elements (LTRs)	Dysregulated in PE women (Oslo Cohort)
1	**CYP11A1** Cytochrome P450 Family 11 Subfamily A Member 1	MLT1F2, MLT1J	✓
2	**ALDH3B2** Aldehyde Dehydrogenase 3 Family Member B2	MLT1F2	✓
3	**EPS8L1** Substrate for the epidermal growth factor receptor (predicted)	MLT1G1	✓
4	**CGA** Chorionic Gonadotropin, Alpha Polypeptide	MLT1F2	N/A
5	**NRK** Nik Related Kinase	MER2, MLT1C	×
6	**C1QTNF6** Complement-C1q Tumor Necrosis Factor-Related Protein 6	MER41B, MLT1C	×
7	**SIGLEC6** Sialic Acid Binding Ig Like Lectin 6	MLT1B	N/A
8	**CRH** Corticotropin Releasing Hormone	MLT2A2, MLT1B	N/A
9	**KIF23** Kinesin Family Member 23	MER39	×
10	**SLC22A11** Solute Carrier Family 22 Member 11	LTR16A, MER39, MER41D	×
11	**RAD51** RecA-Like Protein	MER1B, LTRC, MER39	×
12	**SPINT1** Serine protease inhibitors	MLT1F2	✓
13	**CSF2RB** GM-CSF/IL-3/IL-5 Receptor common beta-chain	MLT1C, MLT1L	✓
14	**DACT2** Modulators of Wnt signaling pathway	MLT2B4, MER41D	✓
15	**PLEKHA8** Pleckstrin Homology Domain Containing A8	MLT2B4	×
16	**PHYHIPL** Phytanoyl-CoA 2-Hydroxylase Interacting Protein Like	MER39B	×

[1] **HGNC-HUGO Gene Nomenclature Committee**

3.2. Expression of trophoblast-specific genes in PE patients

To validate the differential expression of candidate genes in PE patients, quantitative polymerase chain reaction (q-PCR) was performed on human placenta tissue samples of Oslo-cohort-II (Control=28, early-onset PE=24, late-onset PE=22, IUGR+PE=6). Quantitative-PCR data indicated no significant differences in the gene expression of *C1QTNF6, KIF23, RAD51*, *PLEKHA8, NRK, PHYHIPL* and *SLC22A11* in PE patients (early-onset and late-onset PE) when compared to healthy controls (*NS*: Not Significant; Kruskal–Wallis test; Dunn's multiple comparisons test). (Figure 7A). In contrast, gene transcripts of *CYP11A1* and *CSF2RB* were found to be significantly upregulated (***P*< 0.001, **P*< 0.01; 1-way ANOVA; Dunnett's multiple comparisons) along with *EPS8L1* (**P*< 0.01; Kruskal–Wallis test; Dunn's multiple comparisons test) in the early-onset PE patients as compared to the healthy controls, whereas the transcripts of *SPINT1, ALDH3B2* and *DACT2* were found to be significantly downregulated (**P*< 0.01; Kruskal–Wallis test; Dunn's multiple comparisons test) in early-onset PE patients as compared to the healthy controls (Figure 7B). In late-onset PE patients, increased expression of *CYP11A1* (***P*< 0.001; 1-way ANOVA; Dunnett's multiple comparisons) and decreased expression of *SPINT1* (**P*< 0.01; Kruskal–Wallis test; Dunn's multiple comparisons test) was detected as compared to the healthy controls. However, *CSF2RB* (1-way ANOVA; Dunnett's multiple comparisons)*, ALDH3B2, EPS8L1* and *DACT2* (Kruskal–Wallis test; Dunn's multiple comparisons test) were not found to be significantly dysregulated in late-onset PE patients as compared to the healthy controls. Altogether, out of 16 short-listed candidate genes, q-PCR confirmed the dysregulation of six genes in PE patients (Figure 7B, Table 9).

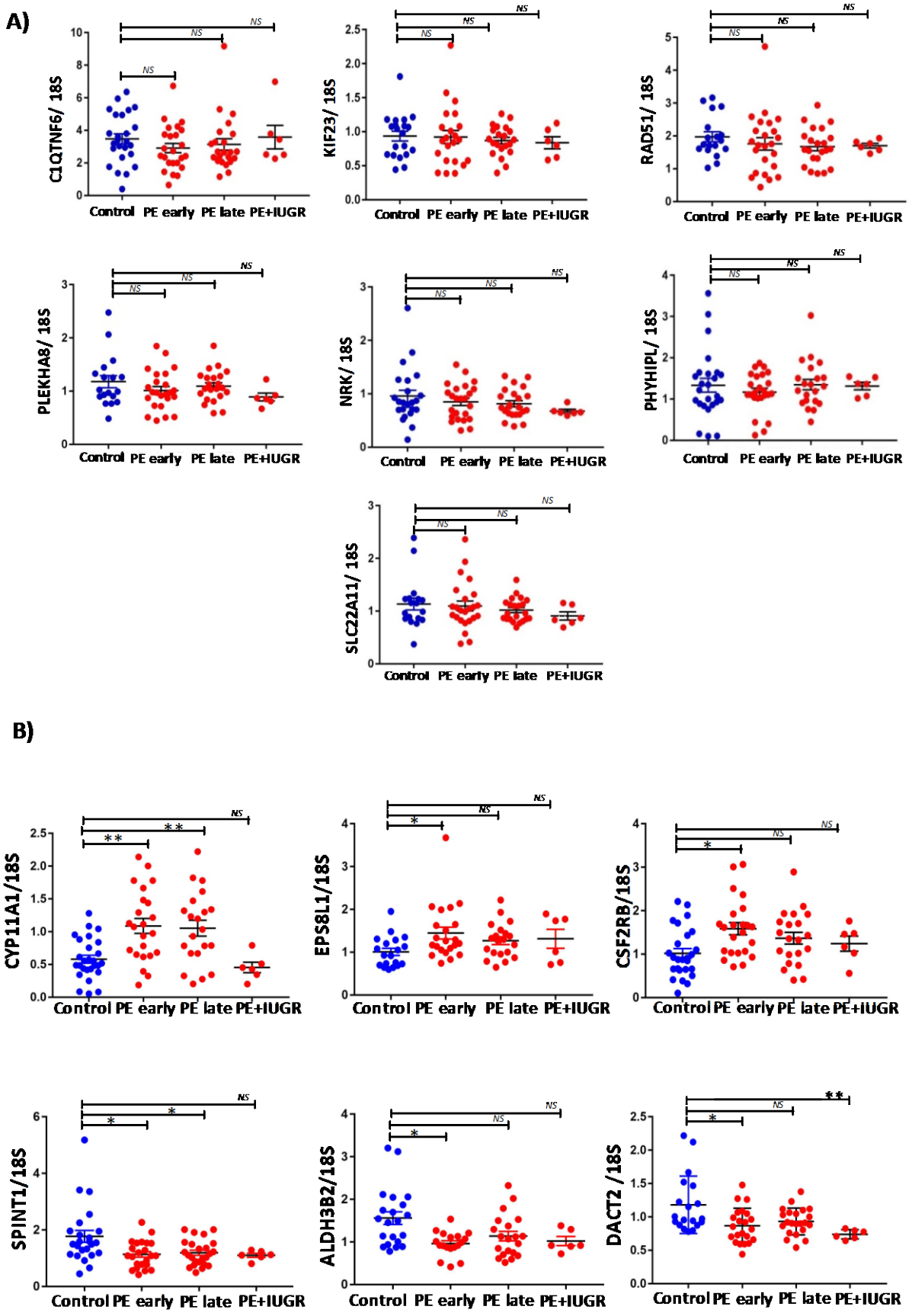

Figure 7. Expression of candidate genes in PE patients

A) Quantitative polymerase chain reaction (q-PCR) did not show any significant differences in the gene expression of *C1QTNF6, KIF23, RAD51*, *PLEKHA8, NRK, PHYHIPL* and *SLC22A11* in PE patients of Oslo-cohort-II as compared to controls (*NS*: Not Significant; values are as mean ± SEM; control n=28; early-onset PE n=24; late-onset PE n=22; IUGR+ PE n=6; Kruskal–Wallis test; Dunn's multiple comparisons test). B) Quantitative polymerase chain reaction (q-PCR) confirmed increased mRNA levels of *CYP11A1* and *CSF2RB* (values are as mean ± SEM; control n= 28; early-onset PE n=24; late-onset PE n=22) (**P< 0.001, *P< 0.01; 1-way ANOVA; Dunnett's multiple comparisons), increased expression of *EPS8L1* (values are as mean ± SEM; control n= 28; early-onset PE n=24; late-onset PE n=22) (*P< 0.01; Kruskal–Wallis test; Dunn's multiple comparisons test) and downregulated expression of *SPINT1, ALDH3B2* and *DACT2* (values are as mean ± SEM; control n= 28; early-onset PE n=24; late-onset PE n=22) (*P< 0.01; Kruskal–Wallis test; Dunn's multiple comparisons test) in early-onset PE patients of Oslo-cohort-II as compared to the control.

3.3. Trophoblast-specific expression of dysregulated genes in PE

The expression of shortlisted dysregulated genes in PE was analyzed in different tissues. The collective average expression of the six candidate genes *(CYP11A1*, *SPINT1*, *CSF2RB*, *ALDH3B2*, *DACT2* and *EPS8L1)* was the highest in the placenta tissue as compared to the rest of the 38 human body tissues (GTEx data; PMID: 25954002) (Figure 8A) indicating their tissue-specific gene regulation.[195] Quantitative-PCR was performed on human pregnancy-associated tissue samples consisting of maternal muscle, fat, decidua, placenta tissues, primary human trophoblasts and Hofbauer cells (predominant macrophage cells of the fetus). The results (normalized with GAPDH) confirmed the enriched expression of the candidate genes in the placenta tissue and primary human trophoblast cells (Figure 8B). Gene transcripts of *CYP11A1, CSF2RB* and *ALDH3B2* were also detected in the decidua but at lower levels (Figure 8B).

In addition, scRNA-seq data analysis (GEO accession: GSE89497; PMID: 30042384) of human placenta tissue from 8th and 24th week of pregnancy further dissected the cell-type specific expression of the 6 candidate genes in human placenta (Figure 8C).[190] The trophoblast sub-types were characterized based on the known marker genes for the specific cell-types as mentioned in Figure 21 in the supplementary information. Expression of all of the six genes were found to be enriched (*FDR< 2e-7) in at least one of the sub-types of the trophoblasts. For example, *EPS8L1* is broadly expressed in the cytotrophoblasts and all the sub-types of extravillous trophoblasts isolated from both first and second trimester placenta but higher

expression levels were found in the syncytiotrophoblast cells. *SPINT1* was also found to be expressed in most of the trophoblast subtypes, but the highest expression was detected in the cytotrophoblasts. Except cytotrophoblasts, *CSF2RB* was found to be expressed in almost all types of trophoblast with higher expression in the extravillous trophoblast cells of the second trimester. *CYP11A1* expression was consistent in all the trophoblast sub-types. *ALDH3B2* also showed stronger expression in syncytiotrophoblasts like *EPS8L1*. *DACT2* was also found to be expressed in all the trophoblast cells (Figure 8C). The single cell analysis further confirmed the trophoblast-specific expression and localization of the candidate genes in the human placenta tissue.

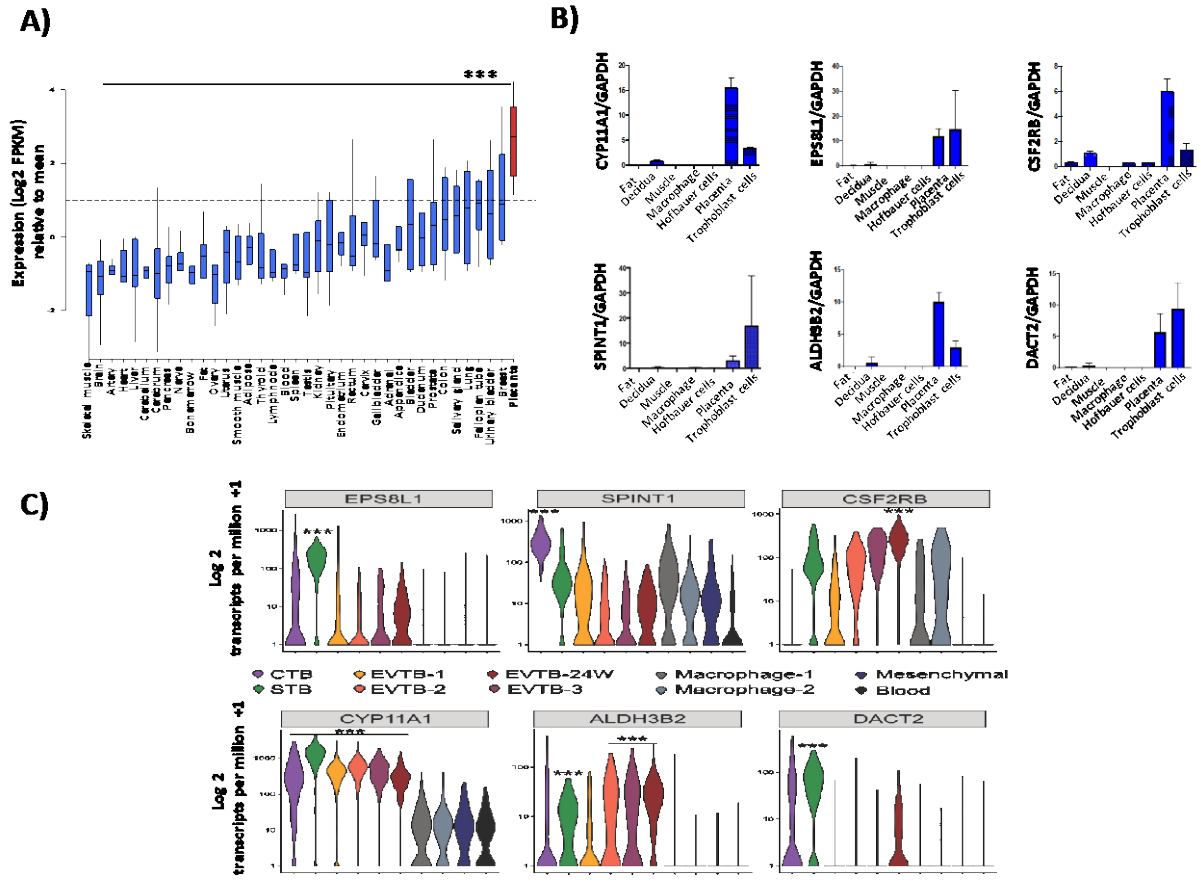

Figure 8. Trophoblast-specific expression of dysregulated genes in PE
A) RNA-seq data analysis (GTEx data; PMID: 25954002) showed enriched (***P< 0.0001) expression of the six candidate genes including *CYP11A1*, *SPINT1*, *CSF2RB*, *ALDH3B2*, *DACT2* and *EPS8L1* in the human placenta as compared to other 38 body tissues. Boxplots shows the distribution of the scaled expression of candidate genes across different human tissues (row-wise z-score) The higher the Z-score, the more tissue-specific enrichment. B) Quantitative polymerase chain reaction (q-PCR) confirmed the higher transcript levels of the *CYP11A1*, *EPS8L1*, *CSF2RB*, *SPINT1*, *ALDH3B2* and *DACT2* in the placenta tissue and

trophoblast cells compared to other pregnancy-related tissues (fat, decidua, muscle, macrophages and hofbauer cells). C) Expression of individual six candidate genes in single cellular transcriptome of human placental tissue (GEO accession: GSE89497; PMID: 30042384). Violin plots showing the enriched expression of *CYP11A1*, *EPS8L1*, *CSF2RB*, *SPINT1*, *ALDH3B2* and *DACT2* across trophoblast lineages of the human placenta from 8th and 24th week of gestation (Log2 transcripts per million +1). (CTB/ VCT: cytotrophoblast, STB/SCT: syncytiotrophoblast, EVTs/ EVTB- (1-3): distinct populations of extravillous trophoblast at 8thW and EVTB-24W: extravillous trophoblast at 24th week), blood, macrophages and Hofbauer cells (macrophage-2)). Given candidate genes are significantly (*FDR < 2e-7) upregulated in either one or all of trophoblast lineages compared with the rest of cell types.

3.4. ERV-LTR mediated trophoblast-specific gene regulation

To delineate the potential *cis*-regulatory effect of the primate-specific ERV-LTRs associated with the candidate genes, the ERV-LTRs were analyzed initially for the transcription factor (TF) binding sites in differentiated trophoblasts (ChIP-seq).[99] The transcription factors (TFs) (e.g. GATA3, GATA2, TFAP2A and TFAP2C) essential for the early trophoblast progenitor cells differentiation *in-vivo* and regulation of the trophoblast-specific gene expression were analyzed (GEO accession: GSE105081; PMID: 29078328).[193] In addition, the active and repressive histone marks (H3K4Me1, H3K27Ac, H3K9Ac and H3K27Me3) over the ERV-LTRs were also analyzed from the data available on human placenta tissue at term (GEO accession:GSE118289; PMID: 30231016) and in differentiated syncytiotrophoblasts (GEO accession: GSE127288; PMID:31294776) (Figure 9A).[191,192] The analysis detected at least one of the four trophectoderm-specific TF binding sites along with the active enhancer marks on the ERV-LTRs, suggesting the possible role of ERV-LTRs as enhancers for their downstream neighbor genes.

To demonstrate the activity of the ERV-LTRs as potential enhancers, an *in-vitro* GFP reporter assay was established. GFP reporter constructs for each of the ERV-LTRs associated with the candidate genes were generated. Multiple upstream human ERV-LTRs of *CYP11A1* and *DACT2* were amplified and cloned individually in a reporter construct. The ERV-LTRs were amplified from the genomic DNA of the human placenta and cloned upstream of the CAGs promoter (Figure 9B). The enhancer assay revealed the highest reporter signal for the MLT1F2-2 located upstream of the *CYP11A1* indicating its enhancer potential (Figure 9C). Notably, the second highest signal was generated by MLT1F2-1 also located upstream of the

CYP11A1, suggesting that the trophoblast-specific *CYP11A1* expression is likely regulated by an ancient ERV-LTR. Interestingly, the MLT1F2 sequence derived from the nearby locus of the *ALDH3B2* generated a significant higher GFP signal whereas a similar sequence but located upstream of the *SPINT1* was not significant (Figure 9C). In addition, MLT1G1, MLT1J, MLT1C and MER41D, located upstream of the *ALDH3B2*, *EPS8L1*, *CYP11A1* and *CSF2RB* genes respectively generated significant GFP signal whereas MLT2B4 neighboring the *DACT2* had no significant enhancement of the GFP signal when compared to the control (Figure 9C). Collectively, the data recapitulate the enhancer potential of the ERV-derived LTRs that have been co-opted in the human genome and are conserved in the primates. Their activation in the human placenta can contribute to the trophoblast-specific gene regulation of certain genes located nearby. Notably, since the human ERV-LTRs are not in their native chromatin context, the reporter assay was not suitable enough to demonstrate the direct link between the *cis*-regulatory activity of the ERV-LTRs and the endogenous expression of the downstream candidate genes.

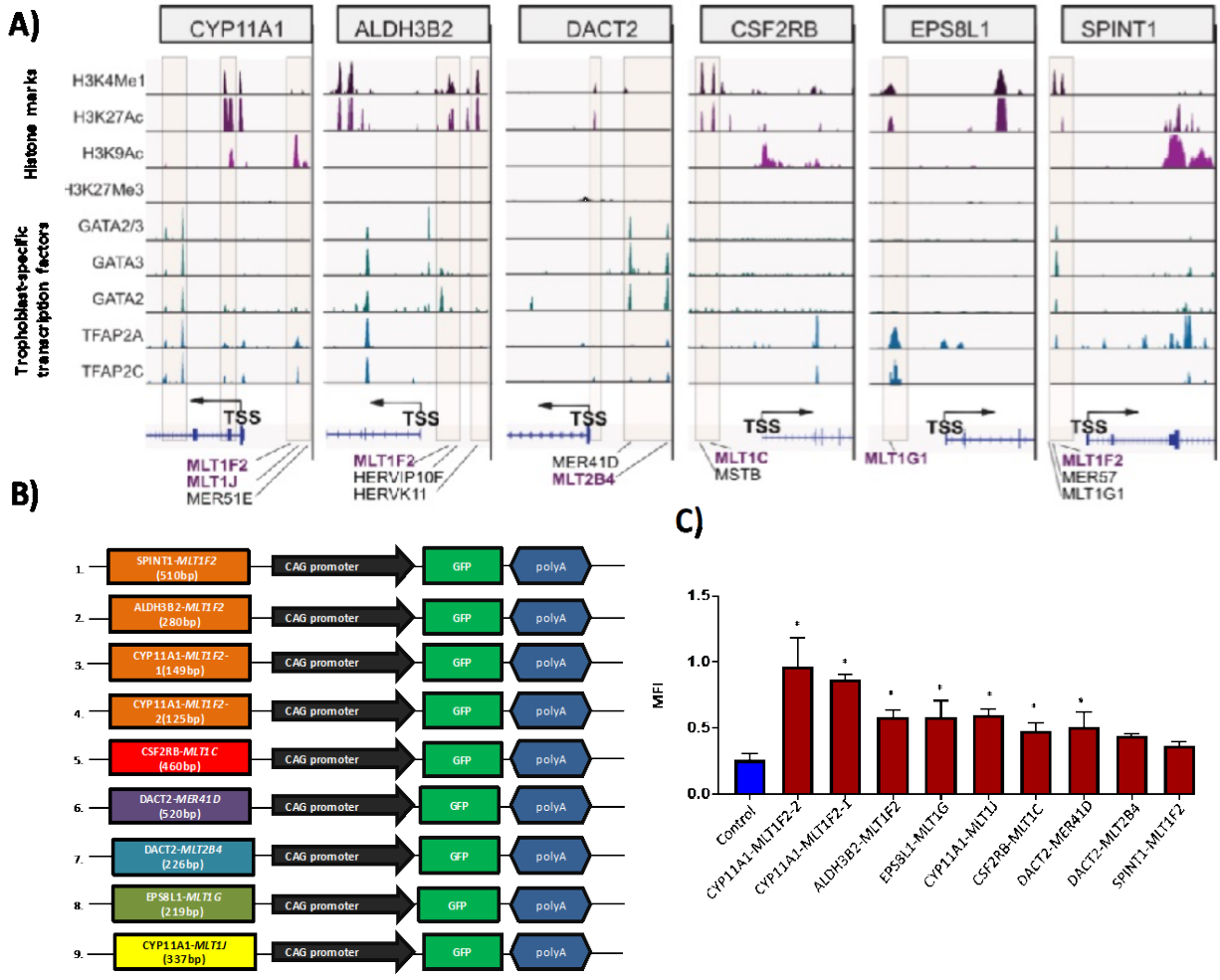

Figure 9. ERV-LTR mediated trophoblast-specific gene regulation

A) Integrative Genome Visualization (IGV) showing raw ChIP-seq signal on the ERV-LTRs associated with *CYP11A1*, *SPINT1*, *CSF2RB*, *ALDH3B2*, *DACT2* and *EPS8L1* for binding sites of active and repressive histone marks (H3K4Me1, H3K27Ac, H3K9Ac and H3K27Me3) in differentiated syncytiotrophoblasts (GEO accession: GSE127288; PMID: 31294776) and human placenta tissue at term (GEO accession: GSE118289; PMID: 30231016). Binding sites of master transcription factors that determines the trophoblast lineages including GATA2, GATA3, TFAP2A and TFAP2C (GEO accession: GSE105081; PMID: 29078328) found in trophoblast progenitors are also shown. Shaded boxes represent the repeat masked ERV-LTR sequences with active chromatin status around the candidate genes. Arrow shows the transcriptional start sites (TSSs) and orientation of respective genes. B) Schematic of the ERV-LTR reporter constructs generated for the GFP-Reporter Assay. C) Bar-plot showing the Median Florescence Intensity (MFI) calculated by normalizing the GFP signal with the m-cherry signal for the 9 ERV-LTR reporter constructs (shown in red) as compared to the control (construct w/o LTR, shown in blue). The data is from three independent experiments (*$P < 0.05$; Unpaired t- test).

3.5. Functional characterization of *EPS8L1* in PE

The candidate genes that were found to be dysregulated in PE patients have been known to mediate different biological processes in humans. Most of the genes have well characterized biological and molecular function, however, their association with the human ERV-LTRs to mediate the trophoblast-specific regulation and their dysregulation in PE patients of Oslo-cohort-II is still novel and interesting to study. Several different biological roles of the candidate genes have been described earlier, for example, *CYP11A1* (Cytochrome P450 Family 11 Subfamily A Member 1) belongs to the family of enzymes that are involved in the steroid hormone production and it was found to be upregulated in PE patients.[196] *ALDH3B2* is involved in the metabolism of alcohol and high levels of *ALDH3B2* has been detected in the human placenta.[197] *SPINT1* (Serine Peptidase Inhibitor, Kunitz Type 1) function as an inhibitor of hepatocyte growth factor (HGF) activator and has been reported to be essential for the development of the mouse placenta.[198] *DACT2* (Dishevelled Binding Antagonist of Beta Catenin 2) is involved in various signaling processes throughout the development (e.g.TGF-beta signaling). *CSF2RB* (Colony Stimulating Factor 2 Receptor Subunit Beta) functions as the common receptor for IL-3, IL-5 and CSF and *EPS8L1* is predicted to be the substrate for the epidermal growth factor receptor pathway, however, the exact function is not known.

In PE data analysis, *EPS8L1* was found to be the most significantly upregulated gene from the 6 short listed genes in PE cohorts after the *CYP11A1,* which has already been associated with PE. *EPS8L1* is an epidermal growth factor receptor pathway substrate 8 (*EPS8*) - like protein 1. Its paralog, *EPS8* is well studied and reported to mediate several pathways that are related to tumorigenesis, metastasis and proliferation of the cancer cells.[199-203] Nevertheless, *EPS8*, by contrast to *EPS8L1* was not found to be dysregulated in PE patients of the Oslo-cohort-II (Figure 10). Since *EPS8L1* is a functionally uncharacterized protein and has not been previously associated with PE, therefore, in this study, I focused on the identification of the role of *EPS8L1* in PE pathophysiology.

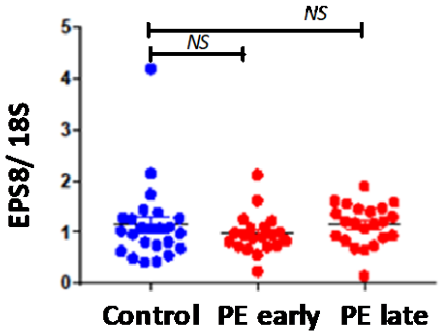

Figure 10. Expression of *EPS8* in PE patients
Quantitative polymerase chain reaction (q-PCR) did not reveal any significant differences in the gene expression of *EPS8* in PE patients (early and late-onset) of the Oslo-cohort-II as compared to the control (values are as mean ± SEM; control n=28; early-onset PE n=24; late-onset PE n=22) (Kruskal–Wallis test; Dunn's multiple comparisons test, *NS*: Not Significant).

3.6. Expression and localization of EPS8L1 in human placenta tissue

The trophoblast-specific expression and the localization of EPS8L1 were additionally analyzed by immunohistochemical (IHC) staining on the human placenta tissue. A healthy human placenta tissue at term and first trimester placental villi were stained. In conjunction with the single-cell transcriptome data, EPS8L1 expression was detected in different types of the trophoblasts, including the extravillous trophoblasts (EVTs), the villous cytotrophoblast (VCTs) and the syncytiotrophoblasts (SCTs) (Figure 11A). No differences in the expression of EPS8L1 were observed between the first trimester villi and the placenta tissue at term, except for SCTs that showed enriched EPS8L1 expression in SCTs of the first trimester villi as compared to the SCTs of the term placenta. To evaluate non-specific staining, the tissue slides were stained with antibody diluent without containing the primary EPS8L1 antibody and was used as a negative control. To monitor the expression of *EPS8L1* during the time course of pregnancy, human placenta samples (n=141) from different gestational-age were analyzed

using q-PCR. The analysis showed that the expression of *EPS8L1* is consistent throughout pregnancy (Figure 11B).

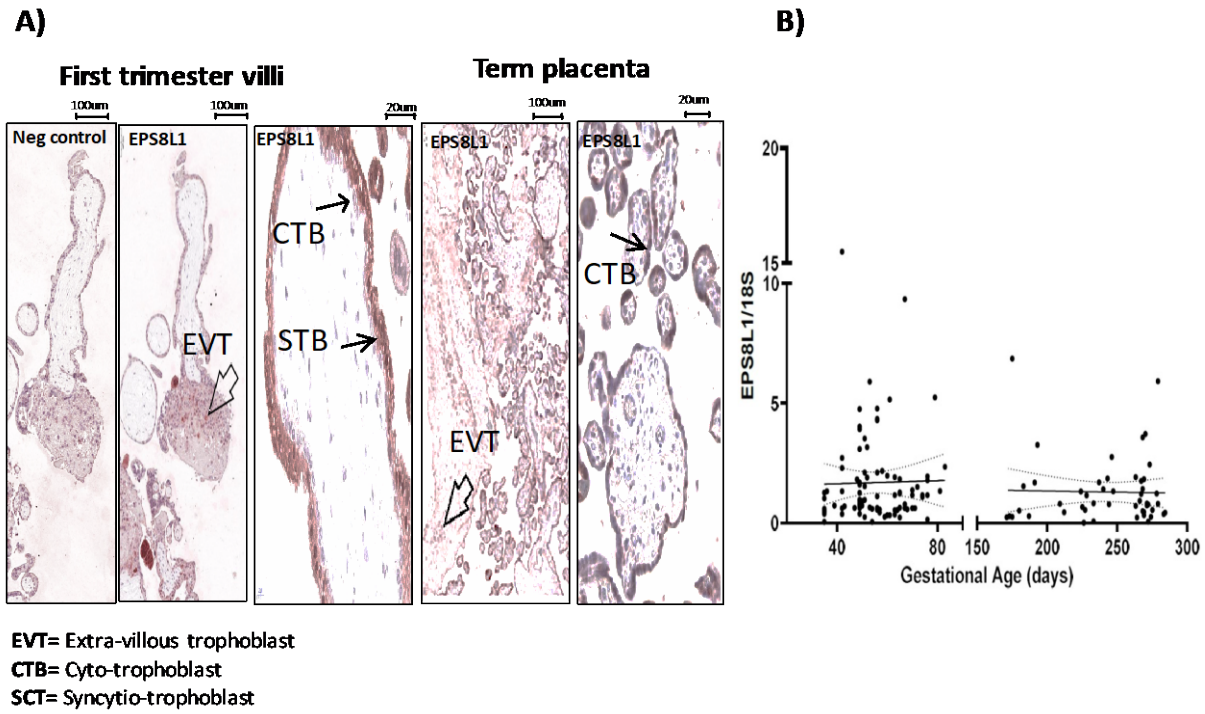

Figure 11. EPS8L1 expression and localization during gestation
A) Immunohistochemical (IHC) staining of placenta villi of first trimester and placenta tissue at term confirmed the trophoblast-specific expression and localization of EPS8L1. The arrows show the EPS8L1 expression in different trophoblasts (Extravillous trophoblast (EVT), villous cytotrophoblast (VCT) and syncytiotrophoblast (SCT)). B) Quantitative polymerase chain reaction (q-PCR) on human placenta tissue (n=141) from early gestation (n=92) and at term (n=49) from women with no self-reported medical history, no smoking habit and age >18 were analyzed for *EPS8L1* mRNA expression. Simple linear regression for expression levels to gestational age were not significant from zero (linear regression line and confidence band shown, alpha was set at 0.05).

3.7. EPS8L1 is upregulated in PE but not in IUGR patients

EPS8L1 was found to be significantly upregulated in early-onset PE patients of the Oslo-cohort-II as compared to the healthy control group. The upregulation was significant both at the transcript level (n=28 controls, n=24 early-onset PE and n=22 late-onset PE, **P*< 0.05; Kruskal–Wallis test; Dunn's multiple comparisons test) and at the protein level in the early-onset PE patients (n=4, **P*< 0.05, ***P*< 0.005; Unpaired t-test) (Figure 12A and 12C). To determine how prevalent *EPS8L1* was dysregulated in PE patients, I analyzed *EPS8L1* levels in other PE cohorts from Kiel and Charite. Quantitative-PCR performed on samples of the

Charite and Kiel cohorts confirmed the increased expression of *EPS8L1* (*P< 0.05; Kruskal–Wallis test; Dunn's multiple comparisons test) in both early-onset (n=14) and late-onset PE patients (n=23) compared to their gestational-age matched healthy controls (Control early (n=32), control late (n=65) (Figure 12B). In contrast, no significant differences were found in the samples derived from IUGR patients of Oslo-cohort-II (n=6), Charite and Kiel cohorts (IUGR early (n=28), IUGR late (n=13) as compared to the controls (Figure 12A and 12B), suggesting that *EPS8L1* dysregulation is specific to PE patients as compared to other pregnancy-related disorders, such as, IUGR.

Notably, EPS8L1 has two protein isoforms (e.g. Isoform A=95 kDa and Isoform B= 66.5 kDa). The shorter isoform EPS8L1_B was predicted to be a secretary protein. Western blot analysis performed on total protein isolated from the placenta tissue of the healthy controls (n=4) and early-onset PE patients (n=4) showed that both isoforms were upregulated in early-onset PE patients (Figure 12C). To confirm that EPS8L1_B is secreted in the human serum, I performed western blot analysis (not shown). To validate the upregulation of EPS8L1_B in PE patients, I used sera collected from 24^{th}-28^{th} week of gestation from controls and early-onset PE patients of Manchester cohort for ELISA. A significant upregulation of EPS8L1_B expression was found in early-onset PE patients (n=12) as compared to the gestational-age matched healthy controls (n=12) (**P< 0.005; Unpaired t-test). As an additional control, sera of non-pregnant women (n=6) were included that had low expression of EPS8L1. Collectively, the results confirmed that the expression of EPS8L1_B is placental (Figure 12D) and suggested that the elevated expression of EPS8L1 in the early-onset PE patients is independent of the changes in the gestational-age (Figure 11). EPS8L1 was found to be upregulated in PE patients from 4 independent cohorts (Oslo-cohort-II, Charite-cohort, Kiel-cohort and Manchester-cohort) highlighting the importance of *EPS8L1* dysregulation in the pathogenesis of PE.

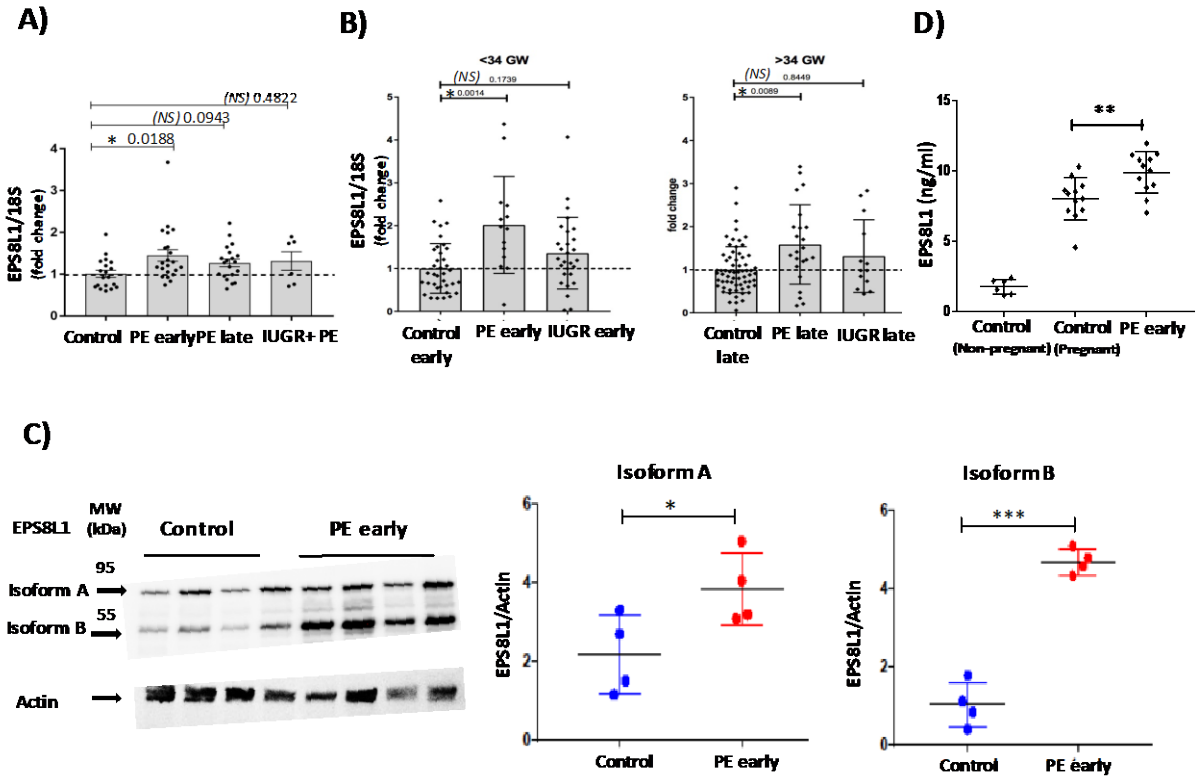

Figure 12. EPS8L1 is upregulated in early-onset PE but not in IUGR patients

A) Quantitative-polymerase chain reaction (q-PCR) confirmed the increased *EPS8L1* mRNA levels in early-onset PE patients (n=24) compared to the healthy controls (n=28) in Oslo-cohort-II (values are as mean ± SEM, **P*< 0.05; Kruskal–Wallis test; Dunn's multiple comparisons test). Late-onset PE (n= 22) also showed increased expression as compared to the healthy controls but not significantly (*P*=0.09; Kruskal–Wallis test; Dunn's multiple comparisons test) B) Quantitative-polymerase chain reaction (q-PCR) performed on samples from Charite and Kiel cohort confirmed upregulation of *EPS8L1* in early-onset PE (n=14) and late-onset PE patients (n=23) compared to gestational-age matched healthy controls (control early (n=32), control late (n=65) (values are as mean ± SEM, **P*< 0.05; Kruskal–Wallis test; Dunn's multiple comparisons test). No significant differences were found in IUGR patients (IUGR early (n=28), IUGR late (n=13) as compared to the healthy controls (Kruskal–Wallis test; Dunn's multiple comparisons test). C) Western blot on placenta tissue of early-onset PE patients confirmed upregulation of EPS8L1 isoform A and B at the protein level (n=4, **P*< 0.05, ****P*< 0.0005; Unpaired t-test). ImageJ software was used to quantify the protein bands of EPS8L1 isoforms from western blot. The quantification data shows the relative amounts as ratio of each isoform band relative to the loading control (Actin). D) ELISA confirmed the upregulation of EPS8L1 in the serum of early-onset PE patients (n=12) as compared to the pregnant controls (n=12) from 24th-28th week of gestation of Manchester cohort (***P*< 0.005;

Unpaired t-test). Serum samples of non-pregnant controls (n=6) were included to confirm that expression of *EPS8L1* is placental.

3.8. EPS8L1 expression correlates with the prognostic marker of PE

The pathogenesis of the multifactorial PE is complex and not fully understood.[13] It is very challenging to detect the disorder early in the pregnancy. Altered circulating levels of anti-angiogenic (sFlt1; soluble FMS-like tyrosine kinase receptor-1) and pro-angiogenic (PlGF; Placental growth factor) factors have been measured in the maternal serum and their ratio sFlt1/ PlGF has been used as a prognostic biomarker of PE prognosis.[204] To find an association between EPS8L1 and an established PE biomarker, the placental expression of EPS8L1 was correlated to the sFlt1/ PlGF ratio determined from the maternal serum of the same PE patients of the Oslo-cohort-II. The correlation analysis revealed that the placental EPS8L1 expression positively correlates and is significant with the sFlt1/ PlGF ratio of the early-onset PE patients (Figure 13).

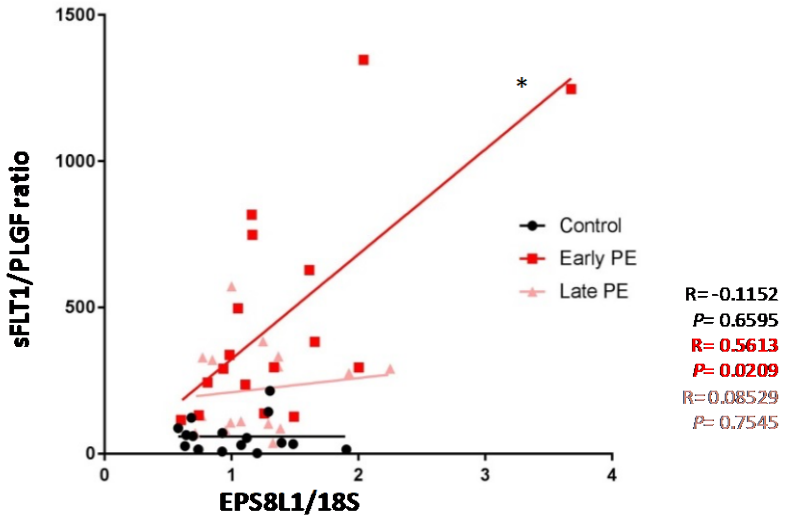

Figure 13. EPS8L1 positive correlation with PE clinical biomarkers
EPS8L1 placental expression significantly positively correlates to the maternal serum sFlt1/ PlGF ratio in early-onset PE patients of the Oslo-Cohort-II as compared to the healthy controls (R= 0.5613; **P*< 0.0209, Spearman Rank Correlation).

3.9. *EPS8L1* knock-out (KO) in trophoblast cells

EPS8L1 expression was detected in different trophoblast sub-types in the human placental tissue. In order to study the function of *EPS8L1* in trophoblast cells, I performed the knock-out

of *EPS8L1* in BeWo cells by using CRISPR-Cas9 genome editing technology. Both isoforms of *EPS8L1* were targeted by the specific sgRNA and three constructs were generated to test their KO efficiency. The constructs containing sgRNA and Cas9 were transfected in the WT-BeWo cells and after 48 hours the GFP positive cells were sorted by FACS. The sorted cells were cultured as a bulk population and were collected for the EPS8L1 protein analysis after 5-7 days. The expression of EPS8L1 was analyzed by western blot in the *EPS8L1*-KO BeWo cells. Loss of *EPS8L1* expression was detected by the activity of all the sgRNAs when compared with the expression in WT-BeWo cells. However, sgRNA2 transfected cells lacked the complete loss of *EPS8L1* expression (Figure 14). The *EPS8L1*-KO cells had a very slow growth and were unable to survive beyond certain time point (5-7 days). Due to this limitation, further analysis on *EPS8L1*-KO cells were unable to perform. The experiment was repeated at least three times with technical replicates but the cells were not able to survive. I also performed the *EPS8L1*-KO in an another trophoblast cell line, SGHPL-4 cells, but the cells couldn't survive after the FACS sorting and were unable to grow as an individual clone. KO studies were not informative, therefore, stably overexpressing *EPS8L1*_SGHPl-4 cells were generated as an *in-vitro* model to mimic the *in-vivo* condition of PE patients. The functional analysis of *EPS8L1* are described later in detail.

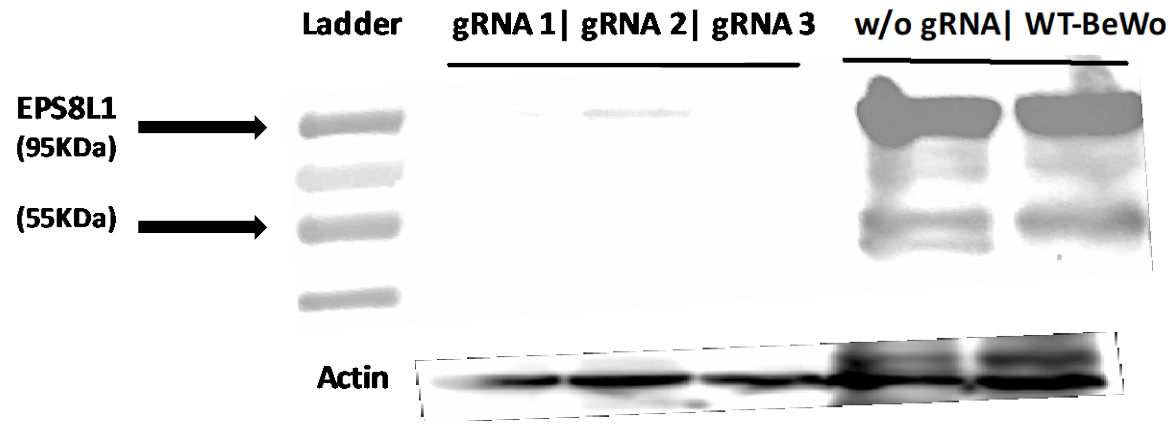

Figure 14. *EPS8L1* knock-out (KO) in trophoblast cells
Western blot confirmed the knock-out of the *EPS8L1* in BeWo cells. Three different sgRNA (sgRNA1, sgRNA2 and sgRNA3) were transfected and EPS8L1 expression was analyzed in each sample as compared to the controls (WT-BeWo cells and cells transfected with construct without sgRNA). Actin was used as a control.

3.10. Characterization of EPS8L1-protein interactors by mass spectrometry (MS)

To understand the potential function of EPS8L1 in human placentation, I performed the mass spectrometry analysis to identify the EPS8L1 protein interactors. Since, EPS8L1 was identified in both syncytiotrophoblasts (SCTs) and extravillous trophoblast cells (EVTs) of the human placental tissue, its interactome was determined in both type of cell lines (BeWo; Syncytio-type and SGHPL-4; EVT-type). *EPS8L1* (HA-tagged) was transiently over-expressed in WT-SGHPL-4 and BeWo cells by transfecting the cells with the expression plasmid (Figure 15A). As an experimental control, a non HA-tagged *EPS8L1* construct was used. After 48 hours of transfection EPS8L1 was pulled down by using the HA-tagged beads. The mass spectrometry data generated almost 400 proteins individually in each cell line as EPS8L1 interactors. A total of 59 proteins came up as significant interactors of EPS8L1 in BeWo cells and 115 proteins came as significant interactors in SGHPL-4 cells. (-Log (p-value > 1.25) (Table 10A and 10B). As expected, the SGHPL-4 cells and BeWo interactomes were different with an overlap of only 17 proteins (bold letters Table 10), suggesting that EPS8L1 has distinct protein interactors in the two trophoblast cell types.

To interpret the interactome data, Gene Ontology Enrichment Analysis was performed by using the ShinyGO v0.61 software to classify the protein interactors in both cell lines, based on their biological function, molecular function and pathways in which EPS8L1 might be involved. In addition, KEGG pathway analysis of protein interactors of both cell lines gave an overview of the enriched pathways based on the interactors. The analysis showed that a lot of interactors are involved in the regulation of translation of proteins in the ribosome, mRNA processing and splicing in the spliceosome complex and processing of proteins in the endoplasmic reticulum. However, pathways like glycolysis, pentose phosphate pathway, cell cell adherent junction that are important for maintaining the cell movement and proliferation and *IL-17* signaling that is involved in both acute and chronic inflammatory responses and activates downstream pathways like NF-kappa B, MAPKs and C/EBPs were found in SGHPL-4 cells specifically (Figure 15C). Antigen processing and presentation, endocytosis and protein export and also a very important signaling pathway for maintenance of pregnancy, Estrogen signaling pathway, was specifically found to be enriched in the interactome data of BeWo cells (Figure 15B). Functionally BeWo and SGHPL-4 cells are different trophoblast cell lines and the KEGG pathway analysis showed that EPS8L1 interacts with different interactors in different cell lines to mediate the corresponding biological process.

To understand EPS8L1 function, I focused on the top most significant EPS8L1 interactors: BAIAP2 and TXNDC5 that were found in both trophoblast cell lines. The functions of these two interactors have been studied extensively and could be important to understand the function of EPS8L1 in human trophoblasts. BAIAP2 (BAR/IMD Domain Containing Adaptor Protein 2 or IRSp53 Insulin Receptor Substrate P53) functions as an adapter protein and regulates the actin cytoskeleton remodeling, filopodia formation and cell migration.[204] Knock-out (KO) studies in mice showed its crucial function in cardiac and placental development during mice embryo development. Besides this the KO placentas showed 27% reduced blood vessels per unit area as compared to the placentas from the healthy mice.[205] This clearly showed that the BAIAP2 has a very important role in the mouse placental angiogenesis. TXNDC5 (Thioredoxin Domain Containing 5) is another significant EPS8L1 interactor and is known to be involved in protein folding, maintaining the cell redox homeostasis, metastasis and invasion of cancer cells by controlling the expression of matrix metalloproteinases.[206] Studies have shown that by repressing TXNDC5 expression, it can lead to inhibition of angiogenesis and metastasis. It is expected that EPS8L1 is also mediating the similar key biological processes of placenta development by interacting with these proteins.

Some of the other important EPS8L1 interactors and their biological functions (what is known in the literature) are summarized in the (Figure 15D). The EPS8L1 MS data gave an insight into the possible pathways in which EPS8L1 might be involved in the human trophoblast cells. The data showed that EPS8L1 has a functional role in angiogenesis, in maintaining cell redox homeostasis, invasion and proliferation of the human trophoblast cells. To further validate the MS data, functional assays were performed which are further explained in detail.

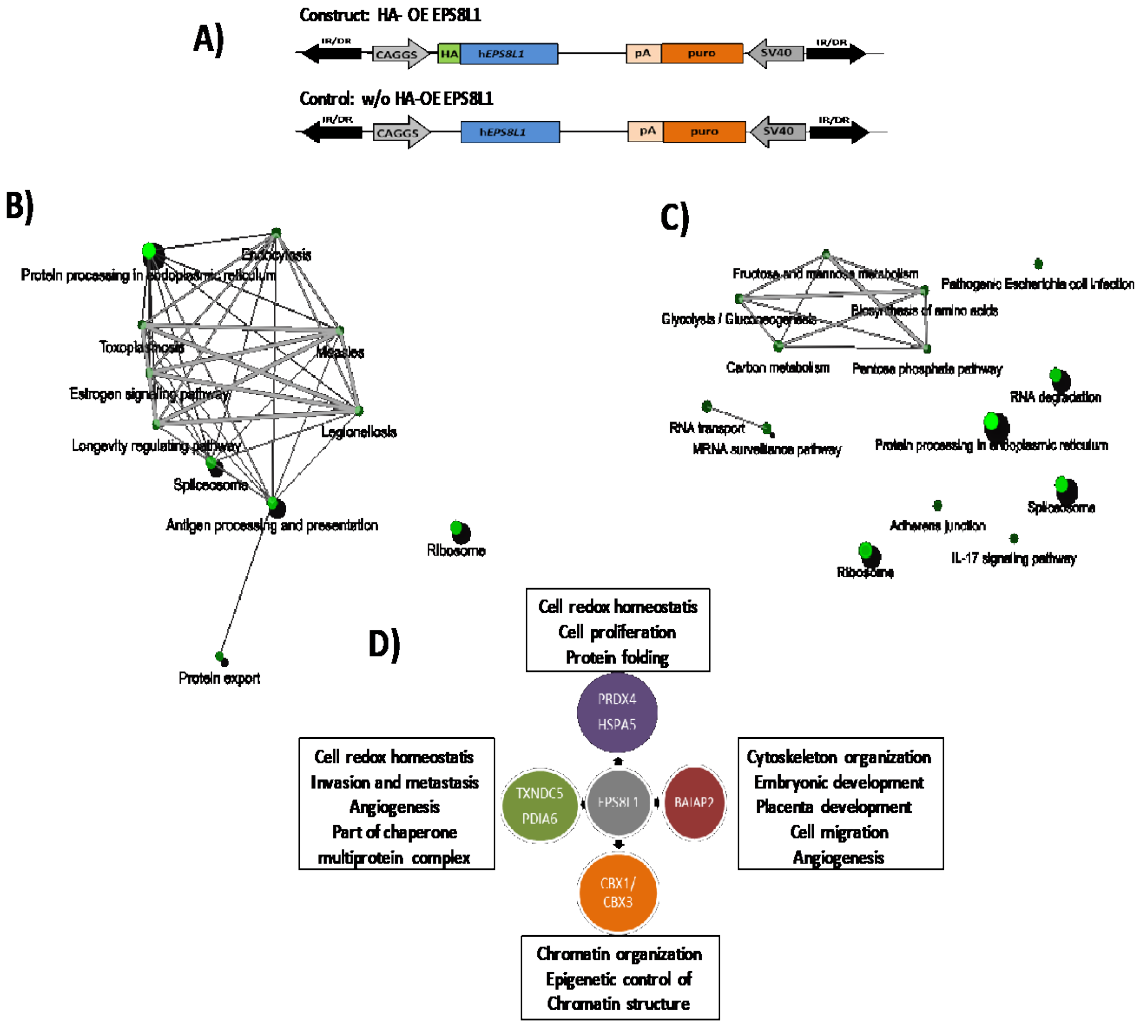

Figure 15. EPS8L1-protein interactors in trophoblast cells
A) *EPS8L1* expression construct (HA and non-HA tagged). B) Interactive-plot showing the enriched pathways among the EPS8L1-protein interactors found in BeWo cells. C) Interactive-plot showing the enriched pathways in EPS8L1-protein interactors found in SGHPL-4 cells (Two pathways (nodes) are connected if they share 20% (default) or more genes. Darker nodes are more significantly enriched gene sets. Bigger nodes represent larger gene sets. Thicker edges represent more overlapped genes). D) Summary of EPS8L1 function based on its interaction with the significant protein interactors.

Table 10. Protein interactors of EPS8L1

A) Significant interactors of EPS8L1 in BeWo cells	
ABI1; ABI2	MYL6
AHNAK	NME1-NME2
ANXA2;ANXA2P2	**P4HB**
ARHGAP27	PABPN1
BAIAP2	PCBP1
BAIAP2L1	**PRDX1**
CALR	RBBP4;RBBP7
CLTA	**RPL15**
DDX21	RPL17;RPL17-C18orf32
EPS8L1	RPL9
FAM98B	RPLP2
FUS	RPS3A
HMGA1	**RPS4X**
HNRNPA0	S100A8
HNRNPA1;HNRNPA1L2	SNRPD1
HNRNPA2B1	SRP72
HNRNPAB	SRSF6
HNRNPD	SSBP1
HNRNPDL	SUB1
HNRNPUL1	TAF15
HSPA1B;HSPA1A	TJP2
HSPA2	**TPM4**
HSPA5	TRMT112
HSPA8	TUFM
HSPA9	**TXNDC5**
HSPH1	UBAP2L
IGF2BP1	**YBX1**
ILF3	YTHDF2
LIN28B	YWHAE
LYAR	

B) Significant interactors of EPS8L1 in SGHPL-4 cells	
ALDOA	**P4HB**
APOOL	PDIA6
ATAD3A	PDLIM7
ATXN2	PFKP
ATXN2L	PKM
AZGP1	**PRDX1**
BAIAP2	PRDX4
BCLAF1	PRKCDBP
C1QBP	PRRC2A
CALML5	PRRC2C
CAPZB	PTGR1
CASP14	RAC1
CBX3;CBX1	RALA
CEP170	RNMT
CKAP4	RPL10
CPOX	RPL13A;RPL13a;RPL13AP3
CSTA	**RPL15**
CTTN	RPL18
DDX6	RPL5
DNAJA1	RPS2
EIF4G1	**RPS4X**
ELAVL1	RPS8
EPS8L1	RPS9
ERP29	RRBP1
EWSR1	RTCB
FKBP11	S100A7
FUBP1	S100A9
FUBP3	SERPINB3
HADHA	SF3B1
HNRNPA2B1	SF3B2
HNRNPD	SKP1
HNRNPH1	SLC25A3
HNRNPH3	SLC25A5
HNRNPK	SMARCE1
HNRNPU	SQRDL
HNRNPUL1	SRRM1
HSPA1B; HSPA1A	SRRT
HSPA5	SRSF7
HSPA8	STAU1
HSPA9	TECR
HSPD1	TIAL1
IGF2BP1	TJP1
ILF3	TOMM70A
IMMT	TRA2B
KHSRP	TRIM21
LARP1	TUBB2A;TUBB2B
LARP4	TUBB3
LMNA	**TXNDC5**
LMO7	U2AF2
LRRFIP1	UBB;RPS27A;UBC;UBA52
LSM14A	UPF1
MKI67	XRCC6
MOGS	XRN2
MTHFD2	**YBX1**
NCL	ZC3HAV1
NQO1	ZNF207

3.11. Functional analysis of *EPS8L1* and its contribution to PE pathogenesis

To study the functional role of *EPS8L1* in PE development, I established an *in-vitro* model, where to mimic the PE phenotype, *EPS8L1* was overexpressed in SGHPL-4 (OE-*EPS8L1*_SGHPL-4) trophoblast cells (derived from first trimester human extravillous trophoblasts). To stably overexpress human *EPS8L1* in SGHPL-4 cells, I used the *Sleeping Beauty (SB)* transposon system (SB100X). The positive clones were selected based on the puromycin selection for more than 12 days. In the selected clones, the increased expression of *EPS8L1* was confirmed at the mRNA and protein levels using q-PCR and Western blotting, respectively (Figure 16C and 16D).

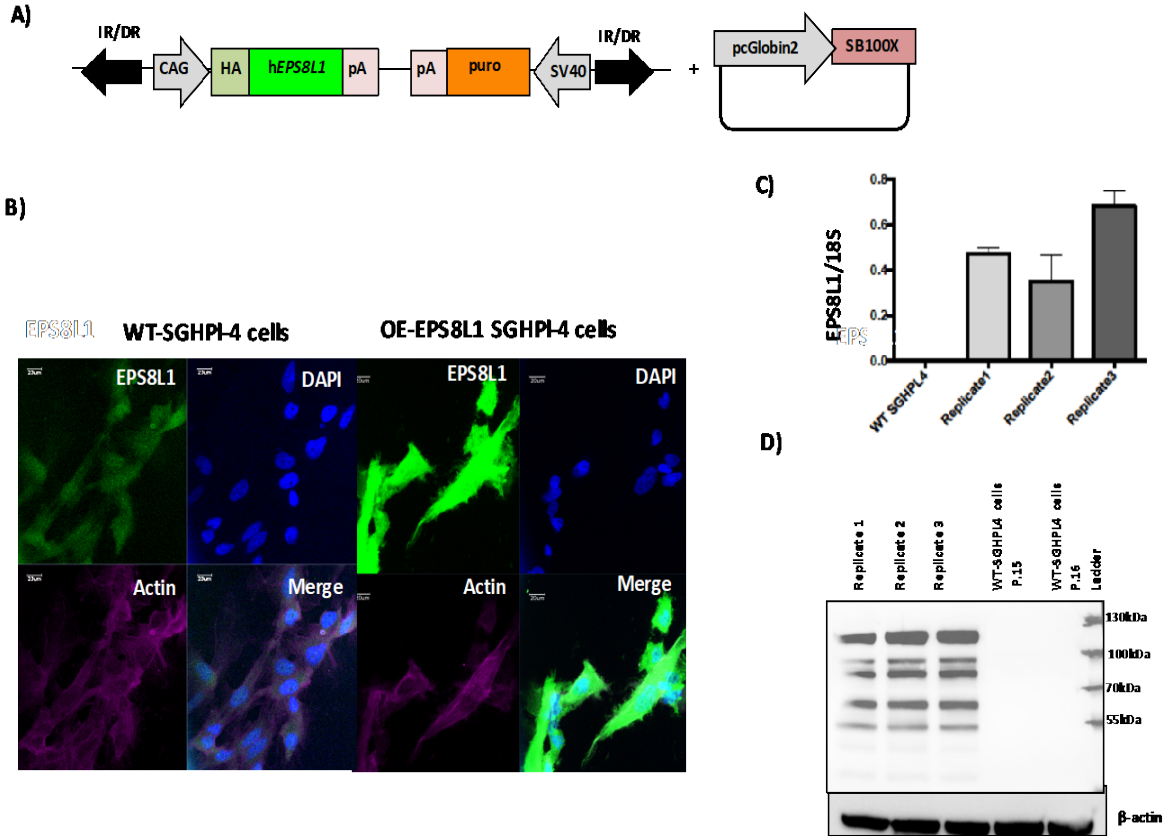

Figure 16. *EPS8L1* overexpression (OE) in trophoblast cells
A) *Sleeping Beauty (SB)* transposon system was used to stably overexpress human *EPS8L1* in SGHPL-4 cells. *EPS8L1* overexpressing cells were selected by puromycin selection for more than 12 days. B) Immunofluorescence staining showing the endogenous expression of *EPS8L1* in WT-SGHPL-4 cells and OE-*EPS8L1*_SGHPL-4 cells. Actin was used as a control. C) Overexpression of *EPS8L1* was confirmed at mRNA level by quantitative polymerase chain reaction (q-PCR) and at D) protein level by western blot (n=3).

3.11.1. Transcriptome analysis of SGHPL-4 cells overexpressing (OE) *EPS8L1*

In order to determine the global transcriptome changes generated by the overexpression of *EPS8L1* in trophoblast cells, RNA-seq was performed on SGHPL-4 cells overexpressing *EPS8L1*. For RNA-seq, total RNA was extracted from both OE-*EPS8L1*_SGHPL-4 cells and WT-SGHPL-4 cells (Control).

Differential gene expression analysis identified 827 significantly (log2 TPM) dysregulated genes in OE-*EPS8L1*_SGHPL-4 cells as compared to the control cells. Among them 615 genes and 210 genes were downregulated and upregulated, respectively. The downregulated genes were involved in biological processes, such as *cell cycle*, *cell invasion*, *cell proliferation* and in *placental angiogenesis* (Figure 17A and 17B). In contrast, genes involved in *oxidative phosphorylation* were upregulated. To see how closely this *in-vitro* model mimics the *EPS8L1* overexpression phenotype in patients, the transcriptome of OE-*EPS8L1*_SGHPL-4 cells was compared with the PE patient transcriptome (PMID: 28904069).[183] This strategy was suitable to identify the global transcriptional changes of the *EPS8L1* overexpression phenotype of PE (Figure 17C). Notably, *EPS8L1* overexpression was significantly represented in the PE cohorts compared with the controls (relatively higher in ~ 60% of the PE patients and ~25% of the controls, hypergeometric $P< 0.03$). Genes that were commonly differentially expressed (n=230) between the two datasets are shown in the scatterplot (Figure 17C). In the *EPS8L1*-overexpressing samples (both *in-vitro* and *in-vivo*), *PFKFB4*, *LDHA*, *CKB*, *PLAC8* and *DIO2* were found to be upregulated and *JUNB*, *FOS*, *C5*, *CX3CL1*, *HES4* and *CLDN1* were downregulated as compared to the control (Figure 17C).

In addition, I used q-PCR on OE-*EPS8L1*_SGHPL-4 cells to confirm the dysregulation of some of the key genes in PE pathogenesis that are involved in the processes like *cell invasion* and *cell proliferation*. Similar to PE patients, *MMP2*, *MMP3* and *MMP9* (*cell invasion*) and *MYC (cell proliferation*) were found to be significantly downregulated in OE-*EPS8L1*_SGHPL-4 cells (Figure 17D). Based on the top dysregulated pathways identified by the transcriptome analysis, I further designed the functional assays using the OE-*EPS8L1*_SGHPL-4 cells as PE model and WT-SGHPL-4 cells as a control.

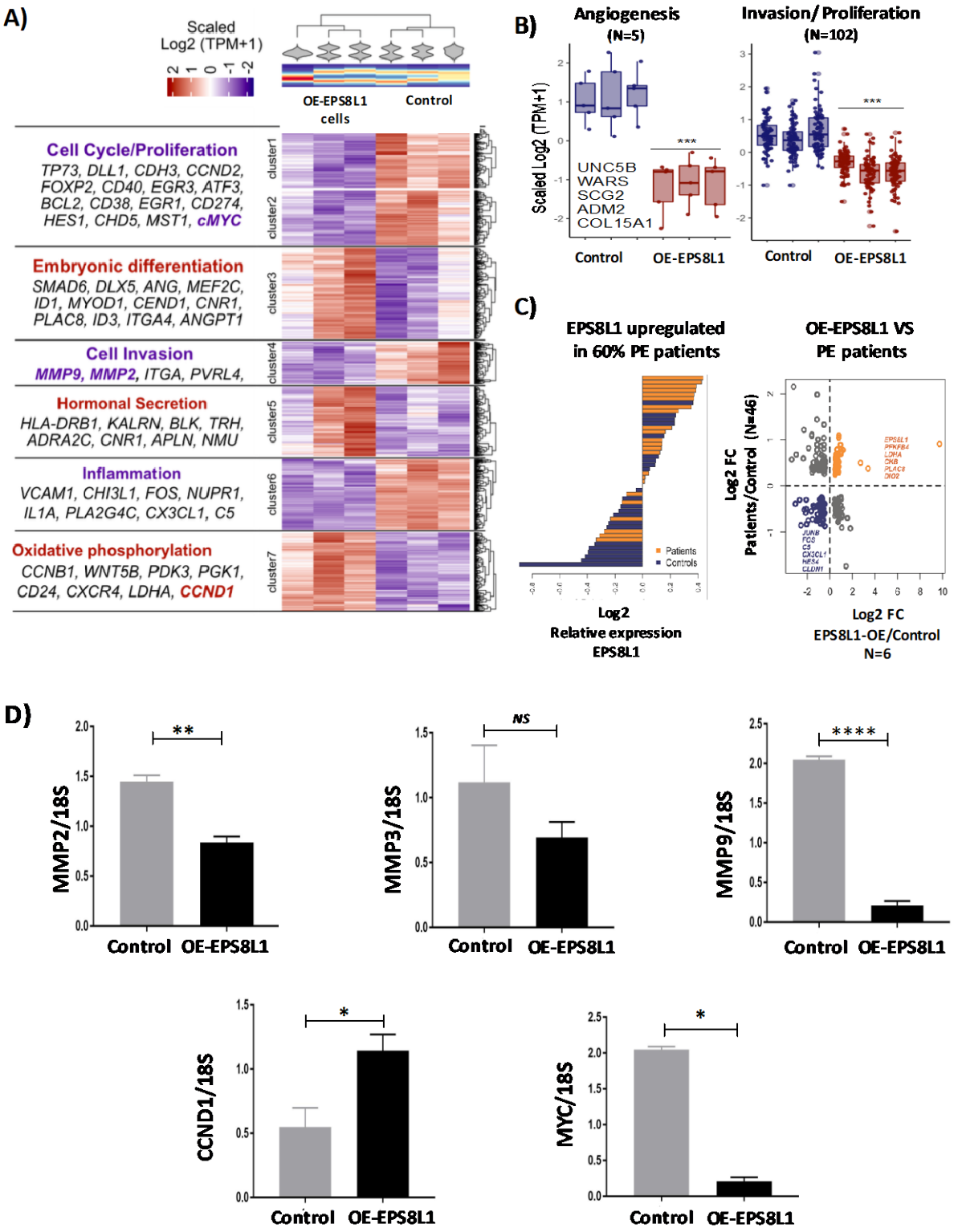
A)
Scaled Log2 (TPM+1)
OE-EPS8L1 cells
Control
Cell Cycle/Proliferation
TP73, DLL1, CDH3, CCND2, FOXP2, CD40, EGR3, ATF3, BCL2, CD38, EGR1, CD274, HES1, CHD5, MST1, cMYC
Embryonic differentiation
SMAD6, DLX5, ANG, MEF2C, ID1, MYOD1, CEND1, CNR1, PLAC8, ID3, ITGA4, ANGPT1
Cell Invasion
MMP9, MMP2, ITGA, PVRL4,
Hormonal Secretion
HLA-DRB1, KALRN, BLK, TRH, ADRA2C, CNR1, APLN, NMU
Inflammation
VCAM1, CHI3L1, FOS, NUPR1, IL1A, PLA2G4C, CX3CL1, C5
Oxidative phosphorylation
CCNB1, WNT5B, PDK3, PGK1, CD24, CXCR4, LDHA, CCND1
cluster1
cluster2
cluster3
cluster4
cluster5
cluster6
cluster7
B)
Angiogenesis (N=5)
Invasion/ Proliferation (N=102)
Scaled Log2 (TPM+1)
UNC5B
WARS
SCG2
ADM2
COL15A1
***
Control
OE-EPS8L1
C)
EPS8L1 upregulated in 60% PE patients
Patients
Controls
Log2 Relative expression EPS8L1
OE-EPS8L1 VS PE patients
Log2 FC Patients/Control (N=46)
Log2 FC EPS8L1-OE/Control N=6
D)
MMP2/18S
**
MMP3/18S
NS
MMP9/18S
****
CCND1/18S
*
MYC/18S
*
Control
OE-EPS8L1

Figure 17. *EPS8L1* overexpression mimics dysregulated pathways in PE
A) Heat map of differentially expressed genes (DEGs, FDR <0.05) in 3 replicates of OE-*EPS8L1*_SGHPL-4 cells vs WT-SGHPL-4 cells (Control) (scaled log2 TPM). DEGs are clustered into 7 clusters defining the gene sets involved in particular gene ontology as annotated on the plot. Genes in clusters 1, 2, 4, and 6 were downregulated, whereas clusters 3, 5 and 7 represent the upregulated genes. B) Jittered box-plot represents the pairwise comparison of scaled expression (at Log2 scale) of genes in OE-*EPS8L1* cells as compared to control cells (***P<0.0005; Wilcoxon rank sum test). Every dot represents a single gene involved in pathways associated with angiogenesis and invasion. C) *EPS8L1* log2 relative expression across healthy (blue) and PE (orange) placenta samples (PMID: 28904069) subjected to the microarray. PE samples displaying higher expression of *EPS8L1* as compared to controls. Comparison between the log2 fold change of the differentially expressed genes in human PE patient's vs healthy controls with the genes differentially expressed on *EPS8L1* overexpression *in-vitro* in SGHPL-4 cells. Hundreds of the genes showed similar differential expression patterns; annotated genes are known to be involved in *invasion* (downregulated) and *oxidative phosphorylation* (upregulated). D) Quantitative polymerase chain reaction (q-PCR) validated the significant downregulation of *MMP2*, *MMP9* and *MYC* (n=4, *P<0.05, **P<0.005, ****P<0.0005; Unpaired t-test) and upregulation of *CCND1* (n=4, *P<0.05; Unpaired t-test) in the OE-*EPS8L1*_SGHPL-4 cells as compared to the control cells. No significant differences were found in the expression of *MMP3* in OE-*EPS8L1*_SGHPL-4 cells as compared to the control cells (n=4, *NS*=Not Significant; Unpaired t-test).

3.11.2. *EPS8L1* overexpression reduces trophoblast cell invasion

The RNA-seq analysis predicted that *EPS8L1* overexpression might affect cell invasion (Figure 17A). To test this hypothesis, transwell invasion assay was designed. The transwell invasion assay was performed using control and OE-*EPS8L1*_SGHPL-4 cells. EGF was used to stimulate the invasion of the control and the OE-*EPS8L1*_SGHPL-4 cells (Figure 18). However, even with the EGF treatment, the number of invaded OE-*EPS8L1*_SGHPL-4 cells were significantly reduced (***P*<0.005, Unpaired t-test) as compared to the EGF treated control cells suggesting that the *EPS8L1* overexpression reduced the invasion of the trophoblasts. FBS (0%) was used as a negative control. However, no significant differences were found in the FBS (0%) treated OE-*EPS8L1*_SGHPL-4 cells as compared to the FBS (0%) treated control cells (n=4, *NS*=Not Significant; Unpaired t-test). The experiment was performed at least three times.

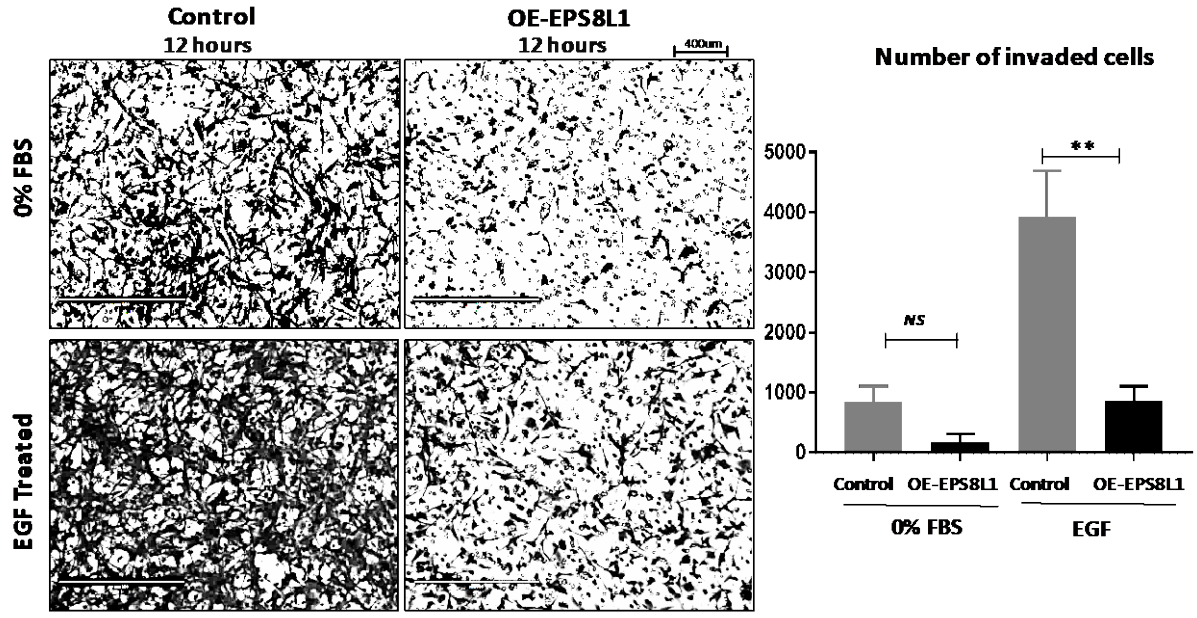

Figure 18. *EPS8L1* overexpression reduces trophoblast cell invasion
Transwell invasion assay was performed on control and OE-*EPS8L1*_SGHPL-4 cells. Images (10X) were taken after 12-14 hours of treatment. Fiji-ImageJ software was used to analyze and calculate the number of invaded cells. A significantly reduced number of OE-*EPS8L1*_SGHPL-4 cells invaded through the matrigel as compared to the control cells with the EGF treatment (n=4, ***P*< 0.005; Unpaired t-test). EGF (10 ng/mL) was applied as a positive control to stimulate the invasion and FBS (0%) was used as a negative control. No significant differences were found in the FBS (0%) treated OE-*EPS8L1*_SGHPL-4 cells as compared to

FBS (0%) treated control cells (n=4, *NS*=Not Significant; Unpaired t-test). The experiment was repeated at least three times and images were taken from different areas of the well.

3.11.3. *EPS8L1* overexpression affects tube formation

Transcriptome analysis of OE-*EPS8L1* cells also showed some genes involved in angiogenesis; a key regulatory process for placentation, to be dysregulated. To find out whether overexpression of *EPS8L1* affects the process of angiogenesis, tube formation assay was performed. Using this *in-vitro* assay, tube characteristics including total number of tubes, tube length, loops and branching points between control and OE-*EPS8L1*_SGHPL-4 cells were analyzed. EGF was used to stimulate the tube formation in the cells. FCS (0%) was used as a negative control. The assay revealed that the OE-*EPS8L1*_SGHPL-4 cells exhibited a significantly reduced number of tubes, loops and branching points as compared to the control cells when treated with EGF (n=6, **P*<0.05, ***P*<0.005; Mann-Whitney test) (Figure 19). The results of the experiment show that the overexpression of *EPS8L1* affects both the invasiveness of the trophoblasts and their ability to form tube-like structures negatively.

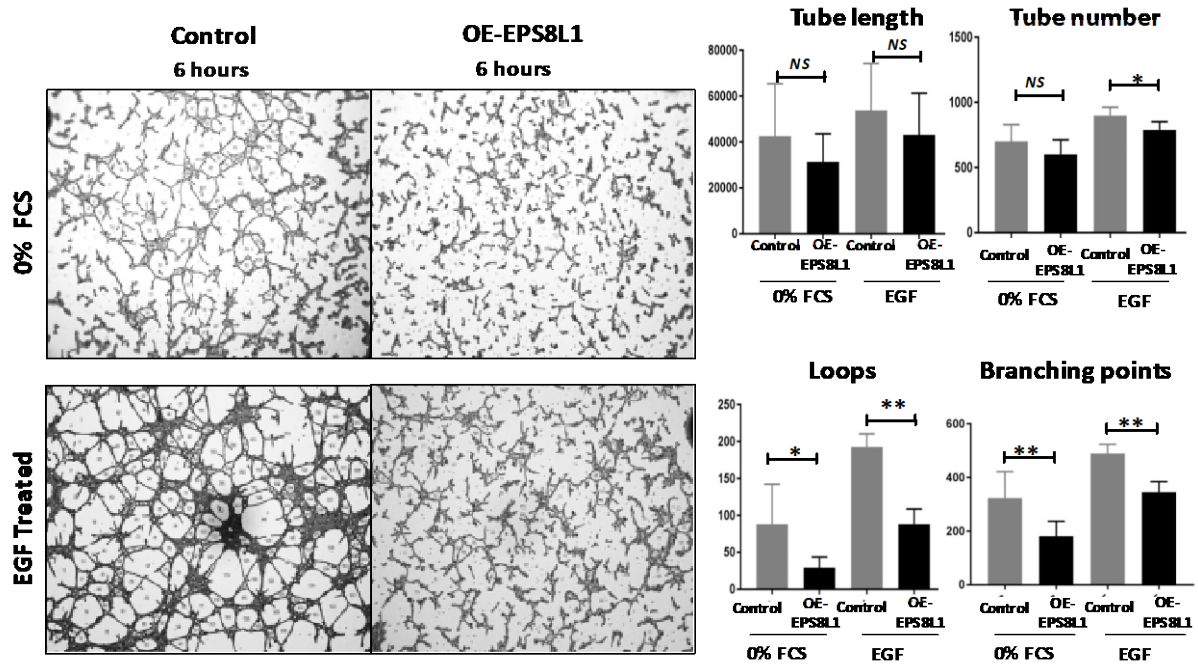

Figure 19. *EPS8L1* overexpression affects tube-formation
Tube formation assay performed on control and OE-*EPS8L1*_SGHPL-4 cells. After 6 hours of treatment, cells were analyzed under microscope and images were taken. Wimasis Image Analysis software was used to analyze and calculate total tube length, number of tubes, number of loops and branching points. In the EGF treated cells, significantly reduced number of tubes, loops formation and branching points were detected in OE-*EPS8L1*_SGHPL-4 cells as compared to the control cells (n=6, **P*<0.05, ***P*<0.005; Mann-Whitney test). No significant

differences were detected in the total tube length of OE-*EPS8L1*_SGHPL-4 cells as compared to the control cells in EGF treated cells (n=6, *NS*: Not Significant; Mann-Whitney test). FCS (0%) was used as a negative control. No significant differences were found in the total number of tubes and length in OE-*EPS8L1*_SGHPL-4 cells as compared to the control cells treated with FCS (0%) (n=6, *NS*=Not Significant; Mann-Whitney test). However, significantly less number of loops and branching points were detected in the FCS (0%) treated OE-*EPS8L1*_SGHPL-4 cells as compared to the control cells (n=6, **P*<0.05, ***P*<0.005; Mann-Whitney test). The data is from six independent experiments.

3.11.4. *EPS8L1* overexpression increases oxidative stress

Transcriptome analysis of OE-*EPS8L1*_SGHPL-4 cells predicted an altered oxidative phosphorylation upon *EPS8L1* overexpression. Oxidative stress has been previously associated with PE placenta.[65] To test that *EPS8L1* overexpression might induce oxidative stress, an *in-vitro* DCFH-DA assay (2', 7'- dichlorodihydrofluorescin diacetate) measuring ROS (reactive oxygen species) level within the cells was designed. Assay was performed on OE-*EPS8L1*_SGHPL-4 cells and was compared to control cells (Figure 20B). In agreement with the prediction, the data suggested that overexpression of *EPS8L1* induces oxidative stress and the percentage increase of ROS was more than 50% (n=3, ***P* < 0.005; Unpaired t-test) (Figure 20C).

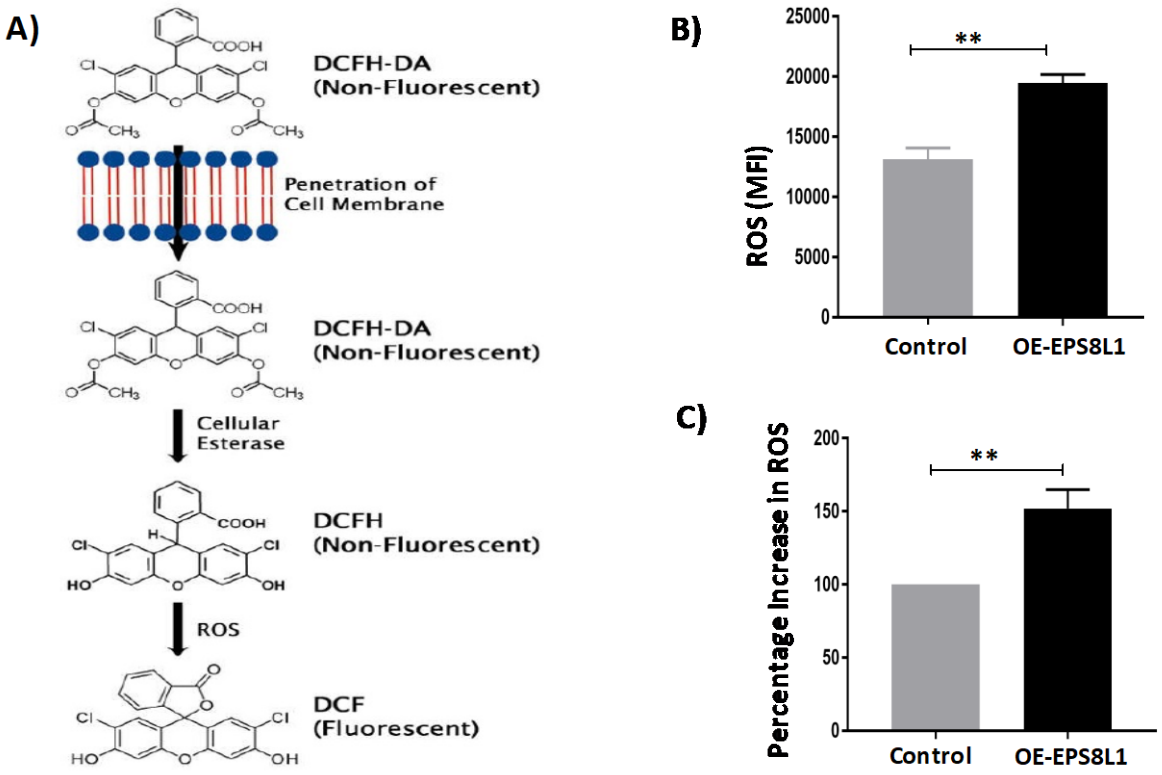

Figure 20. *EPS8L1* overexpression increases oxidative stress
A) Mechanism of DCFH-DA assay. B) ROS production in control and OE-*EPS8L1* SGHPL-4 cells measured by flow cytometry. Data presented as a mean fluorescent intensity (MFI) of the fluorescent signal from the dichlorodihydrofluorescein (DCF) oxidized by ROS. C) ROS production in OE-*EPS8L1*_SGHPL-4 cells detected almost 50% more as compared to the control cells. The data is from three independent experiments (n=3, $^{**}P < 0.005$; Unpaired t-test).

3.12. Conclusion

During the course of my PhD project, I accomplished the following targets:

- Validation of the enriched expression of new set of genes in the human primary trophoblasts in the pregnancy-related tissue panel samples as identified by the *Illumina* Body Map and the GTEx consortium. The expression of the candidate genes was not detected in the maternal decidua tissue (Figure 8).
- Confirmation of the enhancer potential of the primate-specific ERV-LTRs (located within 10-kb region upstream of the TSS) associated with the trophoblast-specific genes (TSGs) by a GFP reporter assay in the trophoblast cell line (Figure 9).
- Identification of a new set of PE associated genes showing enriched expression in trophoblasts mediated by the primate-specific ERV- LTRs (Figure 7).
- Identification of significant upregulation of a previously uncharacterized gene *EPS8L1* in PE patients, enrichment of expression in human trophoblasts and absence of enrichment in the maternal decidua, possibly mediated by a primate-specific ERV-LTR (MLT1G1) located approximately 5-kb upstream of the transcription start site (TSS) of the *EPS8L1* (Figure 9).
- Confirmation of EPS8L1 expression in the Syncytiotrophoblast (SCTs), the villous cytotrophoblasts (VCTs) and the Extravillous trophoblasts (EVTs) of the human placenta tissue by analysis of stained human primary villi and the placenta tissue at term (Figure 11).
- Confirmation that EPS8L1 expression is independent of the gestational-age changes and has a very low expression level in the non-pregnant females (Figure 11 and 12).
- Confirmation of upregulation of *EPS8L1* in the placenta tissue of PE patients at mRNA and protein level in Oslo-cohort-II, at mRNA level in Charite and Kiel cohort and in the serum of early-onset PE patients of high-risk pregnancy cohort of Manchester. *EPS8L1* is found to be upregulated in both early-onset and late-onset PE patients but predominantly it is found to be upregulated in the early-onset PE patients. An important paralog of *EPS8L1* is *EPS8*, however *EPS8* was not observed to be dysregulated in PE patients of Oslo-cohort-II (Figure 10 and 12).
- Validation of EPS8L1 isoform_B secretion and detection in the maternal blood. EPS8L1 was found to be significantly upregulated already in the 24th week of early-onset PE patients of Manchester cohort (Figure 12).

- Positive correlation of EPS8L1 expression in the placenta tissue with the sFLT1/PLGF ratio in the maternal serum of early-onset PE patients, highlighting the importance of potential contribution of EPS8L1 in the prognosis of PE (Figure 13).
- Identification of protein interactions of EPS8L1 in the two trophoblast cell lines (BeWo and SGHPL-4 cells). Most of the interactors observed were specific to the two cell lines. However, a few common interactors such as BAIAP2 and TXNDC5, known to be involved in the placenta development were identified (Figure 15).
- Demonstration that *in-vitro EPS8L1* overexpression model mimics the PE patient phenotype. Reduced trophoblast cell invasion, decreased tube formation (angiogenesis) and increased oxidative stress were detected in overexpressing *EPS8L1* SGHPL-4 cells (Figure 18, 19 and 20). Global transcriptome analysis of overexpressing *EPS8L1* SGHPL-4 cells also led to identification of a number of genes dysregulated in the key pathways of human trophoblasts and the placenta tissue development (Figure 17).

Overall, this study lead to a better understanding of the development of the healthy and PE placenta. Different set of genes were found to be associated with PE. Moreover, I identified the role of human-specific factors, such as human ERV-LTRs in healthy and PE placenta development.

4. Discussion

4.1. PE- A human-specific disorder

PE is a major cause of maternal and fetal death worldwide.[1,2] Approximately 15-20% of all preterm births (PTB) and 12-25% of fetal growth restrictions (FGR) are due to PE.[41,42] Various genetic, immunological, endocrine and environmental factors contribute to PE pathogenesis.[13] The only available treatment of the disorder to date is the removal of the placenta tissue and delivery of the infant.[4] Curiously, PE has not been observed in animals and therefore it is considered to be human-specific.[2] Additional studies are required to understand the molecular mechanisms of the PE pathogenesis with the focus on understanding the question that why is this disorder specific to humans. This study was designed to identify certain human-specific aspects of PE**.**

Compared to other placental mammals, human placenta and trophoblast cells have multiple unique features including morphology, stages of development, mechanism of invasion, embryo implantation and mediating physiological exchange between the fetus and the mother.[207-209] The observed diversity is explained by the rapid evolutionarily changes in the mammalian placentation.[201] Human placenta so far is the least understood of all human organs. Despite its conserved role in supporting the fetal development, it is difficult to model human placentation in other mammals. Defects in the placentation result in pregnancy-related complications such as PE, which is also poorly characterized so far. In this study, I focused on certain human-specific factors of placenta development and their potential contribution to PE pathogenesis.

Human placental evolution had significant assistance from the ancient retroviruses. Human endogenous retroviruses (HERVs) are remains of exogenous retroviral infections of the primate lineage which occurred approximately 25–40 million years ago.[210] The transcription of these elements is epigenetically regulated. Although there are no active ERVs in the human genome, human ERV-derived sequences can be transcriptionally upregulated upon epigenetic changes. Compared to other somatic tissues, the placenta has a unique epigenetic modification (e.g. low DNA methylation) that promotes the expression of human ERVs. Human ERVs have aggressive transcriptional regulatory sequences, the long terminal repeats (LTRs), that have the potential to rewire the expression of the surrounding genes.[152,5,6] In addition,

domesticated ERV-derived genes (e.g envelope gene) have been incorporated into the placentation process and have granted physiological benefit to the host.[211,116,148,147] The functional evolution of hemochorial placentation in primates as well as the invasive phenotype of the human trophoblasts have been associated with ancient retroviral infection.[124,212, 123]

Certain retroviral (ERV) genes have been co-opted to perform key placental functions such as, mediating cell-cell fusions during syncytiotrophoblast differentiation, suppressing maternal immunity or protecting the fetus by exogenous viruses.[147,213,138] The human Syncytin genes, Syncytin 1/HERV-W and Syncytin 2/HERV-FRD, are well-studied trophoblast-specific domesticated genes that are derived from the envelope gene of the ERV and play a very important role in the human placenta development. Abnormal expression of Syncytin-1 has been linked with PE development.[134,147,214,9] Surprisingly, rodents have co-opted retroviral envelope proteins from a distinct retrovirus (ERV) but performing the similar role in placenta morphogenesis and physiology, suggesting the independent co-option of the retroviral genes during evolution for placental functions in rodents and primates.[115] Importantly, the co-option of unrelated retroviruses (ERVs) in rodents and human placenta might have resulted in significant differences (e.g. the invasion process of trophoblasts) justifying the diversified placental evolution of mammals. Such differences make it difficult to model human trophoblast invasion and pregnancy associated disorders in other species.

In addition, due to the activity of the endogenous retroviral LTRs (ERV-LTRs), a limited number of genes gained unique placenta-specific gene expression in humans. The ERV-LTRs are regulatory sequences containing sites for binding of various transcription factors.[99,189,215] Certain ERV-LTRs located nearby to a cellular gene can regulate the expression of the gene by acting as an alternative enhancer or promoter. These events are frequent in the human placenta. In fact, most of the known proteins whose expression is regulated by the ERV-LTRs have an essential role in human placenta development. Notably, their dysregulation have been associated with adverse placental morphology and birth outcome.[216]

There are various examples of placenta-specific genes driven by human ERV-LTRs such as *CYP19* in humans. *CYP19* is expressed in multiple tissues of the human body driven by 10 different promoters.[156] In human placenta, the expression is mediated by a placenta-specific promoter, MER21A, derived from a human ERV-LTR.[157] Rodents lack the placenta-specific isoform however, similar to humans, bovine and ovine *CYP19* have their own placenta-specific promoter. The human *CYP19* is involved in the production of Estrogen and regulates

placental growth and differentiation.[217,218,160,161] Moreover, *INSL4* (insulin-like peptide 4) is the sole member of the insulin-like growth factor gene family that has a specific expression in the human trophoblasts due to the presence of an ERV-derived element in the gene promoter. Furthermore, an ERV-LTR, MER11 is known to control the expression of human Leptin during pregnancy.[131] The absence of MER11 in the murine genome might explain the absence of leptin expression in its placenta. In addition, the trophoblast-restricted expression of the midline1 (*MID1),* human growth factor pleiotrophin (*PTN*) and endothelin B receptor gene (*EDNRB)* is mediated by an ERV-LTR derived from a HERV-E family that act as an alternative placenta-specific promoter/ enhancer containing placenta-specific transcription factor binding sites, such as Sp1 site in the ERV-LTR associated with the *PTN* gene.[164,219,220] Interestingly, these genes exist solely in the humans and Old World monkeys. Among them *CYP19, INSL4*, *PTN*, and Leptin have been reported previously to be dysregulated in PE patients. [176,177,178 179, 180]

Our approach identified already known (Siglec-6*, HCG and CRH*) but also new candidate genes that have unique expression in the human trophoblasts (Table 9).[221,222,223,224,225] With the help of computational biology tools and by performing wet lab experiments (q-PCR) 16 candidate genes were shortlisted. Out of the 16 shortlisted genes, 6 genes (e.g. *ALDH3B2, CYP11A1, SPINT1, DACT2, EPS8L1* and *CSF2RB*) that were significantly dysregulated in the early-onset PE patients of the Oslo-cohort-II were selected for further analysis (Figure 7). With the exception of *CYP11A1*, the rest of the candidate genes were not known to be associated with early-onset PE before.[196] *ALDH3B2, CYP11A1, SPINT1, DACT2, EPS8L1* and *CSF2RB* showed a unique expression in the human trophoblasts and were not detected in the maternal decidua except a very low expression of *CSF2RB* was detected in the human pregnancy-related tissue panel (fat, muscle, decidua, placenta tissue, macrophages, trophoblast and Hofbauer cells) (Figure 8B). Expression analysis of the 6 candidate genes in scRNA-seq data of the human placenta tissue further supported their trophoblast-restricted expression (Figure 8C). The results supported the validity of both the working hypothesis and the strategy that was used to identify the genes with trophoblast-specific expression and their possible contribution in the human placenta development.

The working hypothesis was based on the previous observations that the ERV-LTRs located in the close vicinity upstream (10-kb) of the genes might function as tissue-specific enhancers and contribute to the tissue-specific gene expression. ERV-LTRs that have an enriched expression in the human placenta tissue (as compared to liver, lung, heart, kidney and brain)

were also identified (Figure 6). Furthermore, their putative activity was predicted by analyzing the epigenetic histone modifications and the transcription factor (TF) binding sites using publicly available datasets (ChIP-seq) of the human placenta tissue and differentiated trophoblasts (Figure 9A). Peaks of histone modifications associated with classical enhancers such as enrichment of histone H3 lysine 27 acetylation (H3K27Ac) or histone H3 lysine 4 monomethylation (H3K4Me1) were detected on almost all of the associated ERV-LTRs. In contrast, peaks of repressive (H3K27Me3) chromatin mark were completely absent in the region upstream of candidate genes where ERV-LTRs are located (Figure 9A). These integrative analyses suggested that these genomic regions are transcriptionally active and accessible for the binding of the transcription factors. Binding sites of at least one of the four transcription factors GATA2, GATA3, TFAP2A and TFAP2C involved in trophoblast differentiation from human embryonic stem cells (ESCs) were found on the identified ERV-LTRs (Figure 9A).[193] High levels of H3K4Me1 (active transcriptional enhancers) together with trophoblast-specific transcription factors (TF) binding sites were reliable predictive markers of the potential enhancer activity of the associated ERV-LTRs authenticating the contribution of the retroviral (ERV) regulatory sequences in cell type specific gene regulation. Previously, primate-specific ERV-LTRs such as, MLT1F2 along with MLT1A, MLT1B, MER39B and MER21A have been *(in-silico)* predicted to provide placenta-specific gene expression of their neighboring genes.[226] It was suggested that in species that lack ERV-LTR-derived sequences the associated genes had no placenta-specific expression.[226] Interestingly, our study revealed that MLT1F2 was associated with three different trophoblast-specific candidate genes, identified in this study, including *CYP11A1*, *ALDH3B2* and *SPINT1* (Figure 9). Surprisingly, the *in-vitro* enhancer activity of the three MLT1F2 sequences was different in the reporter assay performed in the cultured trophoblast cells (Figure 9C). MLT1F2 associated with *CYP11A1* and *ALDH3B2* showed clear enhancer activity, whereas the enhancer activity of the MLT1F2, isolated from the upstream region of the *SPINT1* gene, was not significant as compared to the control (Figure 9C). The different trophoblast-specific enhancer activity of the three MLT1F2-derived regulatory regions might be explained by their non-identical sequences, potentially providing different transcription regulatory elements. These observations highlight the complexity of the mechanisms by which human ERV-LTRs have been co-opted to regulate the trophoblast-restricted expression of the nearby genes.

Trophoblast ChIP-seq analysis and *in-vitro* enhancer assay revealed the potential role of ERV-LTRs as trophoblast-specific enhancers of the candidate genes. Notably, the ChIP-seq

signal could not be precisely mapped on all the human ERV-LTRs since the mapping of the reads on the repetitive sequences is rather challenging. Furthermore, in the reporter assay the human ERV-LTRs were not in their native chromatin state. Nevertheless, collectively these two techniques confirmed our hypothesis about ERV-LTRs and their possible role as enhancers. The selected candidate genes associated with these ERV-LTRs were found to be dysregulated in the early-onset PE patients of the Oslo-cohort-II. A unifying feature of the human ERV-LTR-associated dysregulation of genes in PE is likely epigenetics, however the underlying mechanisms are yet to be deciphered. At present, there is no ChIP-seq data available on primary human trophoblasts isolated from PE patients, thus the differential transcription factor (TF) binding and histone modifications on the ERV-LTRs in PE patients correlated to healthy controls could not be analyzed.

4.2. EPS8L1 in PE pathogenesis

Dysregulated trophoblast-specific genes (TSGs) identified in PE patients which include *ALDH3B2* (Aldehyde dehydrogenase), *CYP11A1* (Steroid hormone production), *SPINT1* (Protease inhibitor), *DACT2* (Signaling molecule), *CSF2RB* (Cytokine receptor) and *EPS8L1* (predicted EPS8-like function) are involved in different cellular processes indicating that several pathways might be simultaneously affected in PE. The most dysregulated gene among the 6 candidates was *CYP11A1* which has been previously associated with early-onset PE pathogenesis (Figure 7).[196] Nevertheless, this study identified MLT1F2 as potential trophoblast-specific enhancer of *CYP11A1* providing a mechanistic explanation of its enriched expression in the human trophoblasts.

The human *EPS8L1 (EPS8-like-1)* was the second most dysregulated gene in the early-onset PE patients of the Oslo-cohort-II and currently there are no studies that have been conducted to identify the precise role of *EPS8L1* in the human trophoblasts or its contribution in PE pathogenesis (Figure 12). Therefore, *EPS8L1* was chosen further for an in-depth analysis. *EPS8L1* was identified to be uniquely expressed in the human trophoblasts possibly mediated by MLT1G1; a trophoblast-specific enhancer derived from a primate-specific ERV-LTR. Elevated levels of EPS8L1 were detected in the human placenta tissue and the serum of early-onset PE patients from four independent cohorts (Oslo-cohort-II, Charite cohort, Kiel cohort and Manchester cohort) (Figure 12). Notably, the dysregulation of *EPS8L1* was detected specifically in early-onset PE patients and not in other pregnancy-related disorders such as IUGR (Oslo-cohort-II, Charite and Keil cohorts), suggesting *EPS8L1* to be a relevant

candidate for PE studies. Due to its high sequence homology to the human *EPS8*, it was named as *EPS8-like 1* and was predicted to function in the epidermal growth factor receptor pathway. In addition to *EPS8L1, in-silico* screening of databases of expressed sequence tags (EST) identified further *EPS8-like-1* genes that were named as *EPS8-Like 2* and *3* (*EPS8L2* and *EPS8L3*) respectively.[227,228] The three *EPS8-like* proteins share an overall amino acid sequence similarity between 27 to 42%, suggesting overlapping functional role among the proteins.[229]

EPS8 has been known to modulate various signaling pathways; it function as a cytoplasmic signaling adapter molecule for receptor and non-receptor tyrosine kinases, including EGFR, FGFR, VEGFR and Src kinase.[230] Moreover, *EPS8* regulates several important biological processes such as *cell motility, proliferation, apoptosis, cell cycle* and *cytoskeleton remodeling*.[230] Interestingly, despite its predicted essential cellular function, *Eps8* null mice were completely healthy and fertile.[231] A possible explanation could be the redundancy of the functions of the *EPS8* family members. Indeed, it was observed that *Eps8* KO mice-derived fibroblasts were unable to establish membrane ruffles when stimulated with the growth factor but could be rescued possibly with *Eps8l1*.[231] Eps8l1 could have restored the RTK-dependent cytoskeletal remodeling possibly by establishing a physiological tri-complex with Abi1-and Sos-1 that had activated the Rac dependent pathways in *Eps8* null mice and rescued the phenotype.[229,231]

In this study, particularly *EPS8L1* but not *EPS8* (*EGF* receptor pathway substrate 8) was dysregulated in early-onset PE patients of the Oslo-cohort-II (Figure 10 and 12). *EPS8* is expressed in numerous tissues in the human body but it was not found among the 335 genes that were enriched in the human trophoblasts as compared to the other tissues (liver, lung, brain, kidney, heart and placenta). However, *EPS8* expression has been detected in a broad spectrum of solid tumors including ovarian, lung, breast, thyroid, cervical, pancreatic, colorectal carcinoma, brain, pituitary, oral, esophageal and prostate.[199,200,201,202,203,232,233,234] The dysregulation of *EPS8L1* in PE patients suggest that these two genes *EPS8* and *EPS8L1* might have an independent function especially in the human placenta tissue development.
The expression of *EPS8L1* was detected in both SCTs and EVTs and more than 400 EPS8L1-protein interactors were identified by mass spectrometry in BeWo (SCT-like) and SGHPL-4 (EVT-like) cell lines suggesting the possible involvement of EPS8L1 in several biological processes (Table 10). However, only a few common EPS8L1-protein interactors were

identified in the two types of trophoblast cell lines, suggesting a functional difference of EPS8L1 in the two cell types. The trophoblasts associated potential functions of EPS8L1-protein interactors are summarized in Figure 15D. Among the previously known interactors, mass spectrometry analysis identified only ABI1, BAIAP2 and TXNDC5 in the interactome of EPS8L1 (Table 10). KEGG pathway analysis of EPS8L1-protein interactors in SGHPL-4 cells identified metabolic pathways, including *glycolysis* and *pentose phosphate pathway*, pathways related to *cell migration* and *proliferation* by mediating cell-cell adherent junction and in *inflammatory responses*, both acute and chronic, involving IL-17 signaling and activating downstream pathways like NF-kappa B, MAPKs and C/EBPs signaling pathways (Figure 15C). In contrast, EPS8L1-protein interactors recognized in BeWo cells were involved in the *estrogen signaling pathway* (Figure 15B). Estrogen is a key pregnancy hormone; it prepares the uterus during early pregnancy and is involved in the synthesis of other pregnancy-related hormones that are essential for sustenance of a healthy pregnancy. Besides *estrogen signaling*, the interactors were found to be involved in *antigen processing and presentation, endocytosis* and *protein export* from cytoplasm to the exterior of the cell (Figure 15B).

The shared molecular functions in the two cell types included *cell adhesion molecule binding* (HSPA5, HSPA8, PRDX1, BAIAP2), *enzyme binding* (HSPA9, HNRNPD, P4HB, HSPA5, YBX1, HNRNPUL1, HSPA8) *misfolded/ unfolded protein binding* (HSPA5 HSPA8 HSPA9) and *mRNA binding* (HNRNPA2B1, HNRNPD) *cadherin binding* (HSPA5, HSPA8, PRDX1, RPL15, BAIAP2). BAIAP2/ IRSp53 (BAR/IMD Domain Containing Adaptor Protein 2 or Insulin Receptor Substrate P53) has been found to be crucial for mouse embryonic development. *IRSp53* KO mice embryos were unable to survive due to improper cardiac and placental development. IRSp53 was suggested to be important for the proper differentiation of the spongiotrophoblast layer in mice.[205] IRSp53 has been reported previously as an interactor of Eps8 to establish the IRSp53/Eps8 complex.[235] The IRSp53/Eps8 complex then activates the Rac signaling that modulates the cell motility and invasion in cancer.[235] Based on my MS data, I postulated that the interaction of EPS8L1 with BAIAP2/ IRSp53 might also be involved in regulating the human trophoblast invasion and placenta tissue development (Figure 15D). TXNDC5 (Thioredoxin Domain Containing 5) another EPS8L1-protein interactor was identified in both type of trophoblasts, it has been identified previously to regulate angiogenesis and metastasis by modulating the expression of *SERPINF1* (serpin peptidase inhibitor, clade F) and *TRAF1* (TNF receptor-associated factor 1). Inhibition of

TXNDC5 expression has been reported to significantly reduce the tube-formation (angiogenesis).[236] The interaction of EPS8L1 with TXNDC5, to mediate the process of angiogenesis, might explain the affected tube-like structure formation in OE-*EPS8L1* trophoblast model. PDIA6 (Protein Disulfide Isomerase Family A Member 6) is another significant EPS8L1-protein interactor from the family of TXNDC5, reported to function as a chaperone and prevent the accumulation of the unfolded proteins inside the cells.[237] PDIA6 has been known to specifically interact with integrin β3 subunit and mediate platelet aggregation.[206] Other EPS8L1-protein interactors included CBX1 (Chromobox 1) and CBX3 (Chromobox 3) that are highly conserved non-histone proteins and are known for the epigenetic control of the chromatin structure and the gene expression (Figure 15D). Proteins involved in maintaining the redox balance, PRDX4 (Peroxiredoxin 4) and PRDX1 (Peroxiredoxin 1), were also identified as significant EPS8L1 interactors (Figure 15D). Both of them have been implicated in physiological and pathological mechanisms involved in regulating the cellular response to the oxidative stress.[238,239] These processes ultimately affect the cell survival and proliferation. The EPS8L1-protein interactors showed that EPS8L1 is involved in regulating several different pathways but particularly it has a functional role in angiogenesis, maintaining cell-redox homeostasis, invasion and proliferation of the trophoblast cells in the human placenta tissue.

Interestingly, EPS8L1 was found to have a stable expression in the human placenta and is independent of the gestational-age changes in the pregnancy (Figure 11A and 11B). Analysis of the stained human placenta tissue identified EPS8L1 expression in human trophoblasts of the first trimester villi and the term placenta tissue (Figure 11A). Moreover, *EPS8L1* expression analysis of healthy samples from early gestation (n=92) and at term (n=49) pregnant females indicated that the expression of EPS8L1 is stable during pregnancy (Figure 11B). Furthermore, the expression of EPS8L1 was detectable in almost all the trophoblast sub-types of the human placenta tissue, including, the extravillous trophoblast (EVTs), the syncytiotrophoblasts (SCTs) and the villous cytotrophoblasts (VCTs) as seen in the stained human placenta tissue and scRNA-seq analysis (Figure 11A and 8C). Other tissues and cells such as fat, muscle, macrophage and Hofbauer cells showed very little or almost negligible expression of EPS8L1 and no expression was detected in the maternal decidua, mesenchymal stromal cells and blood cells, suggesting that the expression of EPS8L1 is specific to the trophoblasts as identified from the scRNA-seq analysis of the human placenta tissue and the pregnancy-associated tissue samples (Figure 8 and 11).

Intriguingly, EPS8L1 (Isoform-B) was found to be secreted and detected in the maternal serum during pregnancy. Western blot analysis performed on the maternal serum identified the EPS8L1 secretory isoform B. ELISA performed on the maternal serum from early-onset PE women at around 24^{th}- 28^{th} week of gestation, indicated elevated levels of EPS8L1 as compared to their gestational age-matched healthy controls. Non-pregnant females, by contrast had very low levels of EPS8L1 in their sera, suggesting the importance of EPS8L1 in the placenta tissue development during pregnancy (Figure 12D). As all the controls in Manchester cohort had a previous history of hypertension, EPS8L1 upregulation was specifically detected in the patients who later developed early-onset PE (Figure 12D). Moreover, the placental expression of EPS8L1 in early-onset PE patients of Oslo-cohort II positively correlated with the clinical ratio of the serum sFLT1/PLGF (clinical biomarker of PE) measured in the same PE patients, suggesting the elevated EPS8L1 levels could be used for the early prognosis of the PE pathogenesis (Figure 13).

For the functional studies of *EPS8L1* in the trophoblast cells, it was not possible technically to perform the *EPS8L1*-KO in SGHPL-4 cells, since the *EPS8L1*-KO SGHPL-4 cells were unable to grow as an individual clone. However, *EPS8L1*-KO in BeWo cells was possible and cells could survive as an individual clone but died gradually after 5-7 days of *in-vitro* culturing (Figure 14). As a possible explanation, *EPS8L1*-KO might have disrupted actin cytoskeleton remodeling via disrupting the interaction with the ABI that was identified in MS data as a specific EPS8L1-protein interactor in BeWo cells and has been implicated previously to have a role in the stimulation of the Rac GEF activity to modulate the cytoskeletal remodeling.[229] Besides this, there could be other reasons as well to understand that why the *EPS8L1*-KO BeWo cells were unable to survive beyond a certain time-point. Additional studies are required to understand the possible mechanisms behind the cell death.

As compared to the KO, *EPS8L1* overexpression (OE) studies were more informative. To characterize the functional role of *EPS8L1* in PE development, *EPS8L1* was overexpressed in SGHPL-4 (OE-*EPS8L1*_SGHPL-4) trophoblast cells by using the *Sleeping Beauty (SB)* transposon system (Figure 16).[183] Global transcriptome analysis of overexpressing *EPS8L1* in SGHPL-4 trophoblast cells identified a number of dysregulated genes that are involved in various crucial biological pathways related to placenta development such as *cell invasion, cell proliferation, cell cycle, oxidative phosphorylation* and *angiogenesis* (Figure 17). Following the predictions of the high throughput analysis, the transwell invasion assay could confirm the affected trophoblast invasion in the *EPS8L1* overexpressing trophoblast cells *in-vitro* (Figure

18). Even when stimulated with EGF, the number of invaded *EPS8L1* overexpressing cells were significantly less in comparison to the wild-type (WT) control cells (Figure 18).

Additionally, q-PCR was performed on OE-*EPS8L1*_SGHPL-4 cells to confirm the dysregulation of some key genes that are involved in *cell invasion (e.g. MMP2* and *MMP9*), contributing to PE phenotype. Similar to PE patients, *MMP2* and *MMP9* were downregulated in OE-*EPS8L1*_SGHPL-4 cells (Figure 17D). The *MMP2* and *MMP9* (matrix metalloproteinases) are essential for vascular and uterine arteries remodeling during pregnancy.[240,241,242] Invading EVTs abundantly express MMP2 and MMP9 and invade into the maternal decidua by degradation of ECM (extracellular matrix proteins).[243,244,245,246] Decreased expression of MMP2 and MMP9 has been reported to affect the process of vasodilation and vasoconstriction of the arteries, leading to PE development.[247] My data suggests, that the overexpression of *EPS8L1* affects the expression of MMPs that ultimately has a negative impact on the invasion of OE-*EPS8L1*_SGHPL-4 cells, leading to significantly reduced/ shallow trophoblast cell invasion. Decrease in MMP2 and MMP9 results in collagen deposition which was shown to affect the normal angiogenesis and spiral arteries remodeling during placenta developement.[248] In the tube formation assay, the OE-*EPS8L1*_SGHPL-4 cells failed to establish tube-like structures, resulting in a significantly reduced number of tubes, loops and branching points even after the EGF stimulation (Figure 19). Collectively, the dysregulated *MMPs*, observed in our RNA-seq data could be validated by q-PCR (Figure 17D) and the predicted affected processes could be confirmed by cell invasion and tube formation assays. In addition to *MMPs*, other genes involved in placental angiogenesis (e.g. *UNC5B, SCG2, ADM2, COL15A1* and *WARS*) were found to be significantly reduced in the OE-*EPS8L1*_SGHPl-4 cells (Figure 17B).

For the healthy placenta tissue development, the trophoblasts undergo rapid proliferation and invasion by expressing significantly higher levels of certain proto-oncogenes such as *c-MYC*.[249] Similar to previous PE patient reports, significantly reduced levels of *c-MYC* were detected in the OE-*EPS8L1*_ SGHPl-4 cells (Figure 17D). Therefore, leading to the decreased trophoblast differentiation, proliferation and invasion as indicated from the data (q-PCR and invasion assay).[250,251,252] The expression of certain other genes involved in the regulation of cell proliferation (e.g. *TP73, DLL1, CDH3, CCND2, BCL2, EGR1* and *EGR3*) were also found to be reduced in the *EPS8L1*-OE_SGHPL-4 cells (Figure 17A). Furthermore, the oxidative stress was also found to be significantly elevated in OE-*EPS8L1*_SGHPL-4 cells as demonstrated by DCFH-DA assay (Figure 20). So far, the exact mechanism resulting in the

elevated oxidative stress in PE is not known. However, it has been hypothesized that the abnormal trophoblast invasion, resulting in insufficient placental perfusion and ischemia induces oxidative stress that can contribute in the pathogenesis of PE.[253,254] Reactive oxidative species (ROS) have been reported to regulate the transcription of genes and pathways that might have an effect on the trophoblast proliferation, invasion and angiogenesis.[255] Trophoblast invasion, proliferation, angiogenesis and oxidative stress are linked biological processes that are crucial for the human placenta development during pregnancy. The data suggests that the overexpression of *EPS8L1* might contribute to PE development by affecting all of the above discussed pathways.

To understand how closely OE-*EPS8L1*_SGHPl-4 cells, an *in-vitro* PE-model, mimics the transcriptome of PE patients, the data of OE-*EPS8L1*_SGHPL-4 cells was intersected with the PE patient's microarray data set of 24 PE patients and 22 healthy control samples (Figure 17C). Importantly, in the OE-*EPS8L1*_SGHPl-4 cells and patients, around 230 genes were commonly differentially expressed, suggesting that the *in-vitro EPS8L1* overexpressing model can faithfully capture some of the features of PE and highlight the significance of upregulation of *EPS8L1* in PE pathogenesis. In the overlapping processes with PE patients, genes involved in *oxidative phosphorylation* (*PFKFB4, LDHA, CKB*, *PLAC8* and *DIO2)* and *cell-invasion* (*JUNB, FOS, C5, CX3CL1, HES4* and *CLDN1*) were identified (Figure 17C). Taken together, the data showed that the upregulation of *EPS8L1* in PE has an effect on the pathways related to trophoblast invasion and proliferation, angiogenesis and oxidative stress (Figure 18, 19 and 20). All of the above mentioned pathways have been reported to be dysregulated in PE patients.

4.3. ERV-LTRs and dysregulated trophoblast-specific gene (TSG) expression in PE

Another important question is what causes the dysregulation of the ERV-LTR-driven genes in PE? It has been hypothesized that disturbed epigenetics in PE can affect the ERV-LTRs and thus the expression of the ERV-LTR-regulated genes. Therefore, I investigated the effects of disturbed epigenetics on ERV-LTR-associated trophoblast-specific gene expression in PE patients and in healthy controls. I specifically analyzed the CpG methylation sites of the ERV-LTRs in the healthy controls and PE patients by treating the placental genomic DNA with sodium bisulphite. Bisulphite treatment (unmethylated C to T conversion) made it difficult to amplify ERV-LTR sequences which are highly repetitive sequences. Due to this technical limitation, I could successfully amplify only one ERV-LTR, MLT1F2-2, associated with the *CYP11A1*, from both healthy and early-onset PE patients. Nevertheless, sequencing

results of MLT1F2-2 from healthy (n=3) and early-onset PE patients (n=3) revealed that the CpG site in early-onset PE patients is unmethylated in all of the 3 PE patients tested, whereas it is methylated in the healthy control (not shown). Demethylation of the CpG site in MLT1F2-2 of PE patients could be one of the factors to augment *CYP11A1* expression in the early-onset PE patients. However, more studies are required to support these results and the hypothesis of global epigenetic deregulation, resulting in the ERV-LTR-derived transcriptional dysregulation in PE patients.

Currently, the specificity and novelty of the human trophoblasts has widely been studied, but still there are many unanswered questions. The complexity of the human placentation makes it very challenging and difficult to study. However, recent technologies such as single-cell RNA-seq, organoid cultures and genetic engineering techniques such as CRISPR has made it easier to better understand the normal human placenta development. More appropriate knowledge of the human placenta development during pregnancy will eventually enhance the ability in early diagnosis and treatment of the pregnancy-associated complications such as PE.

5. References

1 WHO Recommendations for Prevention and Treatment of Pre-Eclampsia and Eclampsia. *WHO Guidelines Approved by the Guidelines Review Committee* (2011).

2 Robillard, P. Y., Hulsey, T. C., Dekker, G. A. & Chaouat, G. Preeclampsia and human reproduction. An essay of a long term reflection. *J Reprod Immunol* 59, 93-100, doi:10.1016/s0165-0378(03)00040-8 (2003).

3 Zhou, Y., Damsky, C. H. & Fisher, S. J. Preeclampsia is associated with failure of human cytotrophoblasts to mimic a vascular adhesion phenotype. One cause of defective endovascular invasion in this syndrome? *The Journal of clinical investigation* 99, 2152-2164, doi:10.1172/JCI119388 (1997).

4 Al-Jameil, N., Aziz Khan, F., Fareed Khan, M. & Tabassum, H. A brief overview of preeclampsia. *Journal of clinical medicine research* 6, 1-7, doi:10.4021/jocmr1682w (2014).

5 Reiss, D., Zhang, Y. & Mager, D. L. Widely variable endogenous retroviral methylation levels in human placenta. *Nucleic Acids Res* 35, 4743-4754, doi:10.1093/nar/gkm455 (2007).

6 Macaulay, E. C., Weeks, R. J., Andrews, S. & Morison, I. M. Hypomethylation of functional retrotransposon-derived genes in the human placenta. *Mamm Genome* 22, 722-735, doi:10.1007/s00335-011-9355-1 (2011).

7 Frendo, J. L. *et al.* Direct involvement of HERV-W Env glycoprotein in human trophoblast cell fusion and differentiation. *Mol Cell Biol*. 23, 3566-3574, doi:10.1128/mcb.23.10.3566-3574 (2003).

8 Dupressoir, A., Lavialle C Fau - Heidmann, T. & Heidmann, T. From ancestral infectious retroviruses to bona fide cellular genes: role of the captured syncytins in placentation. *Placenta*. 33, 663-671, doi: 10.1016/j.placenta.2012.05.005 (2012).

9 Schlecht-Louf, G. *et al.* Retroviral infection in vivo requires an immune escape virulence factor encrypted in the envelope protein of oncoretroviruses. *Proc Natl Acad Sci U S A* 107, 3782-3787, doi:10.1073/pnas.0913122107 (2010).

10 Cohen, C. J. *et al.* Placenta-specific expression of the interleukin-2 (IL-2) receptor β subunit from an endogenous retroviral promoter. *Journal of Biological Chemistry* 286, 35543-35552, doi:10.1074/jbc.M111.227637 (2011).

11 Hypertension in pregnancy. Report of the American College of Obstetricians and Gynecologists' Task Force on Hypertension in Pregnancy. *Obstet Gynecol*. 122, 1122-1131, doi: 10.1097/01.AOG.0000437382.03963.88 (2013).

12 Acosta-Sison, H. The relationship of hydatidiform mole to pre-eclampsia and eclampsia; a study of 85 cases. *Am J Obstet Gynecol* 71, 1279-1282, doi:10.1016/0002-9378(56)90437-9 (1956).

13 Chaiworapongsa, T., Chaemsaithong, P., Yeo, L. & Romero, R. Pre-eclampsia part 1: current understanding of its pathophysiology. *Nat Rev Nephrol* 10, 466-480, doi:10.1038/nrneph.2014.102 (2014).

14 SH, A. *Fetal and Neonatal Physiology* 4th Edition, , 46–47 (*Elsevier*, 2011).

15 Soares, M. J., Varberg, K. M. & Iqbal, K. Hemochorial placentation: development, function, and adaptations†. *Biology of Reproduction* 99, 196-211, doi:10.1093/biolre/ioy049 (2018).

16 Yetter, J. F., 3rd. Examination of the placenta. *Am Fam Physician* 57, 1045-1054 (1998).

17 Uhlén, M. *et al.* Proteomics. Tissue-based map of the human proteome. *Science* 347, 1260419, doi:10.1126/science.1260419 (2015).

18 Brosens, I., Pijnenborg R., Vercruysse, L. & Romero, R. The "Great Obstetrical Syndromes" are associated with disorders of deep placentation. *Am J Obstet Gynecol*. 204, 193-201, doi: 10.1016/j.ajog.2010.08.009 (2011).

19 Benirschke, K. The human placenta. J. D. Boyd and W. J. Hamilton. Heffer, Cambridge, 365 pp. 1970. *Teratology* 8, 77-78, doi:10.1002/tera.1420080118 (1973).

20 Schlafke, S. & Enders, A. C. Cellular Basis of Interaction Between Trophoblast and Uterus at Implantation. *Biology of Reproduction* 12, 41-65, doi:10.1095/biolreprod12.1.41 (1975).

21 Kapila, V. & Chaudhry, K. Physiology, Placenta. In: *StatPearls,* Treasure Island (FL). (2020).

22 Turco, M. Y. & Moffett, A. Development of the human placenta. *Development* 146, dev163428, doi:10.1242/dev.163428 (2019).

23 Maltepe, E., Bakardjiev, A. I. & Fisher, S. J. The placenta: transcriptional, epigenetic, and physiological integration during development. *The Journal of clinical investigation* 120, 1016-1025, doi:10.1172/JCI41211 (2010).

24 Pijnenborg, R. & Vercruysse, L. A.A.W. Hubrecht and the naming of the trophoblast. *Placenta* 34, 314-319, doi:10.1016/j.placenta.2013.01.002 (2013).

25 Moffett, A. & Loke, C. Immunology of placentation in eutherian mammals. *Nat Rev Immunol* 6, 584-594, doi:10.1038/nri1897 (2006).

26 Simpson, R. A., Mayhew Tm & Barnes, P. R. From 13 weeks to term, the trophoblast of human placenta grows by the continuous recruitment of new proliferative units: a study of nuclear number using the disector. *Placenta*. 13, 501-512, doi:10.1016/0143-4004(92)90055-x (1992).

27 Handwerger, S. New insights into the regulation of human cytotrophoblast cell differentiation. *Molecular and cellular endocrinology* 323, 94-104, doi:10.1016/j.mce.2009.12.015 (2010).

28 Pijnenborg, R., Dixon, G., Robertson, W. B. & Brosens, I. Trophoblastic invasion of human decidua from 8 to 18 weeks of pregnancy. *Placenta* 1, 3-19, doi:https://doi.org/10.1016/S0143-4004(80)80012-9 (1980).

29 Huppertz, B., Weiss, G. & Moser, G. Trophoblast invasion and oxygenation of the placenta: measurements versus presumptions. *Journal of Reproductive Immunology* 101-102, 74-79, doi:https://doi.org/10.1016/j.jri.2013.04.003 (2014).

30 Kaufmann, P., Black, S. & Huppertz, B. Endovascular Trophoblast Invasion: Implications for the Pathogenesis of Intrauterine Growth Retardation and Preeclampsia. *Biology of Reproduction* 69, 1-7, doi:10.1095/biolreprod.102.014977 (2003).

31 Boyd, J. D. & Hamilton, W. J. *The human placenta*. Vol. 212, doi: https://doi.org/10.1002/tera.1420080118 (1970).

32 Burton, G. J., Jauniaux, E. & Watson, A. L. Maternal arterial connections to the placental intervillous space during the first trimester of human pregnancy: The Boyd Collection revisited. *American Journal of Obstetrics and Gynecology* 181, 718-724, doi:https://doi.org/10.1016/S0002-9378(99)70518-1 (1999).

33 Hustin, J. & Schaaps, J.-P. Echocardiograhic and anatomic studies of the maternotrophoblastic border during the first trimester of pregnancy. *American Journal of Obstetrics and Gynecology* 157, 162-168, doi:https://doi.org/10.1016/S0002-9378(87)80371-X (1987).

34 Brosens, I., Robertson, W. B. & Dixon, H. G. The physiological response of the vessels of the placental bed to normal pregnancy. *The Journal of Pathology and Bacteriology* 93, 569-579, doi:10.1002/path.1700930218 (1967).

35 Moser, G., Weiss, G., Gauster, M., Sundl, M. & Huppertz, B. Evidence from the very beginning: endoglandular trophoblasts penetrate and replace uterine glands in situ and in vitro. *Human reproduction (Oxford, England)* 30, 2747-2757, doi:10.1093/humrep/dev266 (2015).

36 Moser, G. *et al.* Endoglandular trophoblast, an alternative route of trophoblast invasion? Analysis with novel confrontation co-culture models. *Human Reproduction* 25, 1127-1136, doi:10.1093/humrep/deq035 (2010).

37 Pijnenborg, R., Bland, J. M., Robertson, W. B., Dixon, G. & Brosens, I. The pattern of interstitial trophoblasticinvasion of the myometrium in early human pregnancy. *Placenta* 2, 303-315, doi:https://doi.org/10.1016/S0143-4004(81)80027-6 (1981).

38 Villar, J. *et al.* Eclampsia and pre-eclampsia: a health problem for 2000 years. In: Critchly, H., MacLean, A., Poston, L. & Walker, J., Eds., Preeclampsia, RCOG Press, London, 189-207 (2003).

39 Sibai, B. M. Diagnosis and management of gestational hypertension and preeclampsia. *Obstet Gynecol*. 102, 181-192, doi:10.1016/s0029-7844(03)00475-7 (2003).

40 Martin, J. A. *et al.* Births: final data for 2009. *Natl Vital Stat Rep*. 60, 1-70 (2011).

41 Duley, L. The global impact of pre-eclampsia and eclampsia. *Semin Perinatol*. 33, 130-137, doi: https://doi.org/10.1053/j.semperi.2009.02.010 (2009).

42 Ghulmiyyah, L. & Sibai, B. Maternal mortality from preeclampsia/eclampsia. *Semin Perinatol*. 36, 56-59. doi: 10.1053/j.semperi.2011.09.011 (2012).

43 Côté AM, Brown MA, Lam E, et al. Diagnostic accuracy of urinary spot protein: creatinine ratio for proteinuria in hypertensive pregnant women: systematic review. *BMJ*. 336, 1003-1006, doi:10.1136/bmj.39532.543947.BE (2008).

44 Redman, C. W. & Sargent, I. L. Latest advances in understanding preeclampsia. *Science*. 308, 1592-1594, doi:10.1126/science.1111726 (2005).

45 Lowe, S. A. *et al.* Guidelines for the management of hypertensive disorders of pregnancy 2008. *Aust N Z J Obstet Gynaecol*. 49, 242-246. doi:10.1111/j.1479-828X.2009. 01003.x (2009).

46 Mostello, D., Kallogjeri D., Tungsiripat, R. & Leet, T. Recurrence of preeclampsia: effects of gestational age at delivery of the first pregnancy, body mass index, paternity, and interval between births. *Am J Obstet Gynecol*. 199, 55. e1-55.e557, doi:10.1016/j.ajog.2007.11.058 (2008).

47 Fisher, S. J. & Roberts J.M., Defects in placentation and placental perfusion.In:Lindheimer M.D., Roberts J.M. & Cunnningham F., *Chelsey's hypertensive disorders in pregnancy*, 377-394 (1999).

48 Conrad, K. P. & Benyo, D. F. Placental cytokines and the pathogenesis of preeclampsia. *American journal of reproductive immunology* 37, 240-249, doi: 10.1111/j.1600-0897. 1997.tb00222.x (1997).

49 Ahmed, A., Dunk C., Ahmad, S. & Khaliq, A. Regulation of placental vascular endothelial growth factor (VEGF) and placenta growth factor (PIGF) and soluble Flt-1 by oxygen--A review. *Placenta*. 21, 16-24, doi:10.1053/plac.1999.0524 (2000).

50 Zhou, Y. *et al.* Human cytotrophoblasts adopt a vascular phenotype as they differentiate. A strategy for successful endovascular invasion? *J Clin Invest*. 99, 2139-2151. doi:10.1172/JCI119387 (1997).

51 Redman, C. W. & Sargent, I. L. Immunology of pre-eclampsia. *Am J Reprod Immunol*. 63, 534-543, doi:10.1111/j.1600-0897.2010. 00831.x (2010).

52 Moffett, A. & Hiby, S. E. How Does the Maternal Immune System Contribute to the Development of Pre-eclampsia? *Placenta* 28, 51-56, doi:https://doi.org/10.1016/j.placenta.2006.11.008 (2007).

53 Lau, S. Y. *et al.* Tumor necrosis factor-alpha, interleukin-6, and interleukin-10 levels are altered in preeclampsia: a systematic review and meta-analysis. *Am J Reprod Immunol*. 70, 412-427. doi:10.1111/aji.12138 (2013).

54 Yiyenoğlu Ö, B. *et al.* Assessment of oxidative stress markers in recurrent pregnancy loss: a prospective study. *Arch Gynecol Obstet*. 289, 1337-1340, doi:10.1007/s00404-013-3113-4 (2014).

55 Dechend, R., Luft Friedrich C. & Lindheimer Marshall D. *Chesleys hypertensive disorders in pregnancy*. 287-296 (Elsevier, 2009).

56 Wallukat, G. *et al.* Patients with preeclampsia develop agonistic autoantibodies against the angiotensin AT1 receptor. *The Journal of clinical investigation* 103, 945-952, doi:10.1172/JCI4106 (1999).

57 Dechend, R. *et al.* Agonistic Autoantibodies to the AT1 Receptor in a Transgenic Rat Model of Preeclampsia. *Hypertension* 45, 742-746, doi:10.1161/01.HYP.0000154785.50570.63 (2005).

58 Burton, G. J., Watson, A. L., Hempstock, J., Skepper, J. N. & Jauniaux, E. Uterine glands provide histiotrophic nutrition for the human fetus during the first trimester of pregnancy. *J Clin Endocrinol Metab*. 87, 2954-2959, doi:10.1210/jcem.87.6.8563 (2002)

59 Burton, G. J., Hempstock, J. & Jauniaux, E. Nutrition of the human fetus during the first trimester—a review. *Placenta* 22, S70-S77, doi:10.1053/plac.2001.0639 (2001).

60 Jauniaux, E. *et al.* Onset of maternal arterial blood flow and placental oxidative stress: a possible factor in human early pregnancy failure. *Am J Pathol*. 157, 2111-2122, doi:10.1016/S0002-9440(10)64849-3 (2000).

61 Burton, G. J., Yung, H. W., Cindrova-Davies, T. & Charnock-Jones, D. S. Placental endoplasmic reticulum stress and oxidative stress in the pathophysiology of unexplained intrauterine growth restriction and early onset preeclampsia. *Placenta* 30, 43-48, doi: 10.1016/j.placenta.2008.11.003 (2009).

62 Huppertz, B., Kadyrov, M. & Kingdom, J. C. P. Apoptosis and its role in the trophoblast. *Am J Obstet Gynecol*.195, 29-39. doi: 10.1016/j.ajog.2005.07.039 (2006).

63 Redman, C. W. G. & Sargent, I. L. Microparticles and immunomodulation in pregnancy and pre-eclampsia. *Journal of reproductive immunology* 76, 61-67, https://doi.org/10.1016/j.jri.2007.03.008 (2007).

64 Lian, I. A. *et al.* Increased endoplasmic reticulum stress in decidual tissue from pregnancies complicated by fetal growth restriction with and without pre-eclampsia. *Placenta* 32, 823-829, doi: 10.1016/j.placenta.2011.08.005 (2011).

65 Burton, G. J. & Jauniaux, E. Oxidative stress. *Best practice & research. Clinical obstetrics & gynaecology* 25, 287-299, doi:10.1016/j.bpobgyn.2010.10.016 (2011).

66 Cindrova-Davies, T., Spasic-Boskovic, O., Jauniaux, E., Charnock-Jones, D. S. & Burton, G. J. Nuclear factor-κB, p38, and stress-activated protein kinase mitogen-activated protein kinase signaling pathways regulate proinflammatory cytokines and apoptosis in human placental explants in response to oxidative stress: effects of antioxidant vitamins. *The American journal of pathology* 170, 1511-1520, 10.2353/ajpath.2007.061035 (2007).

67 Vaughan, J. E. & Walsh, S. W. Oxidative stress reproduces placental abnormalities of preeclampsia. *Hypertension in pregnancy* 21, 205-223, doi: 10.1081/PRG-120015848 (2002).

68 Ferrara, N. *et al.* Heterozygous embryonic lethality induced by targeted inactivation of the VEGF gene. *Nature* 380, 439-442, doi: https://doi.org/10.1038/380439a0 (1996).

69 Maynard, S. E. *et al.* Excess placental soluble fms-like tyrosine kinase 1 (sFlt1) may contribute to endothelial dysfunction, hypertension, and proteinuria in preeclampsia. *The Journal of clinical investigation* 111, 649-658, doi: 10.1172/JCI17189 (2003).

70 Kendall, R. L. & Thomas, K. A. Inhibition of vascular endothelial cell growth factor activity by an endogenously encoded soluble receptor. *Proceedings of the National Academy of Sciences* 90, 10705-10709, doi:10.1073/pnas.90.22.10705 (1993).

71 Venkatesha, S. *et al.* Soluble endoglin contributes to the pathogenesis of preeclampsia. *Nature medicine* 12, 642-649, doi: 10.1038/nm1429 (2006).

72 McAllister, K. A. *et al.* Endoglin, a TGF-β binding protein of endothelial cells, is the gene for hereditary haemorrhagic telangiectasia type 1. *Nature genetics* 8, 345-351, doi: 10.1038/ng1294-345 (1994).

73 Romero, R. *et al.* A longitudinal study of angiogenic (placental growth factor) and anti-angiogenic (soluble endoglin and soluble vascular endothelial growth factor receptor-1) factors in normal pregnancy and patients destined to develop preeclampsia and deliver a small for gestational age neonate. *The journal of maternal-fetal & neonatal medicine* 21, 9-23, doi: 10.1080/14767050701830480 (2008).

74 Chaiworapongsa, T. *et al.* The maternal plasma soluble vascular endothelial growth factor receptor-1 concentration is elevated in SGA and the magnitude of the increase relates to Doppler abnormalities in the maternal and fetal circulation. *The Journal of Maternal-Fetal & Neonatal Medicine* 21, 25-40, doi: 10.1080/14767050701832833 (2008).

75 Chaiworapongsa, T. *et al.* A subset of patients destined to develop spontaneous preterm labor has an abnormal angiogenic/anti-angiogenic profile in maternal plasma: evidence in support of pathophysiologic heterogeneity of preterm labor derived from a longitudinal study. *The Journal of Maternal-Fetal & Neonatal Medicine* 22, 1122-1139, doi: 10.3109/14767050902994838 (2009).

76 Romero, R. *et al.* An imbalance between angiogenic and anti-angiogenic factors precedes fetal death in a subset of patients: results of a longitudinal study. *The Journal of Maternal-Fetal & Neonatal Medicine* 23, 1384-1399, doi: 10.3109/14767051003681121 (2010).

77 Chaiworapongsa, T. *et al.* Maternal plasma concentrations of angiogenic/antiangiogenic factors in the third trimester of pregnancy to identify the patient at risk for stillbirth at or near term and severe late preeclampsia. *American journal of obstetrics and gynecology* 208, 287-281, doi: 10.1016/j.ajog.2013.01.016 (2013).

78 Relton, C. L. & Davey Smith, G. Epigenetic Epidemiology of Common Complex Disease: Prospects for Prediction, Prevention, and Treatment. *PLOS Medicine* 7, e1000356, doi:10.1371/journal.pmed.1000356 (2010).

79 Yuen, R. K., Peñaherrera, M. S., von Dadelszen, P., McFadden, D. E. & Robinson, W. P. DNA methylation profiling of human placentas reveals promoter hypomethylation of multiple genes in early-onset preeclampsia. *European journal of human genetics : EJHG* 18, 1006-1012, doi:10.1038/ejhg.2010.63 (2010).

80 Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev*. 16, 6-21, doi:10.1101/gad.947102 (2002).

81 Chatterjee, A. & Eccles, M. R. DNA methylation and epigenomics: new technologies and emerging concepts. *Genome Biology* 16, 103, doi:10.1186/s13059-015-0674-5 (2015).

82 Ehrlich, M. *et al.* Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells. *Nucleic acids research* 10, 2709-2721, doi: 10.1093/nar/10.8.2709 (1982).

83 Tsien, F. *et al.* Prolonged culture of normal chorionic villus cells yields ICF syndrome-like chromatin decondensation and rearrangements. *Cytogenetic and genome research* 98, 13-21, doi: 10.1159/000068543 (2002).

84 Novakovic, B. *et al.* DNA methylation-mediated down-regulation of DNA methyltransferase-1 (DNMT1) is coincident with, but not essential for, global hypomethylation in human placenta. *Journal of Biological Chemistry* 285, 9583-9593, doi: 10.1074/jbc.M109.064956 (2010).

85 Schroeder, D. I. *et al.* The human placenta methylome. *Proceedings of the national academy of sciences* 110, 6037-6042, doi: 10.1073/pnas.1215145110 (2013).

86 Maja, V. *et al.* Changes in the placenta and in the rat embryo caused by demethylating agent 5-azacytidine. *The International journal of developmental biology* 43, 843-846 (1999).

87 Chelbi, S. T. *et al.* Expressional and epigenetic alterations of placental serine protease inhibitors: SERPINA3 is a potential marker of preeclampsia. *Hypertension*. 49, 76-83, doi: 10.1161/01.HYP.0000250831. 52876.cb (2007).

88 Sun, Z. *et al.* Integrated analysis of gene expression, CpG island methylation, and gene copy number in breast cancer cells by deep sequencing. *PLOS one* 6, e17490, doi: 10.1371/journal.pone.0017490 (2011).

89 Sundrani, D. P. *et al.* Differential placental methylation and expression of VEGF, FLT-1 and KDR genes in human term and preterm preeclampsia. *Clinical Epigenetics* 5, 6, doi:10.1186/1868-7083-5-6 (2013).

90 Reik, W. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* 447, 425-432, doi: 10.1038/nature05918 (2007).

91 Gama-Sosa, M. A., Wang, R. Y. H., Kuo, K. C., Gehrke, C. W. & Ehrlich, M. The 5-methylcytosine content of highly repeated sequences in human DNA. *Nucleic acids research* 11, 3087-3095, doi:10.1093/nar/11.10.3087 (1983).

92 Schulz, W. A., Steinhoff, C. & Florl, A. R. *DNA Methylation: Development, Genetic Disease and Cancer.* 211-250 (Springer, 2006).

93 Wilson, A. S., Power, B. E. & Molloy, P. L. DNA hypomethylation and human diseases. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer* 1775, 138-162, doi: 10.1016/j.bbcan.2006.08.007 (2007).

94 Fuke, C. *et al.* Age related changes in 5-methylcytosine content in human peripheral leukocytes and placentas: an HPLC-based study. *Annals of human genetics* 68, 196-204, doi: 10.1046/j.1529-8817.2004.00081. x. (2004).

95 Cotton, A. M. *et al.* Inactive X chromosome-specific reduction in placental DNA methylation. *Human molecular genetics* 18, 3544-3552, doi: https://doi.org/10.1093/hmg/ddp299 (2009).

96 Matoušková, M., Blažková, J., Pajer, P., Pavlíček, A. & Hejnar, J. CpG methylation suppresses transcriptional activity of human syncytin-1 in non-placental tissues. *Experimental cell research* 312, 1011-1020, doi: 10.1016/j.yexcr.2005.12.010 (2006).

97 De Koning, A. P. J., Gu, W., Castoe, T. A., Batzer, M. A. & Pollock, D. D. Repetitive Elements May Comprise Over Two-Thirds of the Human Genome. *PLOS Genetics* 7, e1002384, doi:10.1371/journal.pgen.1002384 (2011).

98 Ravindran, S. Barbara McClintock and the discovery of jumping genes. *Proc Natl Acad Sci U S A*. 109, 20198-20199, doi:10.1073/pnas.1219372109 (2012).

99 Wang, J. *et al.* Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. *Genome Res*. 22, 1798-1812, doi:10.1101/gr.139105.112 (2012).

100 Ohno, S. So much "junk" DNA in our genome. *Brookhaven Symp Biol*. 23, 366-370 (1972).

101 Faulkner, G. J. *et al.* The regulated retrotransposon transcriptome of mammalian cells. *Nature genetics* 41, 563-571, doi: https://doi.org/10.1038/ng.368 (2009).

102 Sundaram, V. *et al.* Widespread contribution of transposable elements to the innovation of gene regulatory networks. *Genome research* 24, 1963-1976, doi: 10.1101/gr.168872.113 (2014).

103 Lindblad-Toh, K. *et al.* A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* 478, 476-482, doi: 10.1038/nature10530 (2011).

104 Ayarpadikannan, S., Lee, H.-E., Han, K. & Kim, H.-S. Transposable element-driven transcript diversification and its relevance to genetic disorders. *Gene* 558, 187-194, doi: 10.1016/j.gene.2015.01.039 (2015).

105 Schön, U. *et al.* Human Endogenous Retroviral Long Terminal Repeat Sequences as Cell Type-Specific Promoters in Retroviral Vectors. *Journal of Virology* 83, 12643, doi:10.1128/JVI.00858-09 (2009).

106 Tomita, N. *et al.* Transcription of human endogenous retroviral long terminal repeat (LTR) sequence in a lung cancer cell line. *Biochemical and biophysical research communications* 166, 1-10, doi: https://doi.org/10.1016/0006-291X(90)91904-7 (1990).

107 Patzke, S., Lindeskog, M., Munthe, E. & Aasheim, H.-C. Characterization of a Novel Human Endogenous Retrovirus, HERV-H/F, Expressed in Human Leukemia Cell Lines. *Virology* 303, 164-173, doi:https://doi.org/10.1006/viro.2002.1615 (2002).

108 Wang-Johanning, F. *et al.* Expression of multiple human endogenous retrovirus surface envelope proteins in ovarian cancer. *International Journal of Cancer* 120, 81-90, doi:10.1002/ijc.22256 (2007).

109 Gimenez, J. *et al.* Custom human endogenous retroviruses dedicated microarray identifies self-induced HERV-W family elements reactivated in testicular cancer upon methylation control. *Nucleic Acids Research* 38, 2229-2246, doi:10.1093/nar/gkp1214 (2010).

110 Lamprecht, B. *et al.* Derepression of an endogenous long terminal repeat activates the CSF1R proto-oncogene in human lymphoma. *Nature Medicine* 16, 571-579, doi:10.1038/nm.2129 (2010).

111 Kim, Y.-J., Jung, Y.-D., Kim, T.-O. & Kim, H.-S. Alu-related transcript of TJP2 gene as a marker for colorectal cancer. *Gene* 524, 268-274, doi:https://doi.org/10.1016/j.gene.2013.04.006 (2013).

112 Lock, F. E. *et al.* Distinct isoform of FABP7 revealed by screening for retroelement-activated genes in diffuse large B-cell lymphoma. *Proceedings of the National Academy of Sciences* 111, E3534, doi:10.1073/pnas.1405507111 (2014).

113 Emera, D. *et al.* Convergent evolution of endometrial prolactin expression in primates, mice, and elephants through the independent recruitment of transposable elements. *Mol Biol Evol*. 29, 239-247, doi:10.1093/molbev/msr189 (2012).

114 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* 6, 107, doi:10.1186/1742-4690-6-107 (2009).

115 Dupressoir, A. *et al.* Syncytin-A and syncytin-B, two fusogenic placenta-specific murine envelope genes of retroviral origin conserved in Muridae. *Proc Natl Acad Sci U S A.* 102, 725, doi:10.1073/pnas.0406509102 (2005).

116 Blaise, S., de Parseval N., Bénit, L., & Heidmann, T. Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution. *Proc Natl Acad Sci U S A*. 100, 13013-13018, doi:10.1073/pnas.2132646100 (2003).

117 Cornelis, G. *et al.* Retroviral envelope capture in an ancestrally diverged mammalian clade for placentation in the primitive Afrotherian tenrecs. *Proc Natl Acad Sci U S A.* 111, E4332, doi:10.1073/pnas.1412268111 (2014).
118 Rebollo, R., Romanish Mt. & Mager, D. L. Transposable elements: an abundant and natural source of regulatory sequences for host genes. *Annu Rev Genet.* 46, 21-42, doi: 10.1146/annurev-genet-110711-155621 (2012).
119 Cordaux, R. & Batzer, M. A. The impact of retrotransposons on human genome evolution. *Nat Rev Genet* 10, 691–703, https://doi.org/10.1038/nrg2640 (2009).
120 Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* 409, 860–921, doi: https://doi.org/10.1038/35057062 (2001).
121 Löwer, R., Löwer, J. & Kurth, R. The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences. *Proceedings of the National Academy of Sciences* 93, 5177, doi:10.1073/pnas.93.11.5177 (1996).
122 McCarthy, E. M. & McDonald, J. F. Long terminal repeat retrotransposons of Mus musculus. *Genome biology* 5, R14, doi:10.1186/gb-2004-5-3-r14 (2004).
123 Graham, C. H. & Lala, P. K. Mechanisms of placental invasion of the uterus and their control. Biochem Cell Biol. 70, 867-74. doi: 10.1139/o92-135 (1992).
124 Bièche, I. *et al.* Placenta-specific INSL4 expression is mediated by a human endogenous retrovirus element. *Biology of Reproduction* 68, 1422–1429, doi: https://doi.org/10.1095/biolreprod.102.010322 (2003).
125 Mi, S. *et al.* Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. *Nature* 403, 785-789, doi:10.1038/35001608 (2000).
126 Boyd, M. T., Bax C.M., Bax, B. E., Bloxam, D. L. & Weiss, R. A. The human endogenous retrovirus ERV-3 is upregulated in differentiating placental trophoblast cells. *Virology* 196, 905-909, doi:10.1006/viro.1993.1556 (1993).
127 Venables, P. J., Brookes S.M., Griffiths, D., Weiss, R. A. & Boyd, M. T. Abundance of an endogenous retroviral envelope protein in placental trophoblasts suggests a biological function. *Virology* 211, 589-592, doi:10.1006/viro.1995.1442 (1995).
128 Johansen, T., Holm T. & Bjørklid, E. Members of the RTVL-H family of human endogenous retrovirus-like elements are expressed in placenta. Gene 79, 259-267, doi: https://doi.org/10.1016/0378-1119(89)90208-4 (1989).
129 Blond, J. L. *et al.* Molecular characterization and placental expression of HERV-W, a new human endogenous retrovirus family. *J Virol*. 73, 1175-1185, doi:10.1128/JVI.73.2.1175-1185.1999 (1999).
130 Kjellman, C., Sjögren H.O., Salford, L. G. & Widegren, B. HERV-F (XA34) is a full-length human endogenous retrovirus expressed in placental and fetal tissues. *Gene* 239, 99-107, doi:10.1016/s0378-1119(99)00372-8 (1999).
131 Bi, S., Gavrilova, O., Gong, D. W., Mason, M. M. & Reitman, M. Identification of a placental enhancer for the human leptin gene. *J Biol Chem* 272, 30583-30588, doi:10.1074/jbc.272.48.30583 (1997).
132 Schulte, A. M. *et al.* Human trophoblast and choriocarcinoma expression of the growth factor pleiotrophin attributable to germ-line insertion of an endogenous retrovirus. *Proc Natl Acad Sci U S A* 93, 14759-14764, doi:10.1073/pnas.93.25.14759 (1996).
133 Sjøttem, E., Anderssen, S. & Johansen, T. The promoter activity of long terminal repeats of the HERV-H family of human retrovirus-like elements is critically dependent on Sp1 family proteins interacting with a GC/GT box located immediately to 3' the TATA box. *J Virol*. 70, 188-198, doi:10.1128/JVI.70.1.188-198.1996 (1996).
134 Prudhomme, S., Bonnaud B F. & Mallet, F. Endogenous retroviruses and animal reproduction. Cytogenetic and Genome Research 110, 353-64, doi: 10.1159/000084967 (2005).

135 Chuong, E. B. Retroviruses facilitate the rapid evolution of the mammalian placenta. *BioEssays* 35, 853-861, doi:10.1002/bies.201300059 (2013).
136 Huh, J. W., Ha, H. S., Kim, D. S. & Kim, H. S. Placenta-Restricted Expression of LTR-Derived NOS3. *Placenta* 29, 602-608, doi:https://doi.org/10.1016/j.placenta.2008.04.002 (2008).
137 Chuong, E. B., Rumi, M. A. K., Soares, M. J. & Baker, J. C. Endogenous retroviruses function as species-specific enhancer elements in the placenta. *Nature Genetics* 45, 325-329, doi:10.1038/ng.2553 (2013).
138 Rawn, S. M. & Cross, J. C. The Evolution, Regulation, and Function of Placenta-Specific Genes. *Annual Review of Cell and Developmental Biology* 24, 159-181, doi:10.1146/annurev.cellbio.24.110707.175418 (2008).
139 Dunn-Fletcher, C. E. *et al.* Anthropoid primate-specific retroviral element THE1B controls expression of CRH in placenta and alters gestation length. *PLOS Biology* 16, e2006337, doi:10.1371/journal.pbio.2006337 (2018).
140 Chuong, E. B. The placenta goes viral: Retroviruses control gene expression in pregnancy. *PLOS Biology* 16, e3000028, doi:10.1371/journal.pbio.3000028 (2018).
141 Cross, J. C. *et al.* Genes, Development and Evolution of the Placenta. *Placenta* 24, 123-130, doi:https://doi.org/10.1053/plac.2002.0887 (2003).
142 Kawai, J. *et al.* Functional annotation of a full-length mouse cDNA collection. *Nature* 409, 685-690, doi:10.1038/35055500 (2001).
143 Iwatsuki, K. *et al.* A Novel Secretory Protein Produced by Rat Spongiotrophoblast1. *Biology of Reproduction* 62, 1352-1359, doi:10.1095/biolreprod62.5.1352 (2000).
144 Silva, W. A. *et al.* PLAC1, a trophoblast-specific cell surface protein, is expressed in a range of human tumors and elicits spontaneous antibody responses. *Cancer Immunity Archive* 7, 18 (2007).
145 Fant, M., Barerra-Saldana, H., Dubinsky, W., Poindexter, B. & Bick, R. The PLAC1 protein localizes to membranous compartments in the apical region of the syncytiotrophoblast. *Molecular Reproduction and Development* 74, 922-929, doi:10.1002/mrd.20673 (2007).
146 Massabbal, E. *et al.* PLAC1 expression increases during trophoblast differentiation: Evidence for regulatory interactions with the fibroblast growth factor-7 (FGF-7) axis. *Molecular Reproduction and Development* 71, 299-304, doi:10.1002/mrd.20272 (2005).
147 Rote, N. S., Chakrabarti S Fau - Stetzer, B. P. & Stetzer, B. P. The role of human endogenous retroviruses in trophoblast differentiation and placental development. Placenta 25, 673-83. doi: 10.1016/j.placenta.2004.02.008 (2004).
148 Okahara, G. *et al.* Expression analyses of human endogenous retroviruses (HERVs): tissue-specific and developmental stage-dependent expression of HERVs. *Genomics* 84, 982-990, doi: 10.1016/j.ygeno.2004.09.004 (2004).
149 Ponferrada, V. G., Mauck, B. S. & Wooley, D. P. The envelope glycoprotein of human endogenous retrovirus HERV-W induces cellular resistance to spleen necrosis virus. *Archives of Virology* 148, 659-675, doi:10.1007/s00705-002-0960-x (2003).
150 Taruscio, D. & Mantovani, A. Factors regulating endogenous retroviral sequences in human and mouse. *Cytogenetic and Genome Research* 105, 351-362, doi:10.1159/000078208 (2004).
151 Malassiné, A. *et al.* Expression of HERV-W Env Glycoprotein (syncytin) in the Extravillous Trophoblast of First Trimester Human Placenta. *Placenta* 26, 556-562, doi:https://doi.org/10.1016/j.placenta.2004.09.002 (2005).
152 Kudaka, W., Oda, T., Jinno, Y., Yoshimi, N. & Aoki, Y. Cellular Localization of Placenta-Specific Human Endogenous Retrovirus (HERV) Transcripts and their

Possible Implication in Pregnancy-Induced Hypertension. *Placenta* 29, 282-289, doi:https://doi.org/10.1016/j.placenta.2007.11.009 (2008).

153 Malassiné, A. *et al.* Expression of the fusogenic HERV-FRD Env glycoprotein (syncytin 2) in human placenta is restricted to villous cytotrophoblastic cells. *Placenta* 28, 185-191, doi: 10.1016/j.placenta.2006.03.001 (2007).

154 Medstrand, P., Landry Jr. & Mager, D. L. Long terminal repeats are used as alternative promoters for the endothelin B receptor and apolipoprotein C-I genes in humans. *J Biol Chem*. 276, 1896-1903, doi:10.1074/jbc.M006557200 (2001).

155 Xie, M. *et al.* DNA hypomethylation within specific transposable element families associates with tissue-specific enhancer landscape. *Nature Genetics* 45, 836-841, doi:10.1038/ng.2649 (2013).

156 Bulun, S. E. *et al.* Organization of the human aromatase p450 (CYP19) gene. *Semin Reprod Med* 22, 5-9, doi:10.1055/s-2004-823022 (2004).

157 Van de Lagemaat, L. N., Landry Jr, Mager, D. L. & Medstrand, P. Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trends Genet*. 19, 530-536, doi: 10.1016/j.tig.2003.08.004 (2003).

158 Fürbass, R., Selimyan, R. & Vanselow, J. DNA methylation and chromatin accessibility of the proximal Cyp19 promoter region 1.5/2 correlate with expression levels in sheep placentomes. *Molecular Reproduction and Development* 75, 1-7, doi:10.1002/mrd.20756 (2008).

159 Simpson, E. R. *et al.* Aromatase Cytochrome P450, The Enzyme Responsible for Estrogen Biosynthesis. *Endocrine Reviews* 15, 342-355, doi:10.1210/edrv-15-3-342 (1994).

160 Kamat, A. & Mendelson, C. Identification of the regulatory regions of the human aromatase P450 (CYP19) gene involved in placenta-specific expression. *The Journal of steroid biochemistry and molecular biology* 79, 173-180, doi:10.1016/S0960-0760(01)00156-X (2001).

161 Vanselow, J., Zsolnai A., Fésüs, L., Fürbass, R., & Schwerin, M. Placenta-specific transcripts of the aromatase encoding gene include different untranslated first exons in sheep and cattle. Eur J Biochem. 265, 318-24. doi: 10.1046/j.1432-1327.1999.00734-X (1999).

162 Ball, M. *et al.* Expression of pleiotrophin and its receptors in human placenta suggests roles in trophoblast life cycle and angiogenesis. *Placenta* 30, 649-653, doi: 10.1016/j.placenta.2009.05.001 (2009).

163 Cervar-Zivkovic, M. *et al.* Endothelin-1 stimulates proliferation of first-trimester trophoblasts via the A- and B-type receptor and invasion via the B-type receptor. *J Clin Endocrinol Metab*. 96, 3408-3415, doi:10.1210/jc.2011-0634 (2011).

164 Landry, J. R., Rouhi A., Medstrand, P., & Mager, D. L. The Opitz syndrome gene Mid1 is transcribed from a human endogenous retroviral promoter. *Mol Biol Evol*. 19, 1934-1942. doi: 10.1093/oxfordjournals.molbev.a004017 (2002).

165 Heiss, C. *et al.* Pleiotrophin induces nitric oxide dependent migration of endothelial progenitor cells. *Journal of Cellular Physiology* 215, 366-373, doi:10.1002/jcp.21313 (2008).

166 Landry, J. R. & Mager, D. L. Functional analysis of the endogenous retroviral promoter of the human endothelin B receptor gene. *J Virol* 77, 7459-7466, doi:10.1128/jvi.77.13.7459-7466.2003 (2003).

167 Magariños, M. P., Sánchez-Margalet, V., Kotler, M., Calvo, J. C. & Varone, C. L. Leptin Promotes Cell Proliferation and Survival of Trophoblastic Cells. *Biology of Reproduction* 76, 203-210, doi:10.1095/biolreprod.106.051391 (2007).

168 Fan, Q.-W., Muramatsu, T. & Kadomatsu, K. Distinct expression of midkine and pleiotrophin in the spinal cord and placental tissues during early mouse development. *Development, Growth & Differentiation* 42, 113-119, doi:10.1046/j.1440-169x.2000.00497.x (2000).
169 Henson, M. C. & Castracane, V. D. Leptin in pregnancy: an update. *Biol Reprod*. 74, 218-229, doi:10.1095/biolreprod.105.045120 (2006).
170 Thaete, L. G., Jilling, T., Synowiec, S., Khan, S. & Neerhof, M. G. Expression of Endothelin 1 and Its Receptors in the Hypoxic Pregnant Rat. *Biology of Reproduction* 77, 526-532, doi:10.1095/biolreprod.107.061820 (2007).
171 Chassin, D., Laurent A., Janneau, J. L., Berger, R., & Bellet, D. Cloning of a new member of the insulin gene superfamily (INSL4) expressed in human placenta. *Genomics* 29, 465-470, https://doi.org/10.1006/geno.1995.9980 (1995).
172 Laurent, A. *et al.* Insulin-like 4 (INSL4) gene expression in human embryonic and trophoblastic tissues. *Molecular Reproduction and Development* 51, 123-129, doi:10.1002/(SICI)1098-2795(199810)51:2<123::AID-MRD1>3.0.CO;2-S (1998).
173 Ma, A., Koka, R. & Burkett, P. Diverse functions of IL-2, IL-15, and IL-7 in lymphoid homeostasis. *Annu. Rev. Immunol.* 24, 657-679, doi: 10.1146/annurev.immunol.24.021605.090727 (2006).
174. Lan, R. Y., Selmi, C. & Gershwin, M. E. The regulatory, inflammatory, and T cell programming roles of interleukin-2 (IL-2). *Journal of autoimmunity* 31, 7-12, doi: 10.1016/j.jaut.2008.03.002 (2008).
175 Moffett-King, A. Natural killer cells and pregnancy. *Nature Reviews Immunology* 2, 656-663, doi:10.1038/nri886 (2002).
176 Enquobahrie, D. A. *et al.* Differential placental gene expression in preeclampsia. *American journal of obstetrics and gynecology* 199, 566.e561-566.511, doi:10.1016/j.ajog.2008.04.020 (2008).
177 Liu, S., Wang, F. & Liu, G. Knockdown of pleiotrophin increases the risk of preeclampsia following vitrified-thawed embryo transfer. *Int J Oncol*. 53, 1847-1856. doi:10.3892/ijo.2018.4540 (2018).
178 Reimer, T. *et al.* Microarray analysis of differentially expressed genes in placental tissue of pre-eclampsia: up-regulation of obesity-related genes. *Molecular Human Reproduction* 8, 674-680, doi:10.1093/molehr/8.7.674 (2002).
179 Shin, Y. Y. *et al.* Regulation of steroid hormones in the placenta and serum of women with preeclampsia. *Mol Med Rep*. 17, 2681-2688, doi:10.3892/mmr.2017.8165 (2017).
180 Berkane, N. *et al.* Abnormal steroidogenesis and aromatase activity in preeclampsia. *Placenta* 69, 40-49, doi:https://doi.org/10.1016/j.placenta.2018.07.004 (2018).
181 Herse, F. *et al.* Cytochrome P450 subfamily 2J polypeptide 2 expression and circulating epoxyeicosatrienoic metabolites in preeclampsia. *Circulation* 126, 2990-2999, doi:10.1161/CIRCULATIONAHA.112.127340 (2012).
182 Staff, A. C., Ranheim, T., Khoury, J. & Henriksen, T. Increased contents of phospholipids, cholesterol, and lipid peroxides in decidua basalis in women with preeclampsia. *Am J Obstet Gynecol* 180, 587-592, doi:10.1016/s0002-9378(99)70259-0 (1999).
183 Zadora, J. *et al.* Disturbed Placental Imprinting in Preeclampsia Leads to Altered Expression of DLX5, a Human-Specific Early Trophoblast Marker. *Circulation* 136, 1824-1839, doi:10.1161/CIRCULATIONAHA.117.028110 (2017).
184 Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923-930, doi:10.1093/bioinformatics/btt656 (2014).

185 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550, doi:10.1186/s13059-014-0550-8 (2014).
186 Kim, J. *et al.* Transcriptome landscape of the human placenta. *BMC Genomics* 13, 115, doi:10.1186/1471-2164-13-115 (2012).
187 Yanai, I. *et al.* Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics* 21, 650-659, doi:10.1093/bioinformatics/bti042 (2005).
188 Consortium, E. P. An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57-74, doi:10.1038/nature11247 (2012).
189 Xie, M. *et al.* DNA hypomethylation within specific transposable element families associates with tissue-specific enhancer landscape. *Nat Genet* 45, 836–841, https://doi.org/10.1038/ng.2649 (2013).
190 Liu, Y. *et al.* Single-cell RNA-seq reveals the diversity of trophoblast subtypes and patterns of differentiation in the human placenta. *Cell Res* 28, 819-832, doi:10.1038/s41422-018-0066-y (2018).
191 Kwak, Y.T., Muralimanoharan, S., Gogate, A. A. & Mendelson, C. R. Human Trophoblast Differentiation Is Associated With Profound Gene Regulatory and Epigenetic Changes. *Endocrinology* 160, 2189-2203, doi:10.1210/en.2019-00144 (2019).
192 Dunn-Fletcher, C. E. *et al.* Anthropoid primate–specific retroviral element THE1B controls expression of CRH in placenta and alters gestation length. *PLOS Biology* 16, e2006337, doi:10.1371/journal.pbio.2006337 (2018).
193 Krendl, C. *et al.* GATA2/3-TFAP2A/C transcription factor network couples human pluripotent stem cell differentiation to trophectoderm with repression of pluripotency. *Proc Natl Acad Sci U S A* 114, E9579-E9588, doi:10.1073/pnas.1708341114 (2017).
194 Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* 8, 2281-2308, doi:10.1038/nprot.2013.143 (2013).
195 Mele, M. *et al.* Human genomics. The human transcriptome across tissues and individuals. *Science* 348, 660-665, doi:10.1126/science.aaa0355 (2015).
196 He, G., Xu W., Chen, Y., Liu, X. & Xi, M. Abnormal apoptosis of trophoblastic cells is related to the up-regulation of CYP11A gene in placenta of preeclampsia patients. *PLOS One*. 8, e59609, doi: 10.1371/journal.pone.0059609 (2013).
197 Michorowska, S. *et al.* Detection of ALDH3B2 in Human Placenta. *Int J Mol Sci* 20, doi:10.3390/ijms20246292 (2019).
198 Kohama, K., Kawaguchi, M., Fukushima, T., Lin, C. Y. & Kataoka, H. Regulation of pericellular proteolysis by hepatocyte growth factor activator inhibitor type 1 (HAI-1) in trophoblast cells. *Hum Cell* 25, 100-110, doi:10.1007/s13577-012-0055-2 (2012).
199 Bashir, M. *et al.* P66shc and its downstream Eps8 and Rac1 proteins are upregulated in esophageal cancers. *Cell Commun Signal* 8, doi:10.1186/1478-811X-8-13 (2010).
200 Chen, Y. J., Shen M.R., Chen YJ., Maa M.C. & Leu, T. H. Eps8 decreases chemosensitivity and affects survival of cervical cancer patients. *Mol Cancer Ther*. 7, 1376-1385, doi: 10.1158/1535-7163.MCT-07-2388 (2008).
201 Maa, M. C. *et al.* Eps8 facilitates cellular growth and motility of colon cancer cells by increasing the expression and activity of focal adhesion kinase. *J Biol Chem* 282, 19399-19409, doi:10.1074/jbc.M610280200 (2007).
202 Welsch, T., Endlich, K., Giese, T., Büchler, M. W. & Schmidt, J. Eps8 is increased in pancreatic cancer and required for dynamic actin-based cell protrusions and intercellular cytoskeletal organization. *Cancer Letters* 255, 205-218, doi:https://doi.org/10.1016/j.canlet.2007.04.008 (2007).

203 Yap, L. F. *et al.* Upregulation of Eps8 in oral squamous cell carcinoma promotes cell migration and invasion through integrin-dependent Rac1 activation. *Oncogene* 28, 2524-2534, doi:10.1038/onc.2009.105 (2009).

204 Kast, D. J. & Dominguez, R. IRSp53 coordinates AMPK and 14-3-3 signaling to regulate filopodia dynamics and directed cell migration. *Molecular Biology of the Cell* 30, 1285-1297, doi:10.1091/mbc.E18-09-0600 (2019).

205 Chou, A. M., Sem, K. P., Lam, W. J., Ahmed, S. & Lim, C. Y. Redundant functions of I-BAR family members, IRSp53 and IRTKS, are essential for embryonic development. *Sci Rep*. 7, 40485. doi:10.1038/srep40485 (2017).

206 Chawsheen, H. A., Ying, Q., Jiang, H. & Wei, Q. A critical role of the thioredoxin domain containing protein 5 (TXNDC5) in redox homeostasis and cancer development. *Genes Dis*. 5, 312-322, doi: 10.1016/j.gendis.2018.09.003 (2018).

207 Gundling, W. E., Jr. & Wildman, D. E. A review of inter- and intraspecific variation in the eutherian placenta. *Philos Trans R Soc Lond B Biol Sci*. 370, 20140072, doi:10.1098/rstb.2014.007 (2015).

208 Roberts, R. M., Green, J. A. & Schulz, L. C. The evolution of the placenta. *Reproduction* 152, 179-189, doi:10.1530/REP-16-0325 (2016).

209 Armstrong, D. L. *et al.* The core transcriptome of mammalian placentas and the divergence of expression with placental shape. *Placenta* 57, 71-78, doi: 10.1016/j.placenta.2017.04.015 (2017).

210 Sverdlov, E. D. Retroviruses and primate evolution. *Bioessays* 22, 161-171, doi:10.1002/(SICI)1521-1878(200002)22:2<161:AID-BIES7>3.0.CO;2-X (2000).

211 Mallet, F. *et al.* The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology. *Proc Natl Acad Sci U S A* 101, 1731-1736, doi:10.1073/pnas.0305763101 (2004).

212 Cohen, M. & Bischof, P. Factors regulating trophoblast invasion. *Gynecol Obstet Invest*. 64,126-130, doi:10.1159/000101734 (2007).

213 Lavillette, D. *et al.* The envelope glycoprotein of human endogenous retrovirus type W uses a divergent family of amino acid transporters/cell surface receptors. *J Virol*. 76, 6442-6452, doi:10.1128/jvi.76.13.6442-6452.2002 (2002).

214 Taruscio, D. & Mantovani, A. Factors regulating endogenous retroviral sequences in human and mouse. *Cytogenet Genome Res*. 105, 351-362, doi:10.1159/000078208 (2004).

215 Sundaram, V. *et al.* Widespread contribution of transposable elements to the innovation of gene regulatory networks. *Genome Res*. 24,1963-1976, doi:10.1101/gr.168872.113 (2014).

216 Vlahovic, M. *et al.* Changes in the placenta and in the rat embryo caused by the demethylating agent 5-azacytidine. *Int J Dev Biol* 43, 843-846 (1999).

217 Simpson, E. R. *et al.* Aromatase cytochrome P450, the enzyme responsible for estrogen biosynthesis. *Endocr Rev*.15, 342-355, doi:10.1210/edrv-15-3-342 (1994).

218 Fürbass, R., Selimyan R Fau - Vanselow, J. & Vanselow, J. DNA methylation and chromatin accessibility of the proximal *Cyp19* promoter region 1.5/2 correlate with expression levels in sheep placentomes. *Mol. Reprod. Dev*. 75, 1-7, doi:10.1002/mrd.20756 (2008).

219 Schulte, A. M. *et al.* Human trophoblast and choriocarcinoma expression of the growth factor pleiotrophin attributable to germ-line insertion of an endogenous retrovirus. *PNAS* 93, 14759-14764, doi: https://doi.org/10.1073/pnas.93.25.14759 (1996).

220 Schulte, A. M. *et al.* Influence of the human endogenous retrovirus-like element HERV-E.PTN on the expression of growth factor pleiotrophin: a critical role of a retroviral

Sp1-binding site. *Oncogene* 19, 3988–3998, doi: https://doi.org/10.1038/sj.onc.1203742 (2000).

221 Rumer, K. K., Uyenishi, J., Hoffman, M. C., Fisher, B. M. & Winn, V. D. Siglec-6 expression is increased in placentas from pregnancies complicated by preterm preeclampsia. *Reproductive sciences (Thousand Oaks, Calif.)* 20, 646-653, doi:10.1177/1933719112461185 (2013).

222 Steier, J. A., Ulstein M. & Myking, O. L. Human chorionic gonadotropin and testosterone in normal and preeclamptic pregnancies in relation to fetal sex. *Obstetrics & Gynecology* 100, 552-556, doi: https://doi.org/10.1016/S0029-7844(02)02088-4 (2002).

223 Steger, D. J., Hecht J. H. & Mellon, P. L. GATA-binding proteins regulate the human gonadotropin alpha-subunit gene in the placenta and pituitary gland. *Mol Cell Biol*. 14, 5592-5602, doi:10.1128/mcb.14.8.5592 (1994).

224 Robinson, B. G., Arbiser, J. L., Emanuel, R. L. & Majzoub, J. A. Species-specific placental corticotropin releasing hormone messenger RNA and peptide expression. *Molecular and Cellular Endocrinology* 62, 337-341, doi:https://doi.org/10.1016/0303-7207(89)90022-1 (1989).

225 Grino, M., Chrousos G.P. & Margioris, A. N. The corticotropin releasing hormone gene is expressed in human placenta. *Biochem Biophys Res Commun*. 148, 1208-1214, doi:10.1016/s0006-291x(87)80261-9 (1987).

226 Pavlicev, M., Hiratsuka, K., Swaggart, K. A., Dunn, C. & Muglia, L. Detecting endogenous retrovirus-driven tissue-specific gene transcription. *Genome Biol Evol* 7, 1082-1097, doi:10.1093/gbe/evv049 (2015).

227 Mongioví, A. M. *et al.* A novel peptide–SH3 interaction. *The EMBO Journal* 18, 5300-5309, doi:10.1093/emboj/18.19.5300 (1999).

228 Tocchetti, A., Confalonieri, S., Scita, G., Di Fiore, P. P. & Betsholtz, C. In silico analysis of the EPS8 gene family: genomic organization, expression profile, and protein structure. *Genomics* 81, 234-244, doi:https://doi.org/10.1016/S0888-7543(03)00002-8 (2003).

229 Offenhäuser, N. *et al.* The eps8 Family of Proteins Links Growth Factor Stimulation to Actin Reorganization Generating Functional Redundancy in the Ras/Rac Pathway. *Molecular Biology of the Cell* 15, 91-98, doi:10.1091/mbc.e03-06-0427 (2003).

230 Huang, R. *et al.* EPS8 regulates proliferation, apoptosis and chemosensitivity in BCR-ABL positive cells via the BCR-ABL/PI3K/AKT/mTOR pathway. *Oncol Rep*. 39, 119-128, doi:10.3892/or.2017.6102 (2018).

231 Scita, G. *et al.* EPS8 and E3B1 transduce signals from Ras to Rac. *Nature* 401, 290-293, doi:10.1038/45822 (1999).

232 Xu, M. *et al.* Epidermal growth factor receptor pathway substrate 8 is overexpressed in human pituitary tumors: role in proliferation and survival. *Endocrinology*. 150, 2064-2071. doi:10.1210/en.2008-1265 (2009).

233 Liu, P. S., Jong T.H., Maa, M. C. & Leu, T. H. The interplay between Eps8 and IRSp53 contributes to Src-mediated transformation. *Oncogene* 29, 3977–3989, https://doi.org/10.1038/onc.2010.144 (2010).

234 Wang, H., Patel V., Miyazaki, H., Gutkind, J. S. & Yeudall, W. A. Role for EPS8 in squamous carcinogenesis. *Carcinogenesis*. 30, 165-174, doi:10.1093/carcin/bgn252 (2009).

235 Funato, Y. *et al.* IRSp53/Eps8 Complex Is Important for Positive Regulation of Rac and Cancer Cell Motility/Invasiveness. *Cancer Research* 64, 5237, doi:10.1158/0008-5472.CAN-04-0327 (2004).

236 Xu, B., Li, J., Liu, X., Li, C. & Chang, X. TXNDC5 is a cervical tumor susceptibility gene that stimulates cell migration, vasculogenic mimicry and angiogenesis by down-regulating SERPINF1 and TRAF1 expression. *Oncotarget* 8, 91009-91024. doi:10.18632/oncotarget.18857 (2017).

237 Kikuchi, M., Doi E., Tsujimoto, I., Horibe, T. & Tsujimoto, Y. Functional analysis of human P5, a protein disulfide isomerase homologue. *The Journal of Biochemistry* 132, 451–455, doi: https://doi.org/10.1093/oxfordjournals.jbchem.a003242 (2002).

238 Cai, C. Y., Zhai, L. L., Wu, Y. & Tang, Z. G. Expression and clinical value of peroxiredoxin-1 in patients with pancreatic cancer. *Eur J Surg Oncol*. 41, 228-235, doi: 10.1016/j.ejso.2014.11.037 (2015).

239 Jiang, H., Wu, L., Mishra, M., Chawsheen, H. A. & Wei, Q. Expression of peroxiredoxin 1 and 4 promotes human lung cancer malignancy. *Am J Cancer Res*. 4, 445-460, (2014).

240 Ulbrich, S. E. *et al.* Bovine endometrial metallopeptidases MMP14 and MMP2 and the metallopeptidase inhibitor TIMP2 participate in maternal preparation of pregnancy. *Molecular and Cellular Endocrinology* 332, 48-57, doi:https://doi.org/10.1016/j.mce.2010.09.009 (2011).

241 Mishra, B. *et al.* Expression of extracellular matrix metalloproteinase inducer (EMMPRIN) and its related extracellular matrix degrading enzymes in the endometrium during estrous cycle and early gestation in cattle. *Reproductive Biology and Endocrinology* 8, 60, doi:10.1186/1477-7827-8-60 (2010).

242 Zhang, X., Qi, C. & Lin, J. Enhanced expressions of matrix metalloproteinase (MMP)-2 and –9 and vascular endothelial growth factors (VEGF) and increased microvascular density in the endometrial hyperplasia of women with anovulatory dysfunctional uterine bleeding. *Fertility and Sterility* 93, 2362-2367, doi: https://doi.org/10.1016/j.fertnstert.2008.12.142 (2010).

243 Shimonovitz, S. *et al.* Developmental regulation of the expression of 72 and 92 kd type IV collagenases in human trophoblasts: A possible mechanism for control of trophoblast invasion. *American Journal of Obstetrics and Gynecology* 171, 832-838, doi:https://doi.org/10.1016/0002-9378(94)90107-4 (1994).

244 Suman, P. & Gupta, S. K. Comparative analysis of the invasion-associated genes expression pattern in first trimester trophoblastic (HTR-8/SVneo) and JEG-3 choriocarcinoma cells. *Placenta* 33, 874-877, doi:https://doi.org/10.1016/j.placenta.2012.06.017 (2012).

245 Su, M.-T., Tsai, P.-Y., Tsai, H.-L., Chen, Y.-C. & Kuo, P.-L. miR-346 and miR-582-3p-regulated EG-VEGF expression and trophoblast invasion via matrix metalloproteinases 2 and 9. *BioFactors* 43, 210-219, doi:10.1002/biof.1325 (2017).

246 Isaka, K. *et al.* Expression and Activity of Matrix Metalloproteinase 2 and 9 in Human Trophoblasts. *Placenta* 24, 53-64, doi:https://doi.org/10.1053/plac.2002.0867 (2003).

247 Chen, J. & Khalil, R. A. Matrix Metalloproteinases in Normal Pregnancy and Preeclampsia. *Prog Mol Biol Transl Sci*. 148, 187-165. doi: 10.1016/bs.pmbts.2017.04.001 (2017).

248 Li, W., Mata, K. M., Mazzuca, M. Q. & Khalil, R. A. Altered matrix metalloproteinase-2 and -9 expression/activity links placental ischemia and anti-angiogenic sFlt-1 to uteroplacental and vascular remodeling and collagen deposition in hypertensive pregnancy. *Biochemical Pharmacology* 89, 370-385, doi:https://doi.org/10.1016/j.bcp.2014.03.017 (2014).

249 Bischof, P., Meisser, A. & Campana, A. Biochemistry and molecular biology of trophoblast invasion. *Annals of the New York Academy of Sciences* 943, 157-162, doi:10.1111/j.1749-6632.2001.tb03799.x (2001).

250 Rahat, B., Hamid, A., Ahmad Najar, R., Bagga, R. & Kaur, J. Epigenetic mechanisms regulate placental c-myc and hTERT in normal and pathological pregnancies; c-myc as a novel fetal DNA epigenetic marker for pre-eclampsia. *Molecular human reproduction* 20, 1026-1040, doi:10.1093/molehr/gau053 (2014).
251 Brosens I.A., Robertson, W. B. & Dixon, H. G. The role of the spiral arteries in the pathogenesis of preeclampsia. Obstet Gynecol Annu. 1, 177-91, (1972).
252 Uzan, J., Carbonnel M., Piconne, O., Asmar, R. & Ayoubi, J. M. Pre-eclampsia: pathophysiology, diagnosis, and management. *Vasc Health Risk Manag*. 7, 467-474, doi:10.2147/VHRM.S20181 (2011).
253 Scifres, C. M. & Nelson, D. M. Intrauterine growth restriction, human placental development and trophoblast cell death. *The Journal of physiology* 587, 3453-3458, doi:10.1113/jphysiol.2009.173252 (2009).
254 Hung, T. H. & Burton, G. J. Hypoxia and reoxygenation: a possible mechanism for placental oxidative stress in preeclampsia. *Taiwan J Obstet Gynecol*. 45, 189-200, doi:10.1016/S1028-4559(09)60224-2 (2006).
255 Wu, F., Tian, F.-J. & Lin, Y. Oxidative Stress in Placenta: Health and Diseases. *BioMed research international* 2015, 293271-293271, doi:10.1155/2015/293271 (2015).

6. Supplementary information

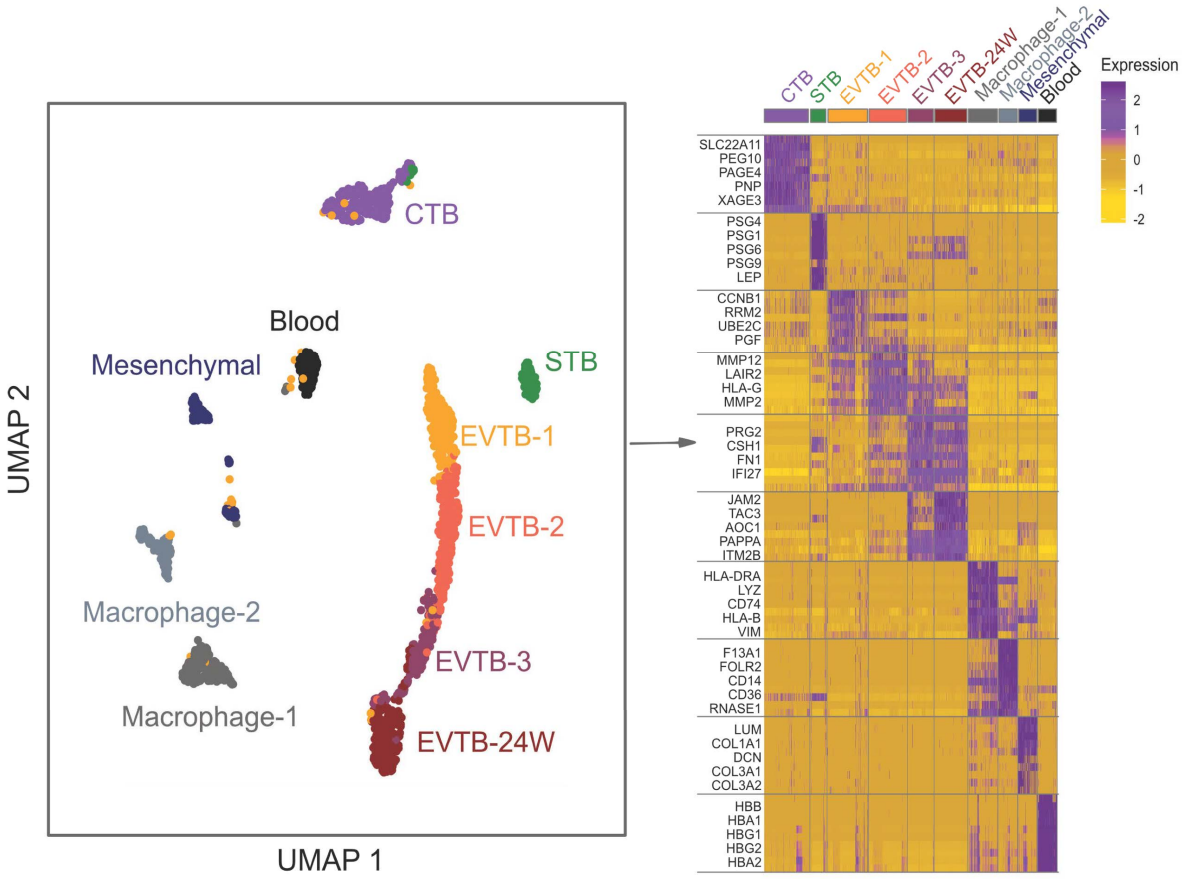

Figure 21. Characterization of single cell RNA-seq data of human placenta tissue
Graphical representation of identified trophoblast sub-type clusters from 8th and 24th week of gestation (PMID: 30042384; GEO accession: GSE89497) as shown in the UMAP plot (left side of the figure). The cell type for each cluster was identified according to the known marker genes as shown in the heat-map (right side of the figure). Purple color shows enriched gene expression.

Table 11. Contributions by others

The following data sets were analyzed and figures were generated by Dr. Manvendra Singh.

	Published Data Set	Figure
1.	**RNA-seq data** on human tissues. (Illumina Body Map 2.0; E-MTAB-513)	Figure 6. Expression of transposable elements (TEs) in human tissues.
2.	**RNA-seq data** on human tissues. (GTEx data; PMID: 25954002)	Figure 8A. Trophoblast-specific expression of dysregulated genes in PE.
3.	**Single cell RNA-seq data** on human placenta tissue from 8th and 24th week of gestation. (GEO accession: **GSE89497**; PMID: 30042384)	Figure 8C. Trophoblast-specific expression of dysregulated genes in PE. Figure 21. Characterization of single cell RNA-seq data of human placenta tissue.
4.	**ChIP-seq data on** binding sites of • GATA2, GATA3, TFAP2A, TFAP2C transcription factors in trophoblast progenitors derived from human embryonic stem cells (H9) (GEO accession: **GSE105081**; PMID: 29078328) • H3K9Ac, H3K27Ac (active) and H3K27Me3 (repressive) histone marks in differentiated syncytiotrophoblasts (GEO accession: **GSE127288**; PMID: 31294776) • H3K4Me1 and H3K27Ac in human term placenta tissue. (GEO accession: **GSE118289**; PMID: 30231016)	Figure 9A. ERV-LTR mediated trophoblast-specific gene regulation.
5.	**Microarray data** on human placenta tissue. (Oslo-Cohort-I) (PMID: 28904069)	Figure 17C. EPS8L1 overexpression mimics dysregulated pathways in PE.
	Data generated in this study	**Figure**
6.	**RNA-seq data** on *EPS8L1*-OE-SGHPl-4 cells and WT-SGHPl-4 cells.	Figure 17A and 17B. EPS8L1 overexpression mimics dysregulated pathways in PE.

7. Appendix

7.1. List of Figures

7.2. List of Tables

8. Abbreviations

A	Absorbance
AA	Antibiotic-Antimycotic
AEC	Aminoethylcarbazole
APS	Ammoniumperoxodisulfat
AT_1	Angiotensin II receptor
BGI	Beijing Genomics Institute, Company
BMI	Body mass index
BSA	Bovine Serum Albumine
Cas9	CRISPR associated protein 9
CCND1	Cyclin D1
ChIP-seq	Chromatin Immunoprecipitation and sequencing
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
Da	Dalton
DAPI	4',6-Diamidino-2-phenylindole
DCFH	2′,7′-Dichlorofluorescin Diacetate
DMEM	Dulbecco's Modified Eagle's medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
EGFP	EGFP enhanced green fluorescent protein
ER	Endoplasmic reticulum
ERVs	Endogenous Retrovirus Elements
EVTs	Extravillous trophoblasts
ELISA	Enzyme-linked immunosorbent assay
FACS	Fluorescence Activated Cell Sorting
FBS	Fetal Bovine Serum
FCS	Fetal Calf Serum
GTEx	The Genotype Tissue Expression
GFP	Green Fluorescent Protein
HBSS	Hanks Balanced Salt
HD	Homeodomain
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HERVH	Human endogenous retrovirus, family H
HERVK	Human endogenous retrovirus, family K
HERVW	Human endogenous retrovirus, family W
HIER	Heat Induced Epitope Retrieval
HTS	High throughput sampler
ICM	Inner cell mass
IUGR	Intrauterine Growth Retardation
IP	Immunoprecipitation
KO	Knock-out
LTRs	Long Terminal Repeats
LINEs	Long Interspersed Elements

M	Molar
MFI	Median Fluorescent Intensity
MS	Mass Spectrometry
MMP2	MMP2 matrix-metalloprotease 2
N/A	Not Applicable
NBCS	New Born Calf Serum
ORF	Open Reading Frame
OE	Overexpression
PBS	Phosphate Buffered Saline
PlGF	Placental Growth Factor
PCR	Polymerase Chain Reaction
PE	Preeclampsia
PVDF	Polyvinylidene fluoride
RIPA	Radioimmunoprecipitation assay
RNA	Ribonucleic Acid
RNA-seq	RNA sequencing
ROS	Reactive Oxygen Species
RPM	Revolutions per Minute
RT	Room Temperature
SB	Sleeping Beauty
SDS-PAGE	Sodium dodecylsulfate-polyacrylamide gel electrophoresis
SEM	Standard error of mean
sFlt-1	Soluble FMS-like tyrosine kinase receptor-1
sgRNA	Single guide RNA
SINEs	Short interspersed elements
SV40	Simian virus 40
SCT	Syncytiotrophoblast
TBS	Tris Buffered Saline
TE	Transposable Elements
TES	Transcriptional end site
TEMED	Tetramethylethylenediamine
TNF	Tumor necrosis factor
Tris	Tris(hydroxmethyl)aminomethane
TSS	Transcriptional start sites
TSG	Trophoblast specific genes
UPR	Unfolded protein response
VCT	Villous cytotrophoblast
VEGF	Vascular Endothelial Growth Factor
VEGFR-1	Vascular Endothelial Growth Factor Receptor 1
WT	Wild Type

9. Publications

- Rabia Anwar, Manvendra Singh, Florian Herse, Ralf Dechend, Zsuzsanna Izsvák "*Dysregulated trophoblast-specific gene expression mediated by retroviral regulatory sequences contributes to preeclampsia (PE)*". *(In process)*

9.1. Active congress participation

09. 2019	**Poster presentation**	Placenta: the origin of pregnancy health and disease, IFPA: "Trophoblast-specific gene expression driven by endogenous retrovirus LTRs in normal vs preeclamptic (PE) pregnancy", Buenos Aires, Argentina. Received *YW Loke early career research (ECR)* travel award and also selected as a finalist for the *Elsevier Trophoblast ECR* award 2019.
04.2019	**Course**	Helmholtz Career and Leadership course, Bonn, Germany.
09. 2018	**Poster presentation**	Clinical growth via placenta, IFPA: "Trophoblast-specific gene expression driven by endogenous retrovirus LTRs in normal vs preeclamptic (PE) pregnancy", Tokyo, Japan.
10. 2017	**Oral presentation**	Keystone Symposia meeting on Maternal-Fetal Crosstalk: Harmony vs. Conflict: "Trophoblast-specific gene expression driven by endogenous retrovirus LTRs in normal vs preeclamptic (PE) pregnancy", Washington D.C., USA. *Poster was selected for short talk*.
10. 2017	**Oral presentation**	Interdisciplinary Autumn School for Reproductive Sciences and related Research Fields: "Trophoblast-specific gene expression driven by endogenous retrovirus LTRs in normal vs preeclamptic (PE) pregnancy", Magdeburg, Germany. *Guest speaker*.
08. 2017	**Poster presentation**	Helmholtz-presentation and communication skills course, Stuttgart-Hohenheim, Germany.
07. 2017	**Course**	Placenta Biology Course, Cambridge, UK.
06. 2017	**Poster presentation**	TransCard Ph.D. Retreat: "Trophoblast-specific gene expression driven by endogenous retrovirus LTRs in normal vs preeclamptic (PE) pregnancy", Valencia, Spain.

10. Acknowledgement

All praises to Almighty Allah, the Creator and the Sustainer of the universe, who is the origin of all knowledge and wisdom. It is He who enabled me to successfully complete my research.
I would like to express my sincere gratitude to my supervisor Dr. Zsuzsanna Izsvák. Without her constant support, trust and guidance, it wouldn't have been possible for me to complete my Ph.D. I would also like to thank my sponsor, National University of Sciences and Technology, Pakistan (Faculty Development Program) for funding my studies for three years.
I would also like to thank our collaborators from Experimental and Clinical Research Center (ECRC), Dr. Ralf Dechend and Dr. Dominik Muller for their valuable advice and scientific discussions related to my project. I would like to thank specially Dr. Dechend for his valuable input and guidance and for his help to get samples from the hospital. My honest gratitude to Dr. Florian Herse for his valuable advice, support and continuous help related to my project. I am thankful to all my colleagues at ECRC. I would like to also acknowledge Julianne Anders for her invaluable technical assistance to my project.
I would like to thank all my previous and current lab fellows from the Mobile DNA group including, Dr. Tamas Rasko, Dr. Christine Romer, Dr. Angelica Garcia-Perez, Dr. Huqiang Cai, Dr. Jichang Wang, Dr. Chuanbo Sun, Dr. Suneel Narayanavari, Dr. Amit Pande, Sandra Neuendorf, Beatrice Heuser, Felix Lundberg, Kathrin Radscheit, Katarina Stevanovic, Ilija Bilic, Aleksandra Kondrashkina, Guo Yong, Himanshu Bhusan Samal, Zhimin Zhou, Yuliang Qu, Karam Ibrahim, Chigozie Joseph Samuel and Bertrand Tangu Teneng for their scientific advices and technical support. I am thankful to Dr. Manvendra Singh for performing all the bioinformatic analysis. I sincerely thank Dr. Julia Zadora for her help in the project. I want to thank Dr. Sanam Bashir for her support not only in the lab but also outside the lab. I want to thank our group's secretory Beate Valeske for her kindness and administrative support.
My deepest gratitude goes to my parents Muhammad Anwar & Lubna Anwar, my brothers Manan & Abdullah, my father-in-law Mr. Khalid, my uncle Mr. Muhammad Akram and rest of the family members for their support, kindness, love and trust throughout my life. I especially would like to thank my beloved (late) grandmother Fehmida for her constant prayers, love and kindness. Without her prayers and support, I wouldn't have successfully achieved my Ph.D. I would like to thank my friends Dr. Miriam and Nida for their support. I would like to extend my gratitude to Dr. Miriam for her guidance in writing. Finally, I would

like to take the opportunity to thank my loving husband, Omair, without his unconditional love, support and faith in me, it wouldn't have been possible for me to achieve my goals.

11. Eigenständigkeitserklärung

Hiermit erkläre ich, Rabia Anwar, dass ich die vorliegende Arbeit selbstständig erarbeitet und verfasst sowie keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Berlin Rabia Anwar