Disturbed Placental Imprinting in Preeclampsia Leads to Altered Expression of DLX5, a Human-Specific Early Trophoblast Marker

BACKGROUND: Preeclampsia is a complex and common human-specific pregnancy syndrome associated with placental pathology. The human specificity provides both intellectual and methodological challenges, lacking a robust model system. Given the role of imprinted genes in human placentation and the vulnerability of imprinted genes to loss of imprinting changes, there has been extensive speculation, but no robust evidence, that imprinted genes are involved in preeclampsia. Our study aims to investigate whether disturbed imprinting contributes to preeclampsia.

METHODS: We first aimed to confirm that preeclampsia is a disease of the placenta by generating and analyzing genome-wide molecular data on well-characterized patient material. We performed high-throughput transcriptome analyses of multiple placenta samples from healthy controls and patients with preeclampsia. Next, we identified differentially expressed genes in preeclamptic placentas and intersected them with the list of human imprinted genes. We used bioinformatics/statistical analyses to confirm association between imprinting and preeclampsia and to predict biological processes affected in preeclampsia. Validation included epigenetic and cellular assays. In terms of human specificity, we established an in vitro invasion-differentiation trophoblast model. Our comparative phylogenetic analysis involved single-cell transcriptome data of human, macaque, and mouse preimplantation embryogenesis.

RESULTS: We found disturbed placental imprinting in preeclampsia and revealed potential candidates, including *GATA3* and *DLX5*, with poorly explored imprinted status and no prior association with preeclampsia. As a result of loss of imprinting, *DLX5* was upregulated in 69% of preeclamptic placentas. Levels of *DLX5* correlated with classic preeclampsia markers. DLX5 is expressed in human but not in murine trophoblast. The DLX5^{high} phenotype resulted in reduced proliferation, increased metabolism, and endoplasmic reticulum stress-response activation in trophoblasts in vitro. The transcriptional profile of such cells mimics the transcriptome of preeclamptic placentas. Pan-mammalian comparative analysis identified *DLX5* as part of the human-specific regulatory network of trophoblast differentiation.

CONCLUSIONS: Our analysis provides evidence of a true association among disturbed imprinting, gene expression, and preeclampsia. As a result of disturbed imprinting, the upregulated *DLX5* affects trophoblast proliferation. Our in vitro model might fill a vital niche in preeclampsia research. Human-specific regulatory circuitry of *DLX5* might help explain certain aspects of preeclampsia.

Julianna Zadora, PhD*
Manvendra Singh, MS*
Florian Herse, PhD
Lukasz Przybyl, PhD
Nadine Haase, PhD
Michaela Golic, MD
Hong Wa Yung, PhD
Berthold Huppertz, PhD
Judith E. Cartwright, PhD
Guy Whitley, PhD
Guro M. Johnsen, PhD
Giovanni Levi, PhD
Annette Isbruch, MD
Herbert Schulz, PhD
Friedrich C. Luft, MD
Dominik N. Müller, PhD
Anne Cathrine Staff, MD
Laurence D. Hurst, PhD†
Ralf Dechend, MD†
Zsuzsanna Izsvák, PhD†

*Dr Zadora and M. Singh contributed equally.

†Drs Hurst, Dechend, and Izsvák contributed equally.

Correspondence to: Zsuzsanna Izsvák, PhD, Mobile DNA, Max Delbrück Center for Molecular Medicine in the Helmholtz Association, Robert Rössle Strasse 10, 13125 Berlin, Germany, Laurence D. Hurst, PhD, Milner Centre for Evolution, Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, UK, or Ralf Dechend, MD, Experimental and Clinical Research Center, Lindenberger Weg 80, 13125 Berlin, Germany. E-mail zizsvak@mdc-berlin.de, l.d.hurst@bath.ac.uk, or ralf.dechend@charite.de

Sources of Funding, see page 1837

Key Words: epigenetics ■ gene expression and regulation ■ genomic imprinting ■ preeclampsia ■ trophoblasts

Clinical Perspective

What Is New?

- Unbiased analysis of genome-wide molecular and clinical data identifies *DLX5* as an imprinted target gene with novel placental function in preeclampsia.
- We observe that *DLX5* is paternally imprinted in the human placenta and its expression is dysregulated in preeclampsia.
- We provide evidence for a mechanistic coupling between preeclampsia and disturbed placental imprinting (loss of imprinting) as a causal role in preeclampsia.
- Our data indicate that DLX5 has a role in trophoblast proliferation and differentiation (syncytialization).
- DLX5-induced overexpression in trophoblasts can faithfully model preeclampsia in a cell culture system, signifying the contribution of a single transcription factor and providing a potential cellular model both for further research and for analysis of drugable targets.
- DLX5 is expressed in human but not in mouse trophoblasts, underlining the human specificity of preeclampsia. Our study highlights the diverged cellular function of DLX5 during mammalian embryonic evolution.

What Are the Clinical Implications?

- Our analysis supports the view that preeclampsia is not a single but a heterogeneous disease with disturbed imprinting commonplace.
- Unsupervised clustering analysis identified 3 distinct transcriptomic classes of preeclampsia, not clustering with the intrauterine growth restriction. Our clusters are in conjunction with previously suggested classification of preeclampsia (early- versus late-onset) but subdivide early-onset preeclampsia further.
- Our study identifies an early preeclampsia cluster that can be clearly characterized by elevated *DLX5* levels, disturbed epigenetics, and similar clinical manifestation.
- We find that levels of *DLX5* correlate with a circulating biomarker, placental growth factor. Whether DLX5 will have utility as a biomarker is unclear because its loss of imprinting is not observed in all instances (69% of all preeclampsia cases).
- Our cellular model has potential for further clinically related research, including analysis of drugable targets.

Preeclampsia is a complex, heterogeneous syndrome characterized by high blood pressure (≥140/90 mm Hg) after the 20th week of pregnancy in association with proteinuria (≥300 mg/L per 24 hours).[1] Preeclampsia is the first sex-specific cardiovascular risk factor,[2] affecting 2% to 8% of human pregnancies, and remains a leading cause of maternal and perinatal mortality.[3]

Despite considerable research efforts, the cause of preeclampsia is not fully understood. Preeclampsia is assumed to be associated with reduced fetal trophoblast invasion and impaired remodeling of maternal spiral arteries leading to poor uteroplacental perfusion. The improper placentation process triggers oxidative and endoplasmic reticulum (ER) stress and results in defective protein synthesis in the placenta. As a consequence, dysregulated expression of inflammatory, angiogenic, and antiangiogenic factors is observed.[4] Currently, the only treatment is delivery, pinpointing the potentially central role of the placenta (or more generally the maternal-fetal interaction) in the disease.[5]

The placenta has unique epigenetic features, including low levels of genomic DNA methylation and a specific expression pattern of imprinted genes, which are the prime candidates to be associated with the evolution of intrauterine development.[6] Because epigenetic disruption of imprinted gene was associated with certain diseases, the potential involvement of disturbed regulation of imprinted genes in preeclampsia has also been intensively discussed.[7–11]

Preeclampsia is not confirmed in other mammals and is considered to be human-specific (although a possible eclampsia event was observed in a gorilla).[12] The human specificity of preeclampsia generates real challenges. First, despite intensive research, the genetic background of the human specificity and the pathogenesis of preeclampsia remain poorly understood. Second, although rodent models have been suggested, no animal models have proven suitable.

Because placentation is a diverse process even among primates,[13] genes associated preferentially with placental expression (eg, imprinted genes) are potential candidates for understanding human specificity. This suggestion is centered largely on the understandings that the dosage of imprinted genes can have phenotypic impact, that the placenta is a hot spot for the activity of imprinted genes, and hence that loss of imprinting (LOI) will be likely to affect matters at the maternal-fetal interface.

Despite the abundant speculation that disruption of imprinting and preeclampsia might be mechanistically coupled, the evidence is at best limited. Indeed, a recent review of the genetics of preeclampsia[14] reports no robust evidence for a mechanistic coupling with imprinting. What evidence there has been is circumstantial, often negative, or not replicable.[15–19] A further barrier to effective research, until recently,[20] has been the lack of a definitive list of genes imprinted in the human placenta.

To get a better understanding of this highly complex disease, we aim to confirm our hypotheses that preeclampsia is not a single but a heterogeneous disease of the placenta; that it is associated with improper tro-

phoblast function; and that epigenetic turbulence can result in abnormal expression of imprinted genes, which in turn compromise proper maternal-fetal interaction and contribute to preeclampsia, specifically in humans. Our experimental strategy is based on a cross-disciplinary approach that uses the current catalog of human imprinted placental genes and combines the generation and analyses of genome-wide molecular data with well-characterized patient material followed by experimental validation. We use molecular and cellular technologies and bioinformatic analyses to predict biological processes affected in preeclampsia, followed by wet-bench validation. We also aimed at establishing an in vitro model to fill a vital niche in preeclampsia research, explicitly because there is no robust animal model system. Finally, we data-mine existing single-cell transcriptome data to possibly shed light on the human-specific nature of preeclampsia.

METHODS

Detailed methods are provided in the online-only Data Supplement.

Patients

Microarray data of human placenta and decidua samples were described earlier.[21] The study also consists of placental and decidual tissues from 56 preeclamptic women and 28 women with normotensive and uncomplicated pregnancies described earlier.[22] The preeclamptic group was subdivided into preeclampsia with intrauterine growth restriction (n=14) and preeclampsia without intrauterine growth restriction (n=42). Furthermore, the preeclamptic group was divided into early-onset preeclampsia (delivery before gestational week 34) and late-onset preeclampsia (delivery during or after gestational week 34). The Regional Committee of Medical Research Ethics in South-Eastern Norway approved the study, and all the subjects gave informed written consent.

Healthy (n=5) and preeclamptic (n=5) primary trophoblast cells were isolated from human placenta samples obtained from HELIOS Klinikum in Berlin. Human placenta sampling was approved by the Regional Committee of the Medical Faculty of Charité Berlin.

Statistics

Data are presented as mean±SEM (for normally distributed data) or median with interquartile range (for nonnormally distributed data). Normality was assessed by Kolmogorov-Smirnov tests. Techniques for each analysis are specified in the figure legends. Values of $P<0.05$ were considered statistically significant.

RESULTS

Cluster Analysis Identifies 3 Distinct Transcriptome Patterns of Preeclamptic Placenta

To determine which genes have aberrant expression in preeclampsia, we sought to first better understand whether preeclampsia is 1 disease or many and then to look for genes misregulated across all preeclamptic subtypes. We compared placenta and decidua samples from 24 patients with preeclampsia and 22 healthy women. Nineteen (of 24) preeclamptic placental samples generated 3 distinct molecular groups (PE_P1 through PE_P3; Figure 1A), whereas control samples (16 of 22) generated 2 groups (C1, C2). Five (of 24) preeclamptic and 4 (of 22) healthy placenta samples grouped with the opposite cluster. Whereas placental PE_P1 and PE_P3 contain samples from early-onset preeclampsia, PE_P2 included samples mainly from late-onset preeclampsia. We did not reveal any significant grouping for the intrauterine growth restriction phenotype. An alternative clustering method based on euclidean distances identified the same 3 preeclampsia clusters (Figure IA in the online-only Data Supplement). In contrast to placental samples, the preeclampsia and control decidua samples appeared to scatter randomly (Figure IB in the online-only Data Supplement), supporting that clinically established preeclampsia is a placental, not a decidual, disorder. The negative result for the decidua also provides a negative control for false/artifactual clustering.

A Subset of Early-Onset Preeclampsia Correlates With Clinical Symptoms

Is there an association between the preeclamptic gene clusters and clinical data? We clustered 36 patients' clinical and biomarker parameters on their relative values (Table I in the online-only Data Supplement). Our analysis suggests that the gene expression profile can be related to clinical disease phenotyping, but only to a certain extent (Figure 1B and Figure IIA and IIB in the online-only Data Supplement). Besides a related transcriptome, PE_P1 shared a similar clinical profile. In contrast, most of the patients from PE_P2 generated a cluster together with preeclamptic samples that had diverse transcriptome, and patients from PE_P3 had a diverse clinical manifestation.

We also performed correlation analyses of clinical and gene expression data across all control and patient samples to identify genes associated with any of the clinical phenotypes. We found correlated gene expression with maternal placental growth factor (PlGF; eg, *PAPPA2*, *SPAG4*, *ENG*, and *LEP)* and sFlt1 (soluble fms-like tyrosine kinase-1) levels (eg, *SPAG4*, *ENG*, *ANKRD37*, and *ERRFI1*). Among the imprinted genes, the expression of both *DLX5* and *GATA3* correlated with diastolic blood pressure and anticorrelated with gestation age (Figure IIC in the online-only Data Supplement). *DLX5* also correlated with sEndoglin (positive) and with baby weight (negative).

All Preeclamptic Clusters Have Disturbed Guidance Signaling

We then analyzed differentially expressed genes (DEGs) in each preeclampsia cluster compared with control

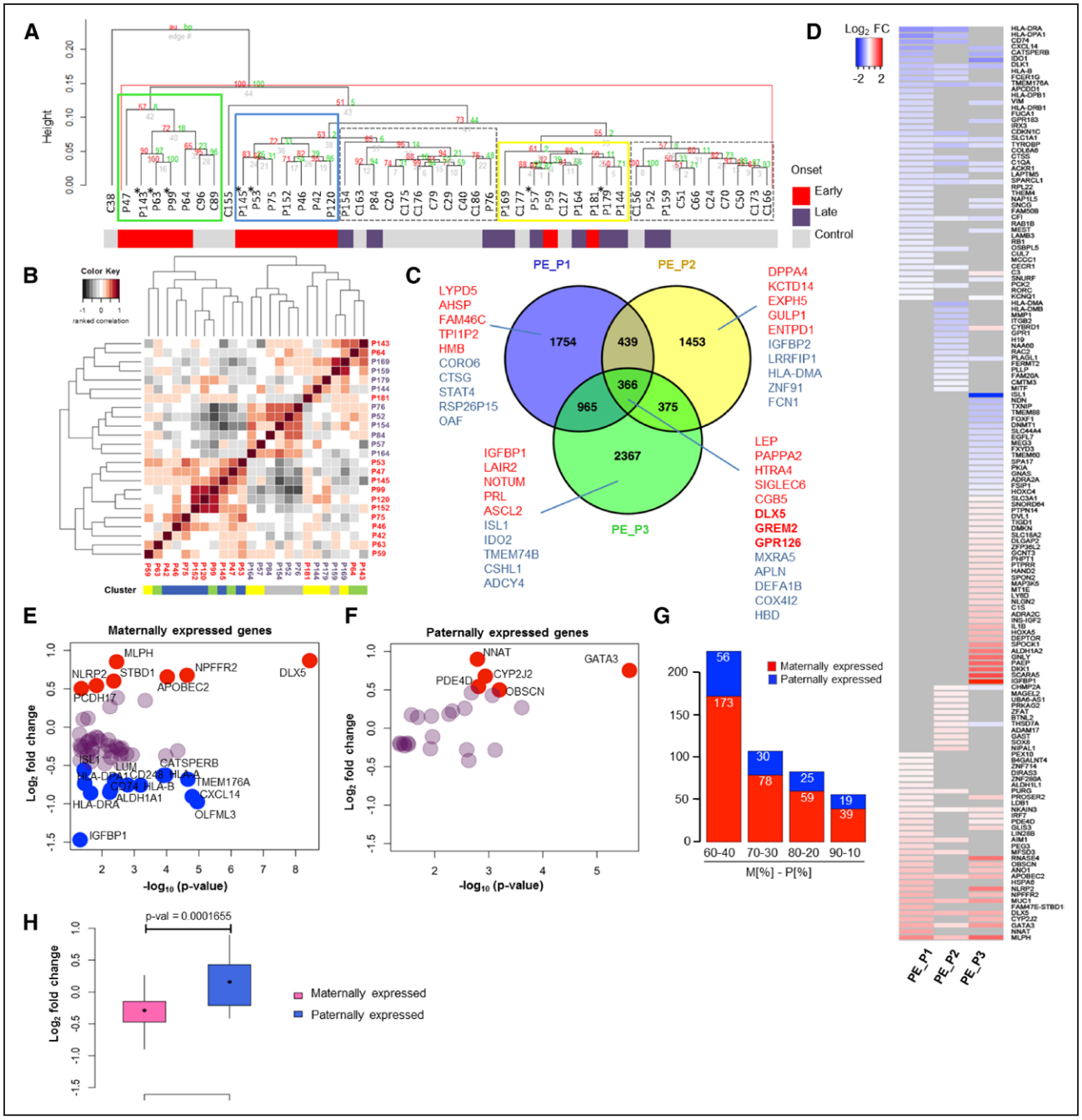


Figure 1. Gene expression analysis of preeclamptic and healthy placenta samples.
A, Hierarchical clustering analysis of microarray data identified 3 preeclamptic groups in placenta, PE_P1 (blue), PE_P2 (yellow), and PE_P3 (green), and 2 control groups, C1 and C2 (gray, dashed) (control placenta, n=22; preeclamptic placenta, n=24). The onset of preeclampsia (by gestational age at delivery) for each sample is indicated. *Preeclampsia+intrauterine growth restriction samples. **B**, Clustering analysis of clinical data. Heat map representing pairwise correlation between patients based on the clinical data (Table I in the online-only Data Supplement). Hierarchical cluster dendrogram was calculated with ranked correlation and complete linkage method on relative values from the clinical variables (n=36). Height of dendrograms represents the euclidian distance. **C**, Venn diagram of differentially expressed genes (DEGs) in 3 preeclamptic clusters (DEG on $P<0.05$). Top upregulated (red) and downregulated (blue) genes for each cluster are shown. **D**, Imprinted genes in preeclampsia. Heat map representing differential expression of imprinted genes in the 3 preeclamptic clusters vs. controls (DEG $P<0.05$). Log_2-fold change of differential expression of maternally **(E)** and paternally **(F)** expressed genes (MEGs and PEGs) in all preeclampsia cases (DEG $P<0.05$). Genes with log_2-fold change ≥0.5 are shown in red and with ≤–0.5 change in blue. **G**, Distribution of differentially expressed MEGs and PEGs in preeclampsia according to different criteria for maternal to paternal allelic expression ratio (MEGs, M[%]-PEGs, P[%]): 60-40, 70-30, and 90-10. **H**, Wilcoxon test on the mean log_2-fold change expression of differentially expressed MEGs and PEGs in preeclampsia.

samples. Each cluster contained >1000 unique DEGs with 366 dysregulated genes common to all 3 clusters ($P<0.05$; Figure 1C and Table II in the online-only Data Supplement). To characterize the clusters and to find possible interactions, data sets containing 3525, 2634, and 4073 DEGs corresponding to preeclamptic clusters PE_P1, PE_P2, and PE_P3, respectively, were analyzed with Ingenuity Pathway Analysis and GOrilla gene ontology tools. Axonal guidance was the top pathway in all 3 preeclamptic clusters, supporting the view that disturbed angiogenesis and disturbed cell–cell communication between endothelial cells and the trophoblasts via the mechanism of axonal guidance are common features of preeclampsia.[23,24] Other than this universal property, the 3 clusters exhibited unique features (Figure III and Table II in the online-only Data Supplement).

Imprinted Genes Exhibit Differential Expression in Preeclamptic Placenta

Next, we focused on differentially expressed imprinted genes in preeclampsia. We used a merged list of genes of the Geneimprint database (184 validated and putative human imprinted genes) and the recent list of imprinted human placental genes (223; maternal:paternal allelic expression ratio, 80:20)[20] (Table III in the online-only Data Supplement). Our analysis identified 150 imprinted/putatively imprinted genes that were dysregulated in at least 1 of the 3 preeclamptic clusters ($P<0.05$; Figure 1D and Table III in the online-only Data Supplement). Maternally expressed genes (MEGs) and paternally expressed genes (PEGs) were affected according to their relative commonality (23%, 59 of 257, versus 17%, 25 of 150; $\chi^2=1.9$, $P=0.16$; Figure 1F and Table III in the online-only Data Supplement). This was observed also when maternal/paternal allelic expression ratios were considered at different stringencies (Figure 1G). The mean expression of MEGs, however, was significantly downregulated compared with PEGs (Figure 1H). Nevertheless, among the most significantly deregulated genes, we identified both MEGs (eg, *DLX5, APOBEC2*, *CD74*) and PEGs (eg, *GATA3*, *CYP2J2*) (Figure 1E and 1F). In our study, the 2 most significantly dysregulated imprinted genes were *GATA3* and *DLX5* ($P<2.51\times10^{-6}$ and $P<3.65\times10^{-9}$, respectively). Furthermore, we focus on *DLX5*, which is a MEG in human lymphoblasts and brain,[25] but its imprinted status in the placenta is not yet explored.

DLX5 Is Upregulated in Preeclamptic Placenta

We used quantitative reverse-transcription–polymerase chain reaction to confirm the upregulation of *DLX5* in preeclampsia (Figure IVB in the online-only Data Supplement). We then confirmed these findings in a second independent patient cohort of 56 compared with 28 controls. Although *DLX5* was significantly upregulated in both early-onset ($P<0.0001$) and late-onset ($P<0.01$) preeclampsia (Figure 2A) and in all 3 preeclamptic clusters, its expression level varied between placental samples. Nevertheless, altogether, 69% of preeclampsia samples (n=56) could be associated with *DLX5* overexpression. We also confirmed increased DLX5 protein expression in preeclamptic placenta tissues by Western blotting (Figure IVC and IVD in the online-only Data Supplement). Correlation analysis of gestation age to *DLX5* expression in control or preeclamptic placentas excluded gestation age–related changes in *DLX5* differential expression (Figure IVE in the online-only Data Supplement). In tissues and cells derived from a pregnancy-related tissue panel, *DLX5* expression was detected in placenta and in trophoblasts but was less pronounced in the decidua (Figure IVF and IVG in the online-only Data Supplement). Immunofluorescent staining detected coexpression of DLX5 with cytokeratin-7, a trophoblast-specific marker (Figure 2B). Immunohistochemistry confirmed elevated DLX5 protein expression in placental tissues of early-onset preeclampsia (Figure 2C).

To associate our findings with clinical preeclampsia biomarkers, we compared the placental expression of *DLX5* with the antiangiogenic sFlt1 and sFlt1/PlGF ratio and with proangiogenic PlGF concentration in maternal serum.[26] Although there was no significant correlation between *DLX5* expression and sFlt1 levels, we observed a negative correlation ($r=-0.35$, $P=0.017$) with PlGF concentrations comparing women with preeclampsia and healthy women (Figure 2D). This observation suggests that placental *DLX5* expression is associated with decreased levels of a maternal circulating placental biomarker in preeclampsia.

Loss of Imprinting Results in Elevated Gene Expression of *DLX5*

To investigate whether alterations in the predicted imprinting status of *DLX5* could be responsible for its overexpression in the preeclamptic placentas, we performed an LOI assay.[27] We measured expression of the silenced allele in placental samples carrying the heterozygous *DLX5* single nucleotide polymorphism (rs73708843). Primers are provided in Table IV in the online-only Data Supplement. Of 97 placental tissues, 42 were genotyped as heterozygous (43.3%), including 16 control (16.5%) and 26 preeclamptic (26.8%) placentas. The mean expression of the putatively inactive *DLX5* allele was 58% in preeclampsia when the expression of the nonimprinted allele for each individual sample was set at 100%. Control samples also exhibited LOI but with significantly less frequent

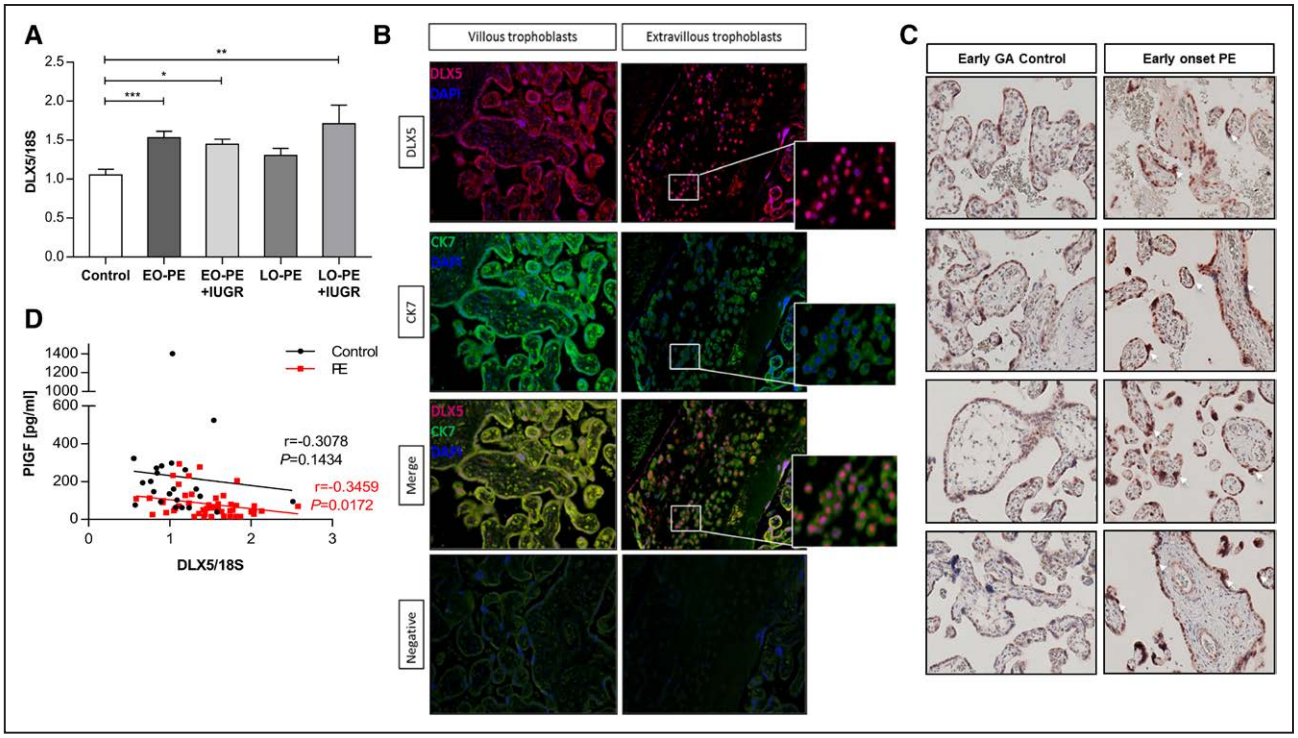


Figure 2. DLX5 is upregulated in preeclamptic placenta.
A, Quantitative polymerase chain reaction confirmed increased *DLX5* mRNA level in placentas of a second preeclamptic cohort (values are mean±SEM; control: 1.053±0.0745, n=28; early-onset (EO) preeclampsia+intrauterine growth restriction (IUGR): 1.447±0.066, n=9; late-onset (LO) preeclampsia+IUGR: 1.71±0.238, n=5; EO preeclampsia: 1.532±0.08, n=20; LO preeclampsia: 1.304±0.09, n=22). **P*<0.05, ***P*<0.005, ****P*<0.0005; ANOVA, Bonferroni multiple-comparisons test. **B**, Double immunofluorescence staining of term human placenta tissue indicated nuclear expression and colocalization of DLX5 to the cytokeratin 7 (CK7), a trophoblast-specific marker. DLX5 staining was positive in the nucleus of both villous and extravillous trophoblasts. **C**, Immunohistochemistry staining on placental villous tissue from healthy pregnancy (early control, gestation age [GA], 31 and 34 weeks) and preeclampsia (GA, 31 and 34 weeks) for human DLX5 confirmed increased DLX5 protein expression in early-onset preeclampsia (PE). **D**, Correlation of *DLX5* expression and placental biomarker. Placental *DLX5* expression significantly negatively correlated to the serum placental growth factor (PlGF) in the preeclampsia group but not in controls. Control, n=27, r=-0.3078, *P*=0.1434; preeclampsia, n=48, r=-0.3459, *P*=0.0172; Spearman rank correlation.

(19%) activation of the imprinted *DLX5* allele (Figure 3A and 3B). Sequencing of cDNA through the single nucleotide polymorphism (rs73708843) on 3 placenta samples confirmed the single allelic expression from *DLX5* (Figure 3C). We found a correlation (*r*=0.314, *P*=0.046) between LOI and *DLX5* expression, suggesting that the overexpression phenotype of *DLX5* was associated with its LOI (Figure 3D). We also inspected CpG methylation levels at the *DLX5* locus in preeclamptic placental compared with healthy control samples (20 versus 20)[28] and identified significant CpG hypomethylation in preeclampsia samples (Figure 3E and Table V in the online-only Data Supplement). Furthermore, the methylation level of several CpGs inversely correlated with *DLX5* expression in these samples (8 versus 8; Figure 3F). Collectively, we interpreted our data as the altered methylation at the *DLX5* locus in preeclampsia results in LOI and affects gene expression.

Upregulated *DLX5* Affects Genes Associated With Cell Growth, Proliferation, Survival, and Movement

To decipher the physiological effect of elevated DLX5 expression on trophoblasts, we stably overexpressed the human DLX5 protein in trophoblast cells in vitro. For the overexpression studies, we used the *Sleeping Beauty* transposon–derived expression system[29] in SGHPL-4 cells, derived from first trimester extravillous trophoblasts (Figure VA–VC in the online-only Data Supplement). Immunohistochemistry revealed a predominant nuclear localization of DLX5 in the DLX5-overexpressing SGHPL-4 cells (DLX-5^{High}; Figure VD in the online-only Data Supplement). To observe the global effects of elevated DLX5 expression, we performed microarray transcriptome profiling of DLX5^{High} and control cells. A total of 3650 DEGs (*P*<0.05; 771 genes at false discovery rate <0.05) were identified upon DLX5 overexpression (Figure VIA and Table VI in the online-only

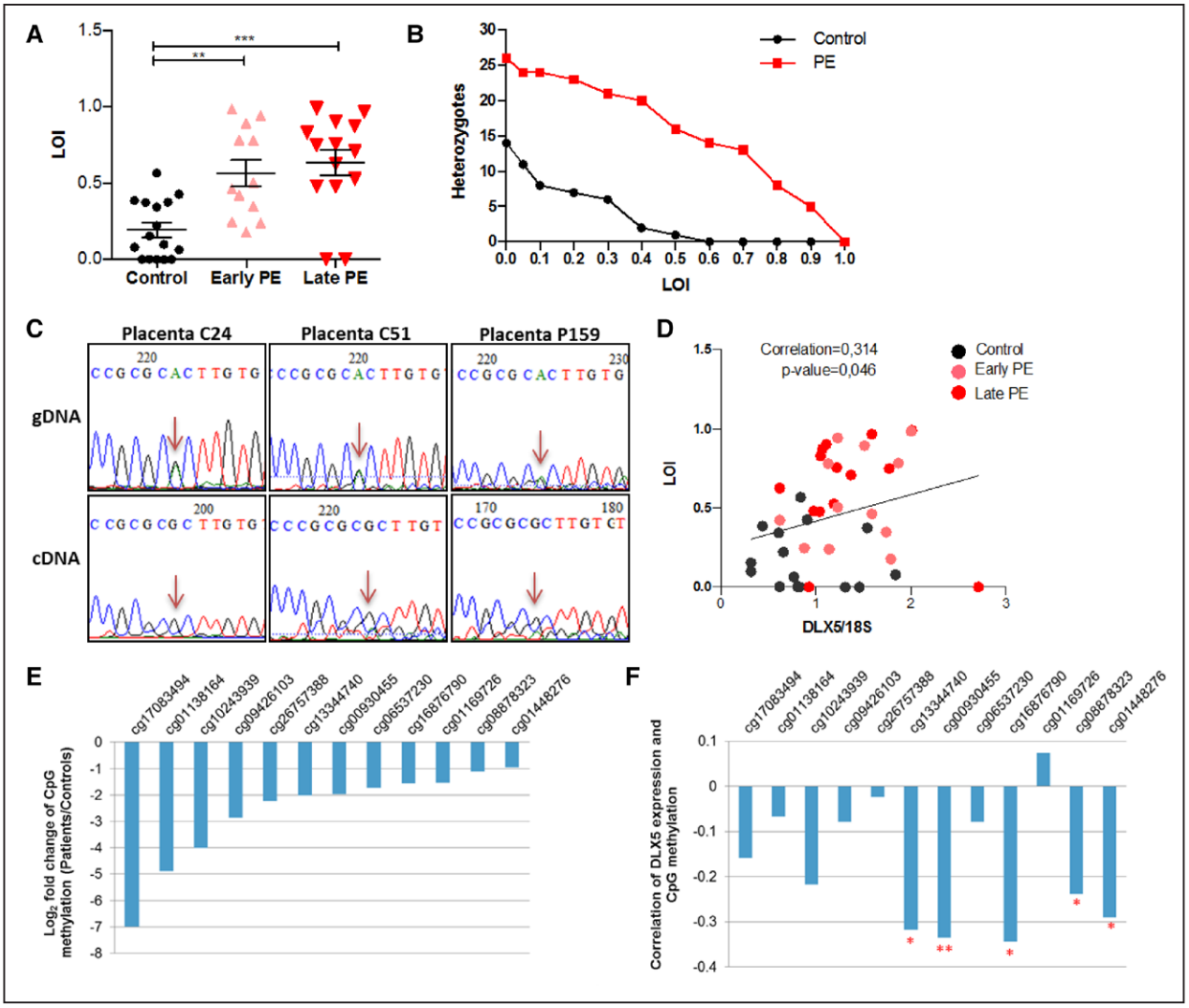


Figure 3. Loss of imprinting (LOI) of DLX5 in preeclamptic placenta.
A, Analysis of mean LOI levels for *DLX5* in healthy and preeclamptic placenta samples. Values are presented as a mean±SEM of LOI (control: 0.1943±0.04765, n=16; early-onset preeclampsia [PE]: 0.5661±0.08516, n=12; late-onset preeclampsia: 0.6349±0.08437, n=14). ***P*<0.001, ****P*<0.0001, 1-way ANOVA, Bonferroni multiple-comparisons test. **B**, Distribution (number) of DLX5 heterozygocities exceeding a particular LOI. **C**, Allelic expression analysis of the imprinted *DLX5* in placenta samples exhibiting the allele-specific expression but no LOI. cDNA of the heterozygous placenta samples for the single nucleotide polymorphism (rs73708843) were sequenced. **D**, LOI correlated with *DLX5* expression in placenta. *P*=0.046, Spearman rank correlation. **E**, CpG methylation of DLX5 locus. Log_2-fold change of CpG methylation level in preeclampsia (control, n=20; early-onset preeclampsia, n=20). Hypomethylated CpG sites are shown (differentially methylated region at adjusted *P*≤0.05). **F**, Pairwise Spearman rank correlation of CpG methylation and *DLX5* expression in placenta (control, n=8; early-onset preeclampsia, n=8; false discovery rate: **P*<0.05, ***P*<0.01). gDNA indicates genomic DNA.

Data Supplement). Ingenuity pathway analysis revealed significant gene enrichment involved in cardiovascular system development and function. The most significant terms describing molecular and cellular functions include cellular growth and proliferation, cell death and survival, and cellular movement and development (Figure VIB in the online-only Data Supplement). The top pathways include interferon and death receptor signaling, and superpathway of cholesterol biosynthesis. Several affected pathways are common between preeclamptic placental samples and the in vitro model. These include deregulated axon guidance, interleukin-8, and neuregulin receptor signaling; thyroid hormone receptor/retinoid X receptor, retinoic acid receptor, and planar cell polarity pathway; antigen presentation pathway; unfolded protein response; and nuclear factor, erythroid 2 like 2-mediated oxidative stress responses (Table VI in the online-only Data Supplement).

Elevated Expression of *DLX5* Models Certain Aspects of Preeclampsia

To test how well our $DLX5^{high}$ in vitro model mimics global transcriptional changes in preeclampsia, we compared transcriptome profiles of control and $DLX5^{high}$

cells with the 3 placental preeclamptic transcriptome clusters (PE_P1 through PE_P3). Hierarchical clustering of relative gene expression levels of control, DLX5^{high} lines (6 versus 6), PE_P1 through PE_P3, and healthy placenta samples revealed that the transcriptome of DLX5^{high} cells clusters with the 3 preeclamptic groups (PE_P1 through PE_P3), whereas control cells cluster with control placenta samples (Figure 4A). In addition, we asked if it was possible to correlate transcriptomes according to their *DLX5* expression levels (Figure VIIA in the online-only Data Supplement). Placenta samples were ordered according to their *DLX5* expression levels. The transcriptomes of DLX5^{high} preeclampsia clustered with cultured DLX5^{high} samples, whereas the low_DLX5

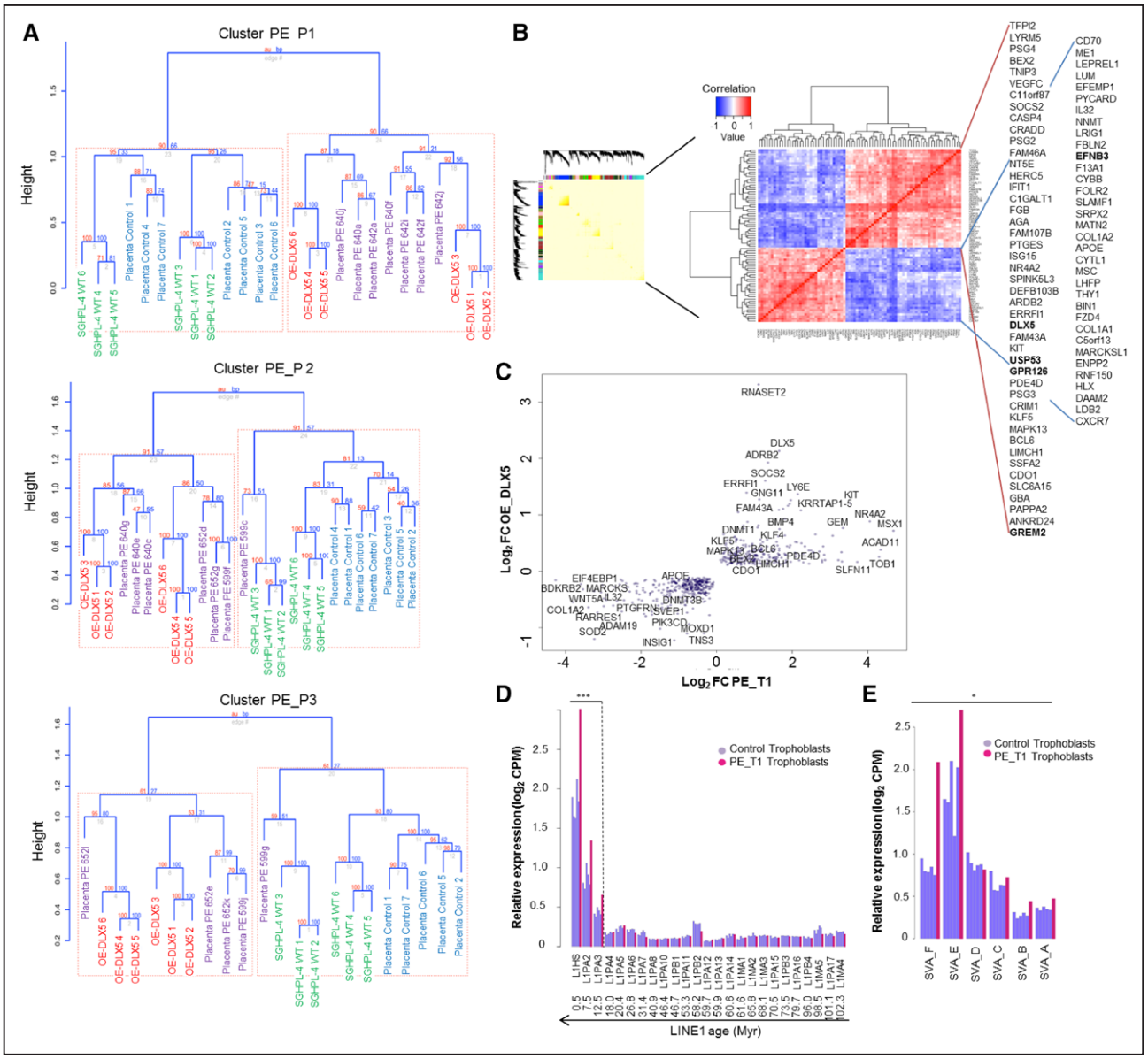


Figure 4. Intersection of preeclamptic transcriptomes with the DLX5^{high} transcriptome.
A, Hierarchical clustering (Spearman rank correlation, average linkage) and bootstrapping (1000 replicates) of the transcriptomes of SGHPL-4 cells with preeclamptic (PE) cluster PE_P1, PE_P2, and PE_P3 transcriptomes (Placenta PE) and control placenta samples (Placenta Control). The equal number of control placenta samples and samples in each preeclamptic cluster was chosen randomly. **B**, Weighted gene coexpression network analysis across 58 samples gave several modules containing a total of ≈3000 genes. Identification of DLX5 target genes in placenta and trophoblast cells. Clustered pairwise correlation matrix of identified 79 genes across 58 samples (Spearman rank correlation, threshold 0.6 and –0.55, $P<0.05$, euclidian distance). **C**, Comparison between the $\log_2$-fold change of the differentially expressed genes in trophoblast sample PE_T1 and genes differentially expressed on DLX5 overexpression in SGHPL-4 trophoblast cell line. Six hundred forty-one genes having the same differential expression pattern were common in both data sets. **D** and **E**, Relative expression of transposable LINE1 and SVA elements in control trophoblast samples and preeclamptic sample PE_T1. In PE_T1, the young members of the TE families L1PA3, L1PA2, L1HS, SVA-E, and SVA-F are upregulated. *$P<0.05$, ***$P<0.005$, Kolmogorov-Smirnov test, Benjamini and Hochberg false discovery rate. CMP indicates counts per million.

preeclamptic transcriptomes were clustered with control SGHPL-4 cells, suggesting that the overexpression of DLX5 in trophoblasts could model certain aspects of preeclampsia (Figure VIIB in the online-only Data Supplement).

Identifying Genes With Correlated Expression Dynamic to DLX5

Because the target genes of the transcription factor DLX5 in placenta are not known, we thought to identify genes with an expression that is correlated with *DLX5*. Thus, we subjected a merged data set of our 2 microarrays (total 58 samples) to weighted gene correlation network analysis. We aimed to identify gene modules of correlated gene expression. This approach allowed us to detect several gene modules containing a total of 3000 genes. Using pairwise ranked correlation analysis of 79 genes associated with *DLX5* across 58 samples (Figure 4B), we identified putative targets of DLX5. The list contains several genes previously associated with preeclampsia,[30–35] including genes involved in cell growth, proliferation, and differentiation (eg, *GREM2*, *KIT*, *ERRFI1*), angiogenesis (eg, *VEGFC*, *PAPPA2, GPR126*), cytokine and growth hormone signaling (*GBA, CXCR7*), immune response (*ISG15*, *HERC5*, *IFIT1*), pregnancy-specific proteins (*PSG2-4*), and the paternally imprinted tissue factor pathway inhibitor-2 (*TFPI2*), involved in regulation of cell invasion and proliferation.

Upregulation of *DLX5* in Trophoblasts Is Associated With Disturbed Epigenetics

To address the limitations of our initial approach using placental microarray data, we performed transcriptome analysis using RNA sequencing on freshly isolated, purified human trophoblasts from control and preeclamptic placentas (5 versus 5). In the PE_T1 trophoblast sample, *DLX5* was highly upregulated, and we observed 1466 genes commonly dysregulated in both PE_T1 and DLX5^{high}, with 641 genes exhibiting the same pattern of expression, including *KIT*, *SOCS2*, *KLF5*, *BEX2*, *ERRFI1*, and DNA methyltransferases (*DNMT1* and *DNMT3B*; Figure 4C). In addition to common features, the PE_T1 sample exhibited further DEGs involved in the regulation of DNA methylation and histone modification such as the TET gene family (*TET1-3*), *SETDB1*, *SIRT1*, and histone deacetylases, indicating that a subset of preeclampsia might be associated with severe epigenetic disturbances. As a likely consequence, the PE_T1 sample is characterized by the deregulation of several imprinted genes (including *DLX5*) but also potentially mutagenic transposable elements such as LINE-1 and SVA[36,37] (Figure 4D and 4E).

Upregulation of *DLX5* Reduces Trophoblast Proliferation

The significant increase of DLX5 expression in preeclamptic placenta samples prompted us to explore the possible mechanism of DLX5 in the pathogenesis of preeclampsia. To characterize the DLX5^{high} phenotype, we have performed several cellular assays inspired by the Pathway analyses. To determine whether the DLX5^{high} phenotype affects trophoblast proliferation, we used time-lapse microscopy and observed reduced cell proliferation of DLX5^{high} cells by 45% compared to control after 48 hours of incubation (Figure 5A). Significantly, reduced trophoblast proliferation of DLX5^{high} was confirmed by a high-throughput sampler cell count and microtiter plate test colorimetric assay (Figure 5B and 5C). DLX5 overexpression had only a slight, not significant, effect on cell apoptosis as indicated by scoring apoptotic cells within 48 hours of incubation (Figure VIIIA and VIIIB in the online-only Data Supplement).

Elevated *DLX5* Expression Affects the Metabolic Profile of the Trophoblast

Cell proliferation, growth, and metabolism are tightly linked processes. To determine whether reduced proliferation was associated with altered metabolism, we monitored metabolic parameters in DLX5^{high} cells. We determined the extracellular acidification rate and oxygen consumption rate, indicators of mitochondrial respiration and glycolytic activity, respectively (Figure 5D–5F). Compared with control, metabolic profiling detected elevated level of extracellular acidification rate and maximal oxygen consumption rate values, suggesting an accelerated metabolism of DLX5^{high} cells (Figure 5E and 5F). Furthermore, DLX5^{high} cells displayed increased spare respiratory capacity values compared with control cells (Figure 5E). In principle, the increased energetic demand could reflect a response to increased stress or cell survival challenges.

DLX5 Expression Responds to ER Stress

Abnormal placentation in preeclampsia results in a series of biological stresses. To investigate the potential role of DLX5 in stress response, we monitored reactive oxidative species production and the effect of induced ER stress. Although we did not detect elevated reactive oxidative species production in DLX5^{high} cells (Figure 5G), DLX5 expression was sensitive to induced ER stress in BeWo choriocarcinoma cells, expressing DLX5 at a readily detectable level. In a hypoxia-reoxygenation challenge assay, DLX5 expression increased significantly on ER stress in a severity-dependent manner (Figure 5H and Figure IXA in the online-only Data Supplement). Furthermore, in our DLX5^{high} tran-

Circulation. 2017;136:1824–1839. DOI: 10.1161/CIRCULATIONAHA.117.028110

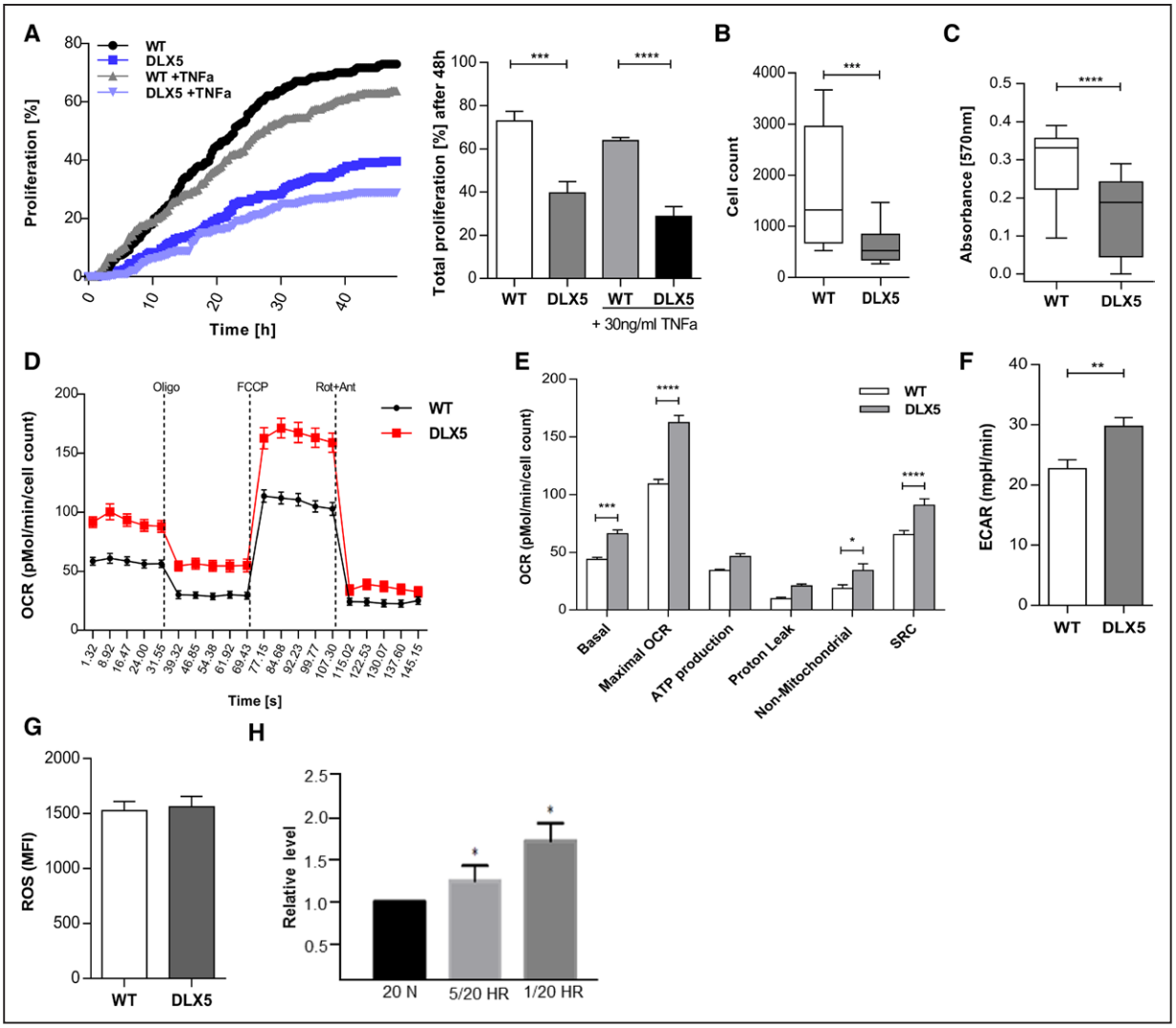


Figure 5. DLX5 decreases SGHPL-4 cell proliferation.

A, DLX5^{high} cells are less proliferative compared with wild-type (WT) cells as indicated by scoring dividing cells over 48 hours of incubation. After 48 hours of incubation, cell proliferation of DLX5^{high} cells (39.58±5.34) is reduced by 45% compared with WT (72.92±4.49). ***P=0.001, ****P<0.0001, 2-way ANOVA, Bonferroni multiple-comparisons test. Tumor necrosis factor-α (TNFa) at a concentration of 30 ng/mL slightly decreased cell proliferation in both DLX5^{high} and WT cells. P=NS. **B**, DLX5^{high} cells exhibited decreased cell proliferation (n=6; median, 524.5; interquartile range [IQR], 343.8–843) compared with WT (n=6; median, 1320; IQR, 679.8–2955) confirmed by cell count high-throughput sampler assay. ***P<0.001, Mann-Whitney test. **C**, Microtiter plate test viability assay confirmed decreased cell proliferation in DLX5^{high} cells (DLX5^{high}: median, 0.1883; IQR, 0.0454–0.242 vs. WT: median, 0.3319; IQR, 0.2226–0.3565). ****P<0.0001, Mann-Whitney test. **D**, Effect of DLX5 on mitochondrial respiration in SGHPL-4 cells. Oxygen consumption rate (OCR) was measured under basal conditions followed by the sequential addition of oligomycin (Oligo; 0.75 μmol/L), FCCP (Carbonyl cyanide-4-[trifluoromethoxy]phenylhydrazone) (1 μmol/L), and antimycin A (Ant 1 μmol/L) plus rotenone (Rot; 0.1 μmol/L) in WT (n=6) and DLX5^{high} cells (n=6). Data are normalized to the cell number. **E**, Individual parameters for basal respiration (WT vs. DLX5^{high}, 43.901±1.705 vs. 66.082±3.213), maximal respiration (WT vs. DLX5^{high}, 109.226±3.913 vs. 162.244±6.431), adenosine triphosphate (ATP) production (WT vs. DLX5^{high}, 34.266±1.105 vs. 46.415±2.2), proton leak (WT vs. DLX5^{high}, 9.635±1.307 vs. 20.85±1.687), nonmitochondrial OCR (WT vs. DLX5^{high}, 18.72±2.989 vs. 34.247±5.773), and reserve capacity (WT vs. DLX5^{high}, 65.325±3.381 vs. 90.866±5.455) were extracted from the assay. *P<0.05, ***P<0.001, ****P<0.0001, 2-way ANOVA, Bonferroni multiple-comparisons test. SRC indicates spare respiratory capacity. **F**, Mean basal extracellular acidification rate (ECAR) level in WT and DLX5^{high} cells (DLX5^{high}: median, 26.27; IQR, 24.27–36.35 vs. WT: median, 22.38; IQR, 17.05–26.99. **P=0.01, Mann-Whitney test. **G**, Reactive oxidative species (ROS) production in WT and DLX5^{high} cells measured by fluorescent cell sorting. Data are presented as a mean fluorescent intensity (MFI) of the fluorescent signal from the dichlorodihydrofluorescein (DCF) oxidized by ROS. DCFH-DA diffuses into the cell and becomes deacetylated by cellular esterases to nonfluorescent 2′, 7′-Dichlorodihydrofluorescin (DCFH), which is next oxidized to fluorescent DCF by ROS. **H**, DLX5 expression level on induction of endoplasmic reticulum (ER) stress in BeWo cells. Quantification of DLX5 level on induction of ER stress indicates significant upregulation of its expression. 5/20 hypoxia-reoxygenation (HR) and 1/20 HR indicate cyclic condition of 6 hours of incubation in 5%/20% O_2 and 1%/20% O_2; and 20 N, normoxia. *P<0.01, Kruskal-Wallis test, Dunn multiple-comparisons test.

scriptome, we observed upregulation of several genes involved in the unfolded protein response pathway that were associated with ER stress response (Figure IXB in the online-only Data Supplement). Eight of these genes (*INSIG1*, *SREBF1*, *HSP90B1*, *ATF6*, *MBTPS1*, *PPP1R15A*, *XBP1*, and *HSPA2*) were also dysregulated in our preeclamptic placenta samples (Figure IXC in the online-only Data Supplement). Although enhanced DLX5 level appears to trigger the cellular stress response, DLX5 expression increased with syncytium formation as evidenced by Forskolin treatment of BeWo cells, suggesting a potential role of DLX5 during the syncytialization process (Figure IXD in the online-only Data Supplement).

Dlx5 Expression Shifts From Postimplantation to Preimplantation Stage of Embryogenesis During Evolution

DLX5 is known primarily as a transcription factor regulating morphogenesis and tissue homeostasis,[38–40] and it is mostly characterized during postimplantation embryogenesis. Because we observed *DLX5* is expressed in trophoblasts, we sought to monitor its expression pattern in preimplantation embryos. We performed comparative single-cell RNA sequencing data analysis on mouse, macaque, and human embryos collected at early embryonic developmental stages.[41–43] In human embryos, the expression of *DLX5* appears at the transition from embryonic day 4 to 5 stage (Figure 6A), rendering *DLX5* one of the earliest expressed genes in the human trophectoderm (Figure 6B). *Dlx5* expression is shifted toward late trophectoderm in macaque embryos (Figure 6C) and is not detectable in murine preimplantation embryos (Figure 6A). To check whether *Dlx5* is expressed at later stages of placenta development in mice, we performed placental immunostaining (embryonic day 14.5 and 15.5) in Dlx5-LacZ$^{+/-}$ animals. Although we observed a weak positive LacZ staining in the Dlx5-LacZ$^{+/-}$ animals on the external muscular layer of the placenta, the signal was not significantly different from the control (Figure X in the online-only Data Supplement), suggesting that *Dlx5* is not involved in murine placentation. Surprisingly, only 33 common trophectoderm marker genes could be identified between mice and humans. In humans, among the early trophectoderm markers, *DLX5* exhibits the highest activation of expression, followed by *RGS13*, *NDRG2*, *ODAM*, and *SLC38A1*, as well as *ID3*, *HAND1*, *DLX3*, and *GCM1* (Figure 6D). Among the genes expressed differentially in human versus mouse preimplantation embryos, *GREM2*, *GPR126*, *USP53*, and *EFNB3* are putative targets of DLX5 (Figure 5B). *GREM2*, *GPR126*, and *USP53* are also upregulated in preeclampsia (Figure 1B), suggesting that the dysregulation of maternally expressed DLX5 and its putative targets might explain certain features of the human-specific nature of preeclampsia. The genes expressed in the same clusters might share transcriptional networks (Figure 6E and 6F and Table VII in the online-only Data Supplement). MEGs but not PEGs form characteristic clusters during human embryogenesis.

DISCUSSION

There certainly has been much speculation that disturbed regulation of imprinted genes might be involved in the development of preeclampsia; however, prior evidence could not establish a significant association between them.[14–19] Here, we provide robust evidence for a mechanistic coupling between preeclampsia and disturbed placental imprinting. Our experimental strategy first aimed at identifying DEGs in preeclamptic placentas by analyzing genome-wide molecular data on well-characterized patient material. The list of DEGs was then intersected with the current catalog of human imprinted placental genes. Using the novel set of genes could clarify certain important issues concerning the long-term debated list of imprinted genes in the human placenta. Our strategy revealed several potential candidates, supporting the hypothesis that disturbed imprinting and preeclampsia could indeed be associated. Our candidate list included imprinted genes that were previously associated with preeclampsia, but their expression deregulation could not be convincingly connected to LOI.[15] *CYP2J2* and *CD74* belonged to a category of genes whose deregulation was already implicated in preeclampsia, but their epigenetic disturbance was not considered as a contributing factor.[21,22] Our strategy also identified genes that were not yet implicated in preeclampsia, and their imprinted status is poorly explored in placenta (APOBEC2, *GATA3*, *DLX5)*. APOBEC2, an enzyme involved in controlling DNA-based parasites such as viruses and transposable elements, appeared on the list of MEGs deregulated in all 3 clusters of preeclampsia. The most significantly affected imprinted genes in preeclampsia were *GATA3* and *DLX5*. *GATA3* could be an excellent candidate for further research because it was previously reported to inhibit trophoblast invasion,[44] thought to be a key process in preeclampsia. Here, we focus on *DLX5*, a transcription factor of the Distal-less homeobox protein family. *DLX5* is involved in developmental processes of the limb, brain, and bone in both mice and humans.[38,45] However, its placental function is not characterized.

We show that *DLX5* is expressed in human villous and extravillous trophoblast and is controlled by imprinting. *DLX5* is upregulated in ≈70% of patients with preeclampsia. The upregulation of *DLX5* in preeclampsia is associated with its leaky expression from the imprinted

Circulation. 2017;136:1824–1839. DOI: 10.1161/CIRCULATIONAHA.117.028110

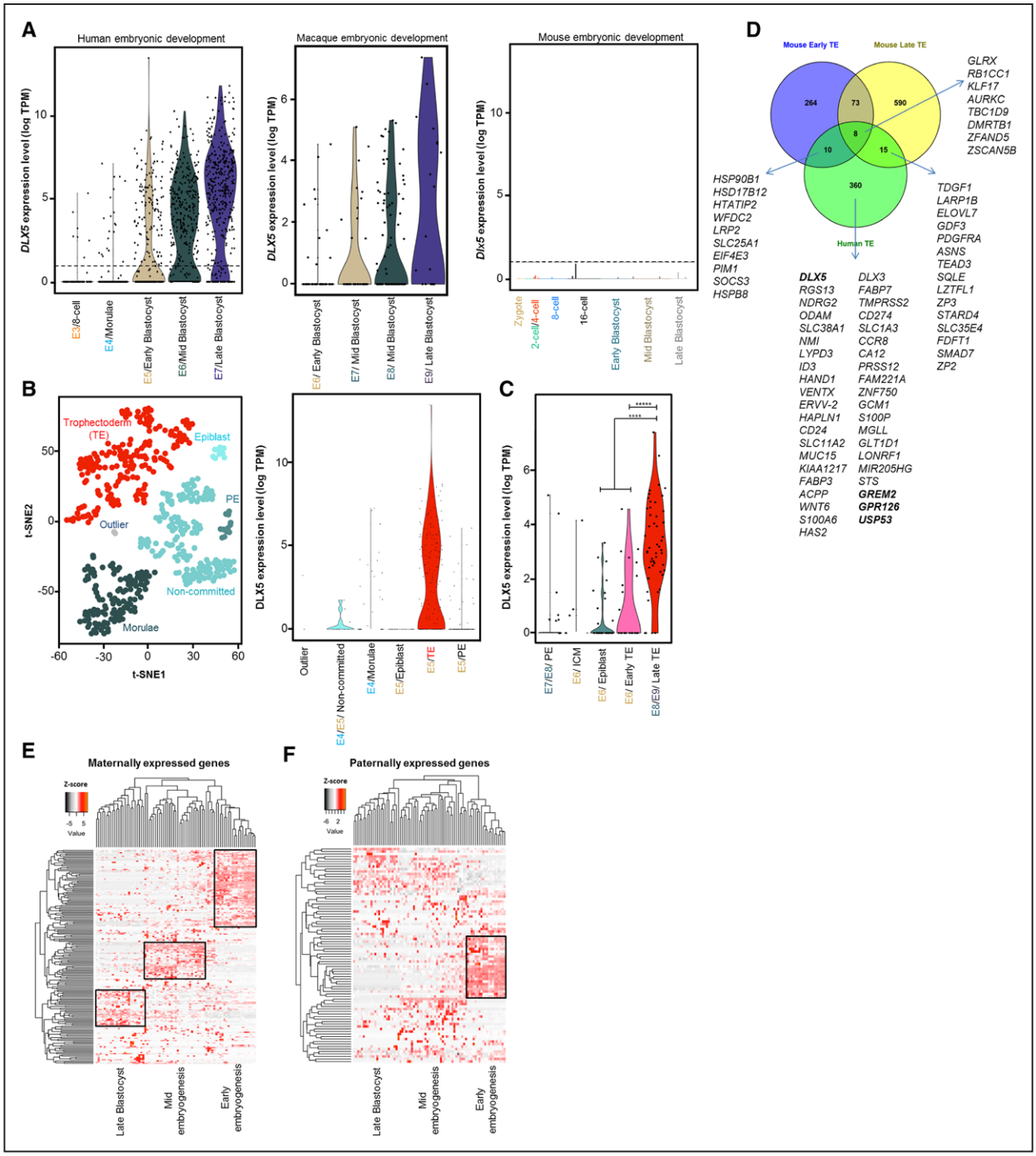


Figure 6. DLX5 expression in human, macaque, and mouse early embryonic development.
A, Violin plots display the expression levels of *DLX5* during different stages of human, macaque, and mouse early embryonic development. In humans, *DLX5* starts to be expressed at embryonic day (E) E4/E5 stage of development. In macaques, *DLX5* is expressed at E6 stage. *Dlx5* is not expressed in mouse preimplantation embryo. **B**, Single-cell transcriptome analysis of human preimplantation embryo reveals trophectoderm (TE)-specific *DLX5* expression. t-SNE analysis on human 353 single cells from E4 to E5 stages when inner cell mass (ICM) and trophectoderm split from Morulae. We defined clusters of cell populations to identify genes expressing exclusively in those clusters. We defined the cell type for each cluster according to the known markers. **C**, Single-cell transcriptome analysis of macaque preimplantation embryo demonstrates enriched *DLX5* expression in the late trophectoderm. *****P=2.602E^{-12}, ****P=1.387E^{-06}, Wilcoxon test. **D**, Comparative analysis between human and mouse trophectoderm markers. For human trophectoderm markers, genes with $\log_2$-fold change >3 are shown. Single-cell transcriptome analysis of human embryogenesis for maternally **(E)** and paternally **(F)** expressed genes. Clustering analysis of imprinted genes expressed during early (oocyte, zygote, 2-cell, and 4-cell stage) and mid (8-stage, Morulae) embryogenesis and late blastocyst (179 of 257 maternally expressed genes, 86 of 150 paternally expressed genes). PE indicates primitive endoderm; TPM, transcripts per million; and t-SNE, t-distributed stochastic neighbor embedding.

allele (LOI). In contrast to previous studies,[27,46] we find a correlation between expression of an imprinted gene and its LOI in preeclampsia. Our data mining[28,47] also reveals differential CpG methylation of the *DLX5* locus in preeclamptic placentas. Although our analysis does not rule out other mechanisms of DLX5 expression regulation such as transcriptional or microRNA regulation, we provide evidence of an association between disturbed imprinting gene expression and preeclampsia.

The spatial and temporal regulation of cell proliferation and differentiation is crucial for successful pregnancy. The first half of gestation is characterized by a series of trophoblast proliferation and differentiation processes, building mature villi and extravillous structures.[48,49] Upregulation of DLX5 resulted in reduced (≈45%) proliferation of the trophoblast. The decreased trophoblast cell proliferation was accompanied by increased oxygen consumption. Why might poorly proliferating trophoblast cells require an accelerated metabolism? We hypothesized that DLX5 overexpression could sensitize trophoblasts to stress. As a result, the cells require increased metabolic activity to overcome this state of affairs. Indeed, the transcriptome analysis of DLX5^{high} trophoblasts revealed several affected pathways acting as stress inducers, such as unfolded protein response pathways and increased interferon and death receptor signaling. Thus, the upregulation of *DLX5* could be a factor contributing to an accelerated placenta aging process and elevated ER stress, resulting in stressed syncytiotrophoblast and consequently increased shedding of inflammatory factors into the maternal circulation.[50,51] Although an enhanced DLX5 level triggers the cellular stress response, DLX5 expression increases with syncytium formation, suggesting a role of DLX5 in regulating the syncytialization. Our single-cell transcriptome data mining of human preimplantation embryos establishes DLX5 as a key marker of trophectoderm differentiation. We propose that DLX5 is involved in regulating a delicate balance between the proliferation and differentiation processes of the trophoblast. A disturbance of this key process has been previously associated with preeclampsia.[52,53]

In contrast with trophoblasts, *DLX5* overexpression is associated with enhanced cell proliferation in various cancer cells.[54–56] Thus, *DLX5* might affect proliferation either negatively or positively during early development or in cancer, respectively. The response to overexpressed DLX5 possibly depends on cell type–specific target genes. Either way, DLX5 appears to be a key gene in determining the developmental decisions of trophoblast cells.

We modeled the effect of DLX5 upregulation in an in vitro system, overexpressing DLX5 in SGHPL-4 trophoblast cells (DLX5^{high}). Because artificial, exogenous overexpression of a gene in cells could alter normal cellular function as a result of the accumulation of unprocessed proteins, we asked how faithfully DLX5^{high} cells mimic preeclampsia. The transcriptome of DLX5^{high} cells resembled that of the preeclamptic transcriptomes, and several dysregulated pathways were commonly seen both in vitro (DLX5^{high}) and in vivo (preeclampsia samples). Thus, the DLX5^{high} phenotype can model several features of preeclampsia in vitro in a cell culture system, signifying the impact of the deregulated DLX5 in the pathogenic phenotype of preeclampsia.

Given the barriers in analysis of preeclampsia resulting from its human specificity, our in vitro model system has considerable potential for downstream analyses. Nevertheless, our study do not suggest that there could be a single explanation for preeclampsia. We identified 3 distinct transcriptomic clusters of preeclamptic placenta (PE_P1 through PE_P3). That the clusters could be related to previously established categories of preeclampsia such as early and late onset of preeclampsia[57] supports the view that preeclampsia is a heterogeneous placental disease and indeed comes in several discrete forms. Although PE_P2 matches late-onset preeclamptic placentas, PE_P1 and PE_P3 can be considered subdivisions of early-onset preeclampsia. The uncovered heterogenic nature of preeclampsia would call for validating a panel of subclass-specific biomarkers for future diagnostic procedures. The *DLX5* overexpression phenotype is detectable in all the 3 clusters but is most pronounced in PE_P1 and PE_P2. Levels of *DLX5* correlated with the placenta-derived PlGF circulating biomarker. Whether DLX5 will have utility as a biomarker is unclear because its LOI was not observed in all instances (69% of preeclampsia cases). Although the PE_P2 and PE_P3 clusters have no clear clinical pattern, patients in a PE_P1 cluster exhibit characteristic clinical phenotypes. Nevertheless, more samples need to be analyzed to securely relate the characteristics of the PE_P1 cluster to clinical disease phenotyping.

Our RNA sequencing data analysis revealed that a subset of preeclampsia is connected to disturbed epigenetic gene regulation. The global epigenetic turmoil is likely associated with the observed differential expression of genes regulating DNA methylation, resulting in the deregulation of transposable retroelements and imprinted genes. The mechanisms of regulating imprinting and repressing retroelements by DNA methylation share several common features.[58] In fact, genomic imprinting is speculated to be a byproduct of the defense mechanisms of the genome against retroviruses and retroelements.[59,60] A domesticated retroelement-derived gene (syncytin-1), implicated to have a key role in placental development,[61–63] has been associated with preeclampsia.[64,65] Here, in a subset of preeclampsia, we observed the reactivation of the human-specific L1_HS and SVA-F elements, capable of transposition in the human genome.[37,66]

Besides trophoblasts, dysregulation of DLX5 expression in other tissues has been reported to contribute to diseases. Downregulation of DLX5 in endometrial glands could also complicate early stages of pregnancy that could later manifest in intrauterine growth restriction or preeclampsia.[67] LOI of the maternally expressed *DLX5* in lymphoblastoid cells contributes to Rett syndrome, a disorder associated with GABAergic dysfunction.[68,69] Dysregulation of DLX5 impairs the differentiation of GABAergic neurons.[70] GABA can increase human chorionic gonadotrophin secretion in human placenta,[71] indicating possible placenta-brain endocrine interactions regulated by imprinting. In contrast with Rett syndrome, we do not detect differential expression of DLX6 in preeclampsia, suggesting that the dysregulation of DLX5 in preeclampsia is not associated with expressional changes of DLX6 (not regulated by imprinting). We propose that DLX5 and DLX6 are regulated differently in brain and placenta.

DLX5 is expressed in human but not mouse trophoblast. Comparative single-cell transcriptome analysis revealed a differential expression of DLX5 between human, macaque, and mouse preimplantation embryogenesis, highlighting the diverged cellular function of DLX5 during mammalian embryonic evolution. Although the *DLX5* gene is highly conserved across different mammalian species (>95% exon sequence similarity in human, macaque, and mouse), its upstream 2- to 10-kb (potential regulatory) region is much faster evolving (>80% and <10% sequence similarity for human versus macaque and for human versus mouse, respectively). Indeed, despite the conserved coding structure, *DLX5* expression appears to be gradually shifted toward earlier developmental stages during mammalian evolution. Although DLX5 is not expressed in preimplantation embryos in mice, its expression peaks at the stage of trophectoderm and inner cell mass/epiblast separation, marks trophectoderm-committed cells, and is regulated by imprinting in humans. In addition to *DLX5*, we could identify further DEGs between human and mice trophectoderm. Among these genes, *GREM2*, *GPR126*, *USP53*, and *EFNB3* are also putative targets of *DLX5,* suggesting that the function of a *DLX5*-regulated circuitry has been redefined during mammalian evolution. *GREM2*, *GPR126*, and *USP53* are also upregulated in preeclampsia and thus might be associated with the human-specific nature of preeclampsia.

ACKNOWLEDGMENTS

The authors thank Juliane Anders and Ute Gerhard for their excellent technical assistance. They also thank technician Lise Øhra Levy, postdoctoral researcher Meryam Sugulle, and Oslo University Hospital PhD students for valuable assistance in patient recruitment and biobank handling at Oslo University Hospital, Norway.

SOURCES OF FUNDING

The Deutsche Forschungsgemeinschaft supported Drs Herse (HE 6249/1-2, HE 6249/4-1), Dechend (DE 631/9-1), and Müller. The German Center for Cardiovascular Research supported Dr Müller. This study was supported by the Wellcome Trust (grant 084804/2/08/Z). The Oslo Pregnancy Biobank work was supported by grants from the South-Eastern Norway Regional Health Authority, Norway, and from Oslo University Hospital. Dr Izsvák is funded by European Research Council Advanced (ERC-2011-AdG 294742).

DISCLOSURES

None.

AFFILIATIONS

From Max Delbrück Center for Molecular Medicine in the Helmholtz Association, Berlin, Germany (J.Z., M.S., F.H., N.H., D.N.M., Z.I.); Experimental and Clinical Research Center, a joint cooperation between the Max-Delbrück Center for Molecular Medicine in the Helmholtz Association and the Charité-Universitätsmedizin Berlin, Germany (J.Z., F.H., L.P., N.H., M.G., H.S., F.C.L., D.N.M., R.D.); Berlin Institute of Health, Germany (J.Z., F.H., L.P., N.H., M.G., F.C.L., D.N.M., R.D., Z.I.); Department of Obstetrics and Department of Gynecology, Charité-Universitätsmedizin Berlin, Germany (M.G.); German Centre for Cardiovascular Research, partner site Berlin, Germany (N.H., D.N.M.); Centre for Trophoblast Research, University of Cambridge, UK (H.W.Y.); Institute of Cell Biology, Histology and Embryology, Medical University of Graz, Austria (B.H.); Molecular and Clinical Sciences Research Institute, St George's University of London, UK (J.E.C., G.W.); Division of Obstetrics and Gynaecology, Oslo University Hospital, Norway (G.M.J., A.C.S.); University of Oslo, Norway (G.M.J., A.C.S.); Évolution des Régulations Endocriniennes, Muséum Nationale d'Histoire Naturelle, Paris, France (G.L.); HELIOS-Klinikum, Berlin, Germany (A.I., R.D.); Cologne Center for Genomics, University of Cologne, Germany (H.S.); and Milner Centre for Evolution, Department of Biology and Biochemistry, University of Bath, UK (L.D.H.).

FOOTNOTES

Received February 27, 2017; accepted August 28, 2017.

The online-only Data Supplement is available with this article at http://circ.ahajournals.org/lookup/suppl/doi:10.1161/CIRCULATIONAHA.117.028110/-/DC1.

Circulation is available at http://circ.ahajournals.org.

REFERENCES

1. ACOG. Hypertension in pregnancy: report of the American College of Obstetricians and Gynecologists' Task Force on Hypertension in Pregnancy. *Obstet Gynecol*. 2013;122:1122–1131.
2. Mosca L, Benjamin EJ, Berra K, Bezanson JL, Dolor RJ, Lloyd-Jones DM, Newby LK, Piña IL, Roger VL, Shaw LJ, Zhao D, Beckie TM, Bushnell C, D'Armiento J, Kris-Etherton PM, Fang J, Ganiats TG, Gomes AS, Gracia

CR, Haan CK, Jackson EA, Judelson DR, Kelepouris E, Lavie CJ, Moore A, Nussmeier NA, Ofili E, Oparil S, Ouyang P, Pinn VW, Sherif K, Smith SC Jr, Sopko G, Chandra-Strobos N, Urbina EM, Vaccarino V, Wenger NK. Effectiveness-based guidelines for the prevention of cardiovascular disease in women–2011 update: a guideline from the American Heart Association. *Circulation*. 2011;123:1243–1262. doi: 10.1161/CIR.0b013e31820faaf8.
3. Steegers EA, von Dadelszen P, Duvekot JJ, Pijnenborg R. Pre-eclampsia. *Lancet*. 2010;376:631–644. doi: 10.1016/S0140-6736(10)60279-6.
4. Redman CW, Sargent IL. Immunology of pre-eclampsia. *Am J Reprod Immunol*. 2010;63:534–543. doi: 10.1111/j.1600-0897.2010.00831.x.
5. Roberts JM, Escudero C. The placenta in preeclampsia. *Pregnancy Hypertens*. 2012;2:72–83. doi: 10.1016/j.preghy.2012.01.001.
6. Reik W, Walter J. Genomic imprinting: parental influence on the genome. *Nat Rev Genet*. 2001;2:21–32. doi: 10.1038/35047554.
7. Graves JA. Genomic imprinting, development and disease: is pre-eclampsia caused by a maternally imprinted gene? *Reprod Fertil Dev*. 1998;10:23–29.
8. Haig D. Genetic conflicts in human pregnancy. *Q Rev Biol*. 1993;68:495–532.
9. Haig D. Genomic imprinting and kinship: how good is the evidence? *Annu Rev Genet*. 2004;38:553–585. doi: 10.1146/annurev.genet.37.110801.142741.
10. Hollegaard B, Byars SG, Lykke J, Boomsma JJ. Parent-offspring conflict and the persistence of pregnancy-induced hypertension in modern humans. *PLoS One*. 2013;8:e56821. doi: 10.1371/journal.pone.0056821.
11. Haig D. The kinship theory of genomic imprinting. *Ann Rev Ecol Systematics*. 2000;31:9–32.
12. Robillard PY, Hulsey TC, Dekker GA, Chaouat G. Preeclampsia and human reproduction: an essay of a long term reflection. *J Reprod Immunol*. 2003;59:93–100.
13. Wildman DE, Chen C, Erez O, Grossman LI, Goodman M, Romero R. Evolution of the mammalian placenta revealed by phylogenetic analysis. *Proc Natl Acad Sci USA*. 2006;103:3203–3208. doi: 10.1073/pnas.0511344103.
14. Williams PJ, Broughton Pipkin F. The genetics of pre-eclampsia and other hypertensive disorders of pregnancy. *Best Pract Res Clin Obstet Gynaecol*. 2011;25:405–417. doi: 10.1016/j.bpobgyn.2011.02.007.
15. Yu L, Chen M, Zhao D, Yi P, Lu L, Han J, Zheng X, Zhou Y, Li L. The H19 gene imprinting in normal pregnancy and pre-eclampsia. *Placenta*. 2009;30:443–447. doi: 10.1016/j.placenta.2009.02.011.
16. Huang GQ, Hu YY, Wang XD. Placental PHLDA2 gene imprinting in patients with pre-eclampsia [in Chinese]. *Sichuan Da Xue Xue Bao Yi Xue Ban*. 2015;46:104–107, 128.
17. Berends AL, Bertoli-Avella AM, de Groot CJ, van Duijn CM, Oostra BA, Steegers EA. STOX1 gene in pre-eclampsia and intrauterine growth restriction. *BJOG*. 2007;114:1163–1167. doi: 10.1111/j.1471-0528.2007.01414.x.
18. Iglesias-Platas I, Monk D, Jebbink J, Buimer M, Boer K, van der Post J, Hills F, Apostolidou S, Ris-Stalpers C, Stanier P, Moore GE. STOX1 is not imprinted and is not likely to be involved in preeclampsia. *Nat Genet*. 2007;39:279–280; author reply 280. doi: 10.1038/ng0307-279.
19. Kivinen K, Peterson H, Hiltunen L, Laivuori H, Heino S, Tiala I, Knuutila S, Rasi V, Kere J. Evaluation of STOX1 as a preeclampsia candidate gene in a population-wide sample. *Eur J Hum Genet*. 2007;15:494–497. doi: 10.1038/sj.ejhg.5201788.
20. Hamada H, Okae H, Toh H, Chiba H, Hiura H, Shirane K, Sato T, Suyama M, Yaegashi N, Sasaki H, Arima T. Allele-specific methylome and transcriptome analysis reveals widespread imprinting in the human placenta. *Am J Hum Genet*. 2016;99:1045–1058. doi: 10.1016/j.ajhg.2016.08.021.
21. Herse F, Lamarca B, Hubel CA, Kaartokallio T, Lokki AI, Ekholm E, Laivuori H, Gauster M, Huppertz B, Sugulle M, Ryan MJ, Novotny S, Brewer J, Park JK, Kacik M, Hoyer J, Verlohren S, Wallukat G, Rothe M, Luft FC, Muller DN, Schunck WH, Staff AC, Dechend R. Cytochrome P450 subfamily 2J polypeptide 2 expression and circulating epoxyeicosatrienoic metabolites in preeclampsia. *Circulation*. 2012;126:2990–2999. doi: 10.1161/CIRCULATIONAHA.112.127340.
22. Przybyl L, Haase N, Golic M, Rugor J, Solano ME, Arck PC, Gauster M, Huppertz B, Emontzpohl C, Stoppe C, Bernhagen J, Leng L, Bucala R, Schulz H, Heuser A, Weedon-Fekjær MS, Johnsen GM, Peetz D, Luft FC, Staff AC, Müller DN, Dechend R, Herse F. CD74-downregulation of placental macrophage-trophoblastic interactions in preeclampsia. *Circ Res*. 2016;119:55–68. doi: 10.1161/CIRCRESAHA.116.308304.
23. Carmeliet P, Tessier-Lavigne M. Common mechanisms of nerve and blood vessel wiring. *Nature*. 2005;436:193–200. doi: 10.1038/nature03875.
24. Liao WX, Laurent LC, Agent S, Hodges J, Chen DB. Human placental expression of SLIT/ROBO signaling cues: effects of preeclampsia and hypoxia. *Biol Reprod*. 2012;86:111. doi: 10.1095/biolreprod.110.088138.
25. Okita C, Meguro M, Hoshiya H, Haruta M, Sakamoto YK, Oshimura M. A new imprinted cluster on the human chromosome 7q21-q31, identified by human-mouse monochromosomal hybrids. *Genomics*. 2003;81:556–559.
26. Stepan H, Herraiz I, Schlembach D, Verlohren S, Brennecke S, Chantraine F, Klein E, Lapaire O, Llurba E, Ramoni A, Vatish M, Wertaschnigg D, Galindo A. Implementation of the sFlt-1/PlGF ratio for prediction and diagnosis of pre-eclampsia in singleton pregnancy: implications for clinical practice. *Ultrasound Obstet Gynecol*. 2015;45:241–246. doi: 10.1002/uog.14799.
27. Lambertini L, Diplas AI, Lee MJ, Sperling R, Chen J, Wetmur J. A sensitive functional assay reveals frequent loss of genomic imprinting in human placenta. *Epigenetics*. 2008;3:261–269.
28. Blair JD, Yuen RK, Lim BK, McFadden DE, von Dadelszen P, Robinson WP. Widespread DNA hypomethylation at gene enhancer regions in placentas associated with early-onset pre-eclampsia. *Mol Hum Reprod*. 2013;19:697–708. doi: 10.1093/molehr/gat044.
29. Mátés L, Chuah MK, Belay E, Jerchow B, Manoj N, Acosta-Sanchez A, Grzela DP, Schmitt A, Becker K, Matrai J, Ma L, Samara-Kuko E, Gysemans C, Pryputniewicz D, Miskey C, Fletcher B, VandenDriessche T, Ivics Z, Izsvák Z. Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. *Nat Genet*. 2009;41:753–761. doi: 10.1038/ng.343.
30. Vaiman D, Calicchio R, Miralles F. Landscape of transcriptional deregulations in the preeclamptic placenta. *PLoS One*. 2013;8:e65498. doi: 10.1371/journal.pone.0065498.
31. Srinivas SK, Morrison AC, Andrela CM, Elovitz MA. Allelic variations in angiogenic pathway genes are associated with preeclampsia. *Am J Obstet Gynecol*. 2010;202:445.e1–445.11. doi: 10.1016/j.ajog.2010.01.040.
32. Xiong Y, Zhou Q, Jiang F, Zhou S, Lou Y, Guo Q, Liang W, Kong D, Ma D, Li X. Changes of plasma and placental tissue factor pathway inhibitor-2 in women with preeclampsia and normal pregnancy. *Thromb Res*. 2010;125:e317–e322. doi: 10.1016/j.thromres.2010.02.017.
33. Jebbink JM, Boot RG, Keijser R, Moerland PD, Aten J, Veenboer GJ, van Wely M, Buimer M, Ver Loren van Themaat E, Aerts JM, van der Post JA, Afink GB, Ris-Stalpers C. Increased glucocerebrosidase expression and activity in preeclamptic placenta. *Placenta*. 2015;36:160–169. doi: 10.1016/j.placenta.2014.12.001.
34. Lu J, Zhou WH, Ren L, Zhang YZ. CXCR4, CXCR7, and CXCL12 are associated with trophoblastic cells apoptosis and linked to pathophysiology of severe preeclampsia. *Exp Mol Pathol*. 2016;100:184–191. doi: 10.1016/j.yexmp.2015.12.013.
35. Moore T, Dveksler GS. Pregnancy-specific glycoproteins: complex gene families regulating maternal-fetal interactions. *Int J Dev Biol*. 2014;58:273–280. doi: 10.1387/ijdb.130329gd.
36. Hancks DC, Mandal PK, Cheung LE, Kazazian HH Jr. The minimal active human SVA retrotransposon requires only the 5′-hexamer and Alu-like domains. *Mol Cell Biol*. 2012;32:4718–4726. doi: 10.1128/MCB.00860-12.
37. Hancks DC, Goodier JL, Mandal PK, Cheung LE, Kazazian HH Jr. Retrotransposition of marked SVA elements by human L1s in cultured cells. *Hum Mol Genet*. 2011;20:3386–3400. doi: 10.1093/hmg/ddr245.
38. Acampora D, Merlo GR, Paleari L, Zerega B, Postiglione MP, Mantero S, Bober E, Barbieri O, Simeone A, Levi G. Craniofacial, vestibular and bone defects in mice lacking the Distal-less-related gene Dlx5. *Development*. 1999;126:3795–3809.
39. Davideau JL, Demri P, Gu TT, Simmons D, Nessman C, Forest N, MacDougall M, Berdal A. Expression of DLX5 during human embryonic craniofacial development. *Mech Dev*. 1999;81:183–186.
40. Merlo GR, Paleari L, Mantero S, Zerega B, Adamska M, Rinkwitz S, Bober E, Levi G. The Dlx5 homeobox gene is essential for vestibular morphogenesis in the mouse embryo through a BMP4-mediated pathway. *Dev Biol*. 2002;248:157–169.
41. Deng Q, Ramsköld D, Reinius B, Sandberg R. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science*. 2014;343:193–196. doi: 10.1126/science.1245316.
42. Petropoulos S, Edsgärd D, Reinius B, Deng Q, Panula SP, Codeluppi S, Plaza Reyes A, Linnarsson S, Sandberg R, Lanner F. Single-cell RNA-seq reveals lineage and X chromosome dynamics in human preimplantation embryos. *Cell*. 2016;165:1012–1026. doi: 10.1016/j.cell.2016.03.023.
43. Nakamura T, Okamoto I, Sasaki K, Yabuta Y, Iwatani C, Tsuchiya H, Seita Y, Nakamura S, Yamamoto T, Saitou M. A developmental coordinate of plu-

ripotency among mice, monkeys and humans. *Nature*. 2016;537:57–62. doi: 10.1038/nature19096.
44. Chiu YH, Chen H. GATA3 inhibits GCM1 activity and trophoblast cell invasion. *Sci Rep*. 2016;6:21630. doi: 10.1038/srep21630.
45. Sowińska-Seidler A, Socha M, Jamsheer A. Split-hand/foot malformation: molecular cause and implications in genetic counseling. *J Appl Genet*. 2014;55:105–115. doi: 10.1007/s13353-013-0178-5.
46. Pozharny Y, Lambertini L, Ma Y, Ferrara L, Litton CG, Diplas A, Jacobs AR, Chen J, Stone JL, Wetmur J, Lee MJ. Genomic loss of imprinting in first-trimester human placenta. *Am J Obstet Gynecol*. 2010;202:391.e1–391.e8. doi: 10.1016/j.ajog.2010.01.039.
47. Chu T, Bunce K, Shaw P, Shridhar V, Althouse A, Hubel C, Peters D. Comprehensive analysis of preeclampsia-associated DNA methylation in the placenta. *PLoS One*. 2014;9:e107318. doi: 10.1371/journal.pone.0107318.
48. Bulmer JN, Morrison L, Johnson PM. Expression of the proliferation markers Ki67 and transferrin receptor by human trophoblast populations. *J Reprod Immunol*. 1988;14:291–302.
49. Reister F, Frank HG, Kingdom JC, Heyl W, Kaufmann P, Rath W, Huppertz B. Macrophage-induced apoptosis limits endovascular trophoblast invasion in the uterine wall of preeclamptic women. *Lab Invest*. 2001;81:1143–1152.
50. Redman CW, Staff AC. Preeclampsia, biomarkers, syncytiotrophoblast stress, and placental capacity. *Am J Obstet Gynecol*. 2015;213(suppl):S9.e1, S9–S11. doi: 10.1016/j.ajog.2015.08.003.
51. Redman CW, Sargent IL, Staff AC. IFPA Senior Award Lecture: making sense of pre-eclampsia: two placental causes of preeclampsia? *Placenta*. 2014;35(suppl):S20–S25. doi: 10.1016/j.placenta.2013.12.008.
52. Redline RW, Patterson P. Pre-eclampsia is associated with an excess of proliferative immature intermediate trophoblast. *Hum Pathol*. 1995;26:594–600.
53. Newhouse SM, Davidge ST, Winkler-Lowen B, Demianczuk N, Guilbert LJ. In vitro differentiation of villous trophoblasts from pregnancies complicated by intrauterine growth restriction with and without pre-eclampsia. *Placenta*. 2007;28:999–1003. doi: 10.1016/j.placenta.2007.04.008.
54. Kato T, Sato N, Takano A, Miyamoto M, Nishimura H, Tsuchiya E, Kondo S, Nakamura Y, Daigo Y. Activation of placenta-specific transcription factor distal-less homeobox 5 predicts clinical outcome in primary lung cancer patients. *Clin Cancer Res*. 2008;14:2363–2370. doi: 10.1158/1078-0432.CCR-07-1523.
55. Xu J, Testa JR. DLX5 (distal-less homeobox 5) promotes tumor cell proliferation by transcriptionally regulating MYC. *J Biol Chem*. 2009;284:20593–20601. doi: 10.1074/jbc.M109.021477.
56. Tan Y, Cheung M, Pei J, Menges CW, Godwin AK, Testa JR. Upregulation of DLX5 promotes ovarian cancer cell proliferation by enhancing IRS-2-AKT signaling. *Cancer Res*. 2010;70:9197–9206. doi: 10.1158/0008-5472.CAN-10-1568.
57. Lisonkova S, Joseph KS. Incidence of preeclampsia: risk factors and outcomes associated with early- versus late-onset disease. *Am J Obstet Gynecol*. 2013;209:544.e1–544.e12. doi: 10.1016/j.ajog.2013.08.019.
58. Suzuki S, Ono R, Narita T, Pask AJ, Shaw G, Wang C, Kohda T, Alsop AE, Marshall Graves JA, Kohara Y, Ishino F, Renfree MB, Kaneko-Ishino T. Retrotransposon silencing by DNA methylation can drive mammalian genomic imprinting. *PLoS Genet*. 2007;3:e55. doi: 10.1371/journal.pgen.0030055.
59. Barlow DP. Methylation and imprinting: from host defense to gene regulation? *Science*. 1993;260:309–310.
60. McDonald JF, Matzke MA, Matzke AJ. Host defenses to transposable elements and the evolution of genomic imprinting. *Cytogenet Genome Res*. 2005;110:242–249. doi: 10.1159/000084958.
61. Varela M, Spencer TE, Palmarini M, Arnaud F. Friendly viruses: the special relationship between endogenous retroviruses and their host. *Ann NY Acad Sci*. 2009;1178:157–172. doi: 10.1111/j.1749-6632.2009.05002.x.
62. Rawn SM, Cross JC. The evolution, regulation, and function of placenta-specific genes. *Annu Rev Cell Dev Biol*. 2008;24:159–181. doi: 10.1146/annurev.cellbio.24.110707.175418.
63. Sugimoto J, Schust DJ. Review: human endogenous retroviruses and the placenta. *Reprod Sci*. 2009;16:1023–1033. doi: 10.1177/1933719109336620.
64. Ruebner M, Strissel PL, Ekici AB, Stiegler E, Dammer U, Goecke TW, Faschingbauer F, Fahlbusch FB, Beckmann MW, Strick R. Reduced syncytin-1 expression levels in placental syndromes correlates with epigenetic hypermethylation of the ERVW-1 promoter region. *PLoS One*. 2013;8:e56145. doi: 10.1371/journal.pone.0056145.
65. Vargas A, Toufaily C, LeBellego F, Rassart É, Lafond J, Barbeau B. Reduced expression of both syncytin 1 and syncytin 2 correlates with severity of preeclampsia. *Reprod Sci*. 2011;18:1085–1091. doi: 10.1177/1933719111404608.
66. Hancks DC, Kazazian HH Jr. SVA retrotransposons: evolution and genetic instability. *Semin Cancer Biol*. 2010;20:234–245. doi: 10.1016/j.semcancer.2010.04.001.
67. Bellessort B, Le Cardinal M, Bachelot A, Narboux-Nême N, Garagnani P, Pirazzini C, Barbieri O, Mastracci L, Jonchere V, Duvernois-Berthet E, Fontaine A, Alfama G, Levi G. Dlx5 and Dlx6 control uterine adenogenesis during post-natal maturation: possible consequences for endometriosis. *Hum Mol Genet*. 2016;25:97–108. doi: 10.1093/hmg/ddv452.
68. Horike S, Cai S, Miyano M, Cheng JF, Kohwi-Shigematsu T. Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. *Nat Genet*. 2005;37:31–40. doi: 10.1038/ng1491.
69. Chao HT, Chen H, Samaco RC, Xue M, Chahrour M, Yoo J, Neul JL, Gong S, Lu HC, Heintz N, Ekker M, Rubenstein JL, Noebels JL, Rosenmund C, Zoghbi HY. Dysfunction in GABA signalling mediates autism-like stereotypies and Rett syndrome phenotypes. *Nature*. 2010;468:263–269. doi: 10.1038/nature09582.
70. Stühmer T, Anderson SA, Ekker M, Rubenstein JL. Ectopic expression of the Dlx genes induces glutamic acid decarboxylase and Dlx expression. *Development*. 2002;129:245–252.
71. Licht P, Harbarth P, Merz WE. Gaba-mediated stimulation of hcg secretion suggests a parallelism in the control of central-nervous and placental gonadotropin release. *Placenta*. 1992;13:151–161

Disturbed Placental Imprinting in Preeclampsia Leads to Altered Expression of DLX5, a Human-Specific Early Trophoblast Marker

Julianna Zadora, Manvendra Singh, Florian Herse, Lukasz Przybyl, Nadine Haase, Michaela Golic, Hong Wa Yung, Berthold Huppertz, Judith E. Cartwright, Guy Whitley, Guro M. Johnsen, Giovanni Levi, Annette Isbruch, Herbert Schulz, Friedrich C. Luft, Dominik N. Müller, Anne Cathrine Staff, Laurence D. Hurst, Ralf Dechend and Zsuzsanna Izsvák

Circulation. 2017;136:1824-1839; originally published online September 13, 2017;
doi: 10.1161/CIRCULATIONAHA.117.028110
Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Print ISSN: 0009-7322. Online ISSN: 1524-4539

SUPPLEMENTAL MATERIAL

Supplemental Methods:

Patients

Microarray analysis of human placenta and decidua samples were described earlier[1]. The clinical biomarkers (CTproAVP, MRproADM, MRproANP, CRP) used for clustering analysis were described earlier[2, 3]. The study comprises of patient samples from bio-bank collection at Oslo University Hospital, Oslo, Norway, approved by the Regional Committee of Medical Research Ethics in South-Eastern Norway. The cohort consists of placental and decidual tissues collected following caesarean sections from 56 preeclamptic women (PE) and 28 women with normotensive and uncomplicated pregnancies. The PE group was subdivided into PE + IUGR (n = 14) and PE without IUGR (n = 42). Furthermore, the PE group was divided into early-onset PE (delivery <34 gestational week) and late-onset PE (delivery ≥34 gestational week)[4]. Placenta biopsies were collected according to a strict protocol, following stat-of-the-art sampling and storage[5]. In brief, biopsies were collected from the center of a macroscopically normal looking central cotyledon within 5 minutes of delivery by study personnel. Biopsies were snap-frozen in liquid nitrogen and stored at –80°C until further processing. Decidual tissue was collected through vacuum suctioning of the placental bed[6]. The levels of serum PlGF and sFLT1 of patients from the second preeclamptic cohort used for correlation analysis were measured on Elecsys (Roche Diagnostics) at HELIOS Kliniken GmbH. sFlt1/PlGF ratios were calculated[4].

The healthy (n = 5) and PE (n = 5) primary trophoblast cells were isolated from human placenta samples obtained from HELIOS Klinikum in Berlin. Human placenta sampling was approved by the Regional Committee of the Medical Faculty of Charité Berlin. All placentas were processed within 2 h after delivery. Control trophoblast cells were isolated from placentae of uncomplicated pregnancies collected following Caesarean section at term. PE

was defined by hypertension (SBP ≥140 mmHg or DBP ≥90 mmHg) and proteinuria (≥0.3 g in a 24-hour urine specimen) according to the American College of Obstetricians and Gynecologists (ACOG) classification.

Cluster analysis of microarray data

We employed R platform (https://cran.r-project.org/) and various Bioconductor packages (https://bioconductor.org/biocLite.R) to analyze the human placenta and decidua microarray data. We extracted significance level of normalized expression values against the background corresponding to each probe IDs using "*lumi*", a R Bioconductor package where the variance-stabilizing transformation (VST) was applied to deal with sample replicates and robust spline normalization (RSN) for normalization[7]. Probes with p-value < 0.05 were further transformed to $\log_2$ scale and IDs were annotated as gene names from "*illuminaHumanV3.db*" from Bioconductor annotation data package containing 47,324 probes. Consequently, we generated expression matrix of 38,382 significant probes in this study. Expression values of a gene with multiple probes were assigned by their mean, resulting in 26,886 unique genes. In order to achieve consistency between arrays, $\log_2$ transformed expression values underwent through quantile normalization using "normalizeQuantile" function built-in "*limma*" package. Quality of processed data was further assured by PCA visualization[8]. Each gene value was further assigned as their relative abundance value which is the ratio of expression value of gene in each sample and the mean value of expression for corresponding gene across the samples. The resulting relative expression matrix was subjected to unsupervised hierarchical clustering which blindly classifies all samples based on their transcriptome pattern without the prior knowledge of their disease status. Hierarchical clustered dendrograms were generated using "average" linkage method (mean distance between an observation in one cluster and an observation in the other cluster) on the dissimilarity matrix. Pairwise distances were calculated using Euclidian distance method on dissimilarity matrix obtained by pairwise

Spearman correlation. P-value threshold for correlation test of matrix was kept up to 0.01. In order to examine statistical reliability of clustering we performed bootstrapping (1000 replicates) on unbiased hierarchically clustered dendrogram of distance matrix by "average" method using pvclust[9]. In present study, fold change of differential gene expression (DEG) between samples on the $\log_2$ scale was analyzed using linear and e-Bayesian model algorithms from "*limma*" R Bioconductor package[10]. P-value of DEGs between the three PE clusters (n < 8) vs clustered Controls (n = 16) was obtained from empirical Bayesian built-in function from "*limma"*. We did not adjust the p-values by using multiple hypothesis tests in order to preserve the family-wise error-rate (FWER) that were supportive for further linking the genes with clinical phenotypes. Pairwise differential gene expression is shown on the heatmap displaying $\log_2$ fold change values, which were detected in all three PE clusters against controls at a significant level p-value < 0.05. Venn diagram was used to represent the distribution of differentially expressed gene in the clusters using "*venny*", an online tool (http://bioinfogp.cnb.csic.es/tools/venny/).

Furthermore, imprinted genes were extracted from the differentially expressed genes to compare their expression in the identified preeclamptic clusters. We analyzed differential expression of maternally (MEGs) and paternally (PEGs) expressed genes in all PE placenta samples. The list of imprinted genes was analyzed according to the Geneimprint website that gives known and predicted human imprinted genes (http://www.geneimprint.com/site/home/), and recently published list of imprinted genes in human placenta[11]. We defined imprinting as genes showing statistically significant allelic expression of < 20% mean expression ratio between maternal and paternal allele as a paternally expressed genes, and those showing > 80% as maternally expressed genes according to the allelic expression in the study[11]. This resulted in a total of 257 MEGs and 150 PEGs (Supplementary Table 3). For analysis of imprinting, we considered also M/P allelic expression ratio (M[%] – P[%]) following values:

60-40, 70-30, as well as 90-10. Wilcoxon test was subjected to analyze mean $\log_2$ fold change of differentially expressed MEGs and PEGs in PE.

Pathway analysis of deregulated genes

Canonical pathways and biological function of the differentially expressed genes identified in microarray data sets were investigated using QIAGEN's Ingenuity® Pathway Analysis (IPA®, QIAGEN Redwood City, www.qiagen.com/ingenuity). The Ingenuity Knowledge Base (www.ingenuity.com) was used as reference gene sets (pathways). Overrepresentation of biological pathways was assessed by Fisher's exact test and corrected for multiple testing by the Benjamini-Hochberg procedure. The ratio (overlap) is calculated as a number of genes from the dataset that map to the pathway divided by the number of total genes included into the pathway.

Clinical data analysis

Clinical data and biomarker parameters of the preeclamptic patients subjected for microarray were analyzed to assess the difference in clinical outcome in each preeclamptic cluster. We leveraged our clinical and transcriptome datasets to infer similarity between clinical phenotypes, patients (including or excluding controls) and moreover, genes contributing to each clinical phenotype by our pairwise correlational analysis. First, we created the matrix of phenotypical values for each individual subject considering only those that had continuous variables (non-binary values of clinical phenotypes) and in addition, we removed those that had missing data points (<75% of samples). It resulted in 33 phenotypes (in black in Suppl. Mat, Table S1. Patient`s clinical data, page) for 46 samples (24 Patients and 22 Controls). Missing data points (>75% of samples) in resulting matrix was assigned as mean value (since we further considered values relative to mean for comparison) of respective clinical variable. Finally, we transformed data points relative to mean of their respective clinical variables. We

then constructed dendrograms via bootstrapping (10 replicates) based hierarchical clustering using ranked correlation and complete linkage method on relative values from the clinical variables as stated above. Height of dendrograms represents the Euclidian distance of dissimilarity matrix; numbers in red and blue indicate au and bp values from bootstrapping. We used similar strategy to perform sample's (Patients with or without Controls) and clinical phenotype's clustering on the basis of variable clinical data. In addition, when analyzing only Patients data we added four more variables (in blue in Suppl. Mat, Table S1. Patient`s clinical data) because this data was available for Patients but not for Controls.

In order to make our analysis consistent and to find association between genes and clinical phenotype, expression data was also assigned as relative to mean values and resulting two dataframes (clinical and expression data) that were merged on the basis of their sample IDs. Pairwise ranked correlation was calculated for each clinical phenotype with each gene across the dataframe; to obtain the genes associated with any of given phenotypes correlational threshold was kept to 60 (any pair doesn't fit to the criterion was given white color in the heat map displaying correlogram, whereas red denotes positive and black as negative associations). As a novelty, we show the list of genes corresponding/associate with the clinical phenotype in the supplement (Suppl. Mat. Figure S2B).

mRNA isolation, RT-PCR

Total RNA was isolated from tissues (homogenized by ceramic beads) and cells using QIAzol lysis reagent and Qiagen RNeasy mini kit (including on-column DNAase I digestion) (Qiagen) according to the manufacturer's protocol. RNA quality and concentration was measured by NanoDrop-1000 spectrophotometer (PeqLab). RNA was reverse transcribed into cDNA using High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) and detected by real-time polymerase chain reaction (PCR) on ABI 7500 Fast Sequence Detection

System (Applied Biosystems). Data was analyzed by 7500 Fast System Software (Applied Biosystems). Primers and probes were designed with PrimerExpress 3.0 (Applied Biosystems), synthetized by Biotez (Germany) and are listed in Table S2. The expression of all of the genes was normalized to 18S expression.

Genomic DNA extraction

Frozen placenta tissues were pulverized with liquid nitrogen. DNA from placentas was extracted using the DNeasy Blood and Tissue Kit (Qiagen) according to the manufacturer's instructions. DNA was stored at −20°C.

Loss of imprinting assay

Loss of imprinting assay (LOI) was carried for DLX5 gene. Detail of assay design has been already described[12]. To test differential allelic mRNA expression a single nucleotide polymorphism (SNP) for DLX5 (rs73708843) was used. gDNA from 97 preeclamptic and controls placental samples were genotyped for heterozygous SNP using standard sequencing protocol. Next cDNAs from heterozygous placental samples were amplified with gene-specific primers bracketing the readout polymorphism using FastStar Essential DNA Green Master (Roche). The reaction was carried on 20 ng cDNA templates with 0.2mM primers in a final volume of 20 µl. Cycling conditions were: 95.0°C for 10 min, followed by 35 cycles of 95.0°C for 30 sec, 70.0°C for 30 sec and 72.0°C for 30 sec. Quantitative allele-specific PCR (qAS-PCR) was detected using allele-specific primers. These primers were optimized for allelic discrimination using heterozygote genomic DNA amplified on the LightCycler480™ (Roche). The reaction was carried with FastStar Essential DNA Green Master (Roche) on 20 ng gDNA template, 0.2mM primers and final volume 20 µl. Cycling conditions were: 95.0°C for 10 min, followed by 45 cycles of 95.0°C for 30 sec, 70.0°C for 30 sec and 72.0°C for 30 sec.

PCR cDNA amplicons from heterozygous placental samples for the selected readout SNPs were diluted 1:40-1:100 accordingly to the abundance of the mRNA and amplified on the LightCycler480™ (Roche) in the reaction containing: FastStar Essential DNA Green Master Mix (Roche), 0.2mM primers, diluted DNA template; final volume 20 µl. Cycling conditions were: 95.0°C for 10 min, followed by 45 cycles of 95.0°C for 30 sec, 70.0°C for 30 sec and 72.0°C for 30 sec. The primers are listed in Table S2. LOI was measured as an expression of the silenced allele, which was calculated as:

$LOI = 2^{-|\Delta Cp|}$

where the $-|\Delta Cp|$ corresponds to the absolute difference between allele-specific Cp values on cDNA level corrected for the specificity of the allele-specific PCR[12].

To validate allelic expression of DLX5 in human placenta we subjected cDNA of the placenta samples showing no LOI for standard sequencing analysis.

DLX5 CpG methylation analysis

We reanalysed methylation and gene expression data of a PE dataset[13]. We explored "*methylumi*", "*lumi*", "*minfi*" for methylation analysis. "*minfi*" gave highest level of Pearson's correlation within the groups of patient and controls (data is available). We used Bioconductor package "*minfi*"[14] for methylation data analysis and "*lumi*" for gene expression data analysis (as described previously). We calculated $\log_2$ fold change of differential gene expression level in PE patients compared to controls using linear modelling algorithm from Bioconductor package "*limma*" with multiple testing correction by applying Benjamini and Hochberg's (BH) false discovery rate (FDR) analysis.

We preprocessed and normalized raw data by performing background correction against controls. Furthermore, we constructed data frame of normalized data by assigning the Beta (β) value for each probe across the samples.

We fetched the probe IDs of CpG methylation data frame, falling in the genomic window of DLX5 locus in addition with that of 10KB upstream region from the normalized data frame of 20 PE patients and 20 controls. We calculated the differential methylation regions (for the probes, which had ≥ 50% of methylation in either of the cases). We calculated $\log_2$ fold change of methylation level in PE patients compared to controls using linear modelling algorithm from Bioconductor package "*limma*".

Secondly, we created data frame of DLX5 expression and methylation level across 8 controls and 8 PE patients. We then assigned relative expression and methylation values to each probe across the samples. We calculated pairwise correlation of relative expression of DLX5 with relative methylation of various methylation sites using "*rcorr"* function from *"Hmisc"* package using Spearman's test. For each pair, we obtained p-values using "*cor.test*" and FDR was calculated using Benjamini and Hochberg's (BH) test for multiple testing using *p.adjust* function.

Single cell RNA sequencing data analysis

We reanalyzed single cell RNA sequencing data from two studies for human embryogenesis[15] and for mouse embryogenesis[16].

Reads were mapped to the human genome (hg19) and mouse genome (mm10) using STAR[17] with our defined settings i.e. *--alignIntronMin 20 --alignIntronMax 1000000 --chimSegmentMin 15 --chimJunctionOverhangMin 15 --outFilterMultimapNmax 20*, and only uniquely mapped reads were considered. We obtained counts using featureCounts[18] at 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 locating between TSS and TES). This we did using our *in-house* R script (available on request). We chose samples expressing more than 5000 genes. We subsequently selected for further analysis those genes that express ($\log_2$ TPM > 2) in at least 5 of these samples. This resulted in 1285

single cells expressing 15501 genes for human samples, whereas for mouse we recovered 259 single cells expressing 15181 genes (Note: DLX5 was not recovered in mouse as it didn`t pass the threshold). We adjusted the batch effects arising from single cell sequencing using COMBAT[19] from SVA[20] package in R.

Each gene was represented by an across sample vector of z-scores. The resulting z vectors for all the genes were clustered by Pearson's correlation method. Single-cell RNA-seq data was analyzed using "Seurat" R package[21]. In order to obtain significant clusters of cell population at E4 and E5 stage, we first excluded all genes that were very lowly expressed on average (log(TPM+1)<2) and identified the set of genes that was most variable across the dataset by z-score cutoffs, as described in[22]. We next reduced the dimensionality of our dataset using PCA. As previously described in[22], we ran PCA using the '*prcomp*' function in R using most variable genes as an input. We did 200 random samplings to find significant genes, each time randomly permute 1% of genes. This gave us p-value for each gene in each PC, based on how likely the gene/PC score would have been observed by chance. Next, we utilized a modified randomization approach ('jack straw'), a built-in function in "Seurat" package to identify 'statistically significant' principal components in the datasets. 'Significant' PCs will have a strong enrichment of genes with low p-values. We found PC 1-9 as statistically significant. We then chosen significant genes for PC1-9, allowed each PC to contribute a max of 200 genes (as described in "Seurat" package). When we run tSNE using our 9 significant PCs as an input (spectral tSNE), we got distinct point clouds. Finally, we calculated density cluster of the tSNE map (G.use parameter is the density parameter for the clustering – we set it to 7). This yielded distinct clusters from scRNA-seq datasets. We used recently developed single-cell visualization approach based on the nonlinear dimensionality reduction t-distribution stochastic neighborhood embedding (t-SNE) machine learning algorithm for clustering the transcriptome of human and mouse single cell transcriptomes. We used t-SNE instead PCA since unlike PCA, t-SNE does not map the points between the two contrasting eigen vectors

(Principle components) i.e. higher and lower dimension spaces; t-SNE maps the points in the way that points that are close in the higher dimensional space will be close (with high probability) in the low-dimensionality embedding. It converts similarities between data points to joint probabilities and tries to minimize the divergence between joint probabilities of low-dimensional embedding and high-dimensional data. Differentially expressed genes between single cell clusters were calculated using SCDE algorithms[23].

We ran t-SNE on mouse and human cells and defined clusters of cell populations to identify genes expressing exclusively in those clusters. For human, we redefined clusters of cell populations from 353 single cells only from E4-E5 stages where inner cell mass (ICM) and trophectoderm (TE) split from Morulae. For human, we identified three major clusters containing approximately 90% of total cells and three minor clusters containing < 10% of total cells from E4 and E5 stage of human embryonic development, whereas for mouse we found seven clusters (data not shown). We defined the cell type for each cluster according to the known markers[24, 25]. We identified human TE defining cells and early and late mouse TE cells.

For cynomolgus monkeys (*Macaca fascicularis*), we obtained processed data frame from GEO datasets GSE74767[26] as RPKM and converted it to TPM. We grouped the cells as their pre-implantation embryonic days from E6-E9 and extracted the level of DLX5 expression for each stage. Furthermore, to analyze cell type specific expression pattern of DLX5 at the pre-implantation stage, we analyzed the expression level of DLX5 in pre-implantation hypoblast, ICM, Epiblast, early and late trophectoderm.

For imprinting expression analysis during human embryogenesis, we obtained raw fastq files from GEO database under accession number GSE36552[27]. We calculated the expression on $\log_2$ TPM as above. The data was further scaled to z-score for each gene in order to demonstrate the relative expression. We removed any gene that had < 1 TPM expression level in all the analyzed samples. We used list of imprinted genes as described above and extracted

from the data frame. Hierarchical clustering analysis of imprinted genes expressed during early (oocyte, zygote, 2-cell and 4-cell stage), mid (8-stage, Morulae) embryogenesis and late blastocyst was subjected separately for maternally (179/257) and paternally (86/150) expressed genes.

RNA sequencing of human trophoblasts

Total RNA was isolated from five control and five early-onset PE human trophoblasts using Trizol lysis reagent and Direct-zolTM RNA MiniPrep kit including DNase I on-column digestion (Zymo Research) according to the manufacturer`s protocol. Concentration of RNA was quantified on NanoDrop Spectrophotometer ND-1000 and the quality of RNA was analysed using Agilent RNA 6000 Nano Kit on Agilent 2100 Bioanalyzer machine. Library for RNA sequencing was prepared from 550 ng of RNA using Illumina TruSeq Stranded mRNA LT Set A kit (cat. no. RS-122-2101) according to TruSeq Stranded mRNA Sample Prep LS Protocol. Samples were indexed with sample-specific indices which allow for the pooling and sequencing of all libraries in two pools of five samples. Expression profiling of trophoblast transcriptomes by high throughput sequencing were performed in BIMSB Genomics Platform of Max Delbrück Center for Molecular Medicine (Berlin, Germany) on Illumina HiSeq2000 sequencing platform. The clustering of the index-coded samples was performed on a cBot Cluster Generation System using PE TruSeq Cluster Kit v3-cBot-HS (Illumina) according to the manufacturer's instructions. After cluster generation, sequencing was performed on an Illumina HiSeq 2000 platform as 100 bp first strand specific paired-end reads.

Sample-specific barcoded sequencing reads were de-multiplexed from multiplexed flow cells. The resulting BCL files were converted to FASTQ format files using CASAVA 1.8.2. The quality of the raw sequence reads was determined with the FastQC[28]. Reads with quality score < 30 were removed. We also truncated 2nt from the end of sequencing reads, since their

average quality score was not same as the rest of nucleotides. This resulted in at least 70 million reads per sample. Next, reads were mapped over the reference genome (Human hg19/GRCh37) and transcriptome model (hg19.refseq.gtf), downloaded from USCS tables (http://hgdownload.cse.ucsc.edu/goldenPath/hg19/bigZips/) using TopHat v2.0.8, samtools 0.1.17.0 and Bowite 2.0.5.0 applying parameters as: *"tophat2 -p 8 –r 150--mate-std-dev 140 --library-type fr-firststrand"*. On average 75% of total reads were uniquely mapped on the annotated gene models, approximately 10% of total reads are uniquely mapped on repeated fraction of genome (data not shown). Transcript assembly for each individual sample was conducted using Cufflinks v2.08 measured as FPKM. For calculation of differentially expressed genes (DEGs) we calculated Counts Per Million (CPM) using featureCounts[18] and algorithms from "*DESeq2*" which performed quantization and statistical inference of systematic changes between conditions, as compared to within-condition variability[29]. The package "*DESeq2*" provides methods to test for differential expression by use of negative binomial generalized linear models; the estimates of dispersion and logarithmic fold changes incorporate data-driven prior distributions. In addition, for the outlier's samples (trophoblast sample PE1) DEGs were calculated using a single-replicate model. The read counts were calculated with *featureCounts* from *subread* package (http://subread.sourceforge.net/), FPKM was calculated using *bamutils* (http://ngsutils.org/modules/bamutils/count/). Next, *Random Variable1 (Var1) = n.l.x* formula was used, where *x* (Random Variable 2) is the expression level of this gene, *n* is reflecting the sequencing depth and *l* is the gene length. The two random variables were used in the published model of GFOLD algorithm, which calculate the normalization constant and variance to extract fold changes from unreplicated RNA-seq data [30].

Distinct loci of repeat elements i.e. obtained from repbase libraries (http://www.girinst.org/repbase/). We counted one alignment per uniquely mapped reads over the distinct loci and calculated counts per million (CPM) normalized to library size. In order

to decipher the expression-set for each repeat elements family, we used both uniquely and non-uniquely mapped reads over the distinct families and calculated expression in similar fashion. Non-unique reads were counted considering only one alignment per read for the each repeat elements family that is mapped exclusively within the respective family. We used Kolmogorov-Smirnov test (KS) on the distribution of RNA-seq read counts per million (log scale) over distinct loci of L1 and SVA elements. Observed p-values from KS test were further processed for multiple testing using Benjamini and Hochberg's (BH) test to calculate FDR.

Immunofluorescent staining

Human control placenta sections were deparaffinized in xylene and followed by rehydration through graded ethanol. Antigen was retrieved by heat-induced epitope retrieval (HIER) method in 0,01 M Citrate buffer pH 7. Non-specific antibody binding was blocked with Ultra V Block (Lab Vision). Double immunostaining of rabbit anti-DLX5 antibody (Sigma-Aldrich) (diluted 1:50 in Antibody Diluent (Dako)) and mouse anti-Cytokeratin-7 (1:1000, Abcam) was applied for 45 minutes. The secondary antibody goat anti-rabbit Alexa 555 (1:200) and goat anti-mouse Alexa 488 (1:200) was incubated for 30 minutes. After extensive washes with phosphate-buffered saline with 0.05% Tween-20 (PBS/T) cell nuclei were stained with DAPI (diluted 1:2000 in PSB) and sections were mounted with ProLong Gold Antifade Mounting Medium (Life Technologies).

Immunohistochemistry staining

Human control and PE placenta sections were deparaffinized and dehydrated. Antigen was retrieved by heat-induced epitope retrieval (HIER) method in 0.01 M Citrate buffer pH 6. Endogenous peroxidase activity was blocked with Hydrogen peroxide (Lab Vision/Thermo scientific) for 10 minutes. Tris-buffered saline with 0.05 % Tween-20 (TBS/T) was used in all

washing steps. Non-specific immunoglobulin binding was blocked with Ultra V Block (Lab Vision). Rabbit anti-DLX5 antibody (Sigma-Aldrich) was diluted (1:200) in Antibody Diluent (Dako) and applied for 45 minutes. The secondary antibody HRP Polymer (ready-to-use; Lab Vision) was incubated for 20 minutes. Peroxidase activity was detected with aminoethylcarbazole (AEC) chromogen (ready-to-use; Lab Vision). All sections were counterstained with hemalaun and mounted with Kaiser's Glyceringelatine (Merck).

Constructs generation

Human DLX5 coding sequence (sequence ID: refNM_005221.5) was amplified by PCR from placental cDNA with primers containing restriction sites for cloning into pCAGGS-HA tag vector. PCR reaction was carried with Phusion High-Fidelity DNA Polymerase (New England BioLabs) on 50ng of cDNA template with 0.2 mM dNTPs, 0.5 μM primers, 3% DMSO in a final volume of 20 μl. Cycling conditions were: 98.0°C for 30s, followed by 35 cycles of 98.0°C for 10 sec, 56.0°C for 30 sec and 72.0°C for 1.5 min and final extension of 10 min at 72.0°C. The PCR product of ~870bp was extracted from 1% agarose gel, digested with EcoRV and NotI restriction enzymes and cloned in pCAGGS-HA vector into the EcoRV/NotI restriction site. It resulted in generation of N-terminal HA-tagged DLX5 expressing construct (pCAGGS-HA-N-DLX5). To generate a construct for stable transgene overexpression, the CAGGS-HA-DLX5 cassette was cut out with SspI restriction enzyme from pCAGGS-HA-N-DLX5 and cloned into EcoRV site, generated by site directed mutagenesis into pT2B-puro *Sleeping Beauty* vector carrying the SV40-puro cassette for puromycin selection.

Cell culture and stable DLX5 overexpression and microarray

SGHPL-4 cells were a kind gift from Judith E. Cartwright (St George's University of London, London, United Kingdom). This cell line is derived from primary human first trimester extravillous trophoblasts (EVT), transfected with the early region of SV40 and show similar

invasive capabilities to primary EVTs and retain features of normal EVTs [31]. SGHPL-4 cells were cultivated in HAM's F10 (Biochrom) media containing 10% (v/v) FCS, 2mM glutamine and 1% (v/v) penicillin/streptomycin (P/S) in 37°C, 5% CO_2.

SGHPL-4 cells were electroporated with 10:1 ratio of vector carrying DLX5 overexpression cassette and plasmid containing *Sleeping Beauty* transposase using Neon Transfection System (Life Technologies). Two days post-transfection puromycin selection was carried for two weeks. DLX5 overexpression was confirmed by RT-qPCR and western blot.

Transcriptome analysis of DLX5-overexpressing SGHPL-4 cells was performed on Illumina HumanHT-12_V4_0_R2 BeadChip according to manufacturer's protocol. Total RNA was isolated from SGHPL-4 cells overexpressing human DLX5 protein using Direct-zol™ RNA MiniPrep kit including DNase I on-column digestion (Zymo Research) according to the manufacturer`s protocol. Six independent transfections as biological replicates were used for the study. Expression values against the background corresponding to each probe IDs were analyzed using "*lumi*" R Bioconductor package. Probes with p-value < 0.05 were further transformed onto $\log_2$ scale. Probe IDs were annotated as gene names from "*illuminaHumanv4.db*" from Bioconductor annotation data package. Expression values of multiple probes for one gene were assigned by their mean. We then adjusted the upper quantile variation, which resulted in normalized data matrix of genes from samples generated in present study. Fold change of differentially expressed genes between wild type and OE-DLX5 samples on the $\log_2$ scale were analyzed using linear and e-Bayesian model algorithms from "*limma*" R Bioconductor package with multiple testing correction by applying Benjamini and Hochberg's (BH) false discovery rate (FDR) analysis on the threshold of 0.05. DEGs are presented on *MA plot* where each point represents the $\log_2$ fold change of genes (y-axis) plotted against the $\log_2$ average expression (x-axis) and the significant level of p-value < 0.05 are colored as red for upregulated genes and blue for downregulated genes.

Intersection of preeclamptic microarray and DLX5 microarray

Quantile normalized datasets from the two microarrays described above were used in the analysis. In order to generate a matrix of expression level for unique genes in each sample, two datasets from different platforms were merged by their unique gene names in total for 58 samples. The batch effect arising from two different platforms was corrected by normalizing surrogate variances from "Combat" package from R Bioconductor[19]. The corrected batch effect was confirmed by Principal Component Analysis (PCA). In order to remove the uneven biases, we considered equal number of patients and controls or equal number of OE-DLX5 and wild type samples prior to intersecting datasets for further analysis. Each gene expression value was further assigned as their relative abundance value, which is the ratio of expression level of a gene in each sample and mean expression values of the gene across the samples. We considered comparison of 6 wild type and 6 OE-DLX5 SGHPL-4 transcriptomes with each of the three preeclamptic clusters and equal number of controls which were randomly chosen. The resulting expression matrix was subjected to hierarchical clustering (Spearman rank correlation, average linkage). P-value threshold for correlation test for matrix was kept up to 0.01 where the hierarchically clustered dendrogram displays gene expression patterns relative to the mean of all the replicated samples included in this analysis. In order to examine statistical reliability of clustering we performed bootstrapping (1000 replicates) on unbiased hierarchically clustered dendrogram of distance matrix using Ward method.

Furthermore, to find a significant impact of DLX5 on PE we clustered the transcriptomes of wild type and OE-DLX5 SGHPL-4 cells with preeclamptic placenta samples that displayed low and high DLX5 expression level, respectively. Six OE-DLX5 and six wild type transcriptomes were analyzed with the equal number of patients expressing higher level of DLX5 and lower level of DLX5. This resulted in two groups of 12 samples: OE-DLX5 and wild type cells, and PE DLX5-high and PE DLX5-low. For each gene, the expression was calculated as relative to the average value of gene expression across the 12

samples for each group. Then both dataframes were merged together to form a single matrix. The batch effect was removed as defined previously. In the next step, genes showing significant variation (we applied Poisson's distribution model on the data, genes lying out of curve were considered as genes showing higher variance in the matrix of 24 samples) were retrieved, which resulted in total of > 3000 genes. Furthermore, we removed 52 genes, which were showing more variation within the group than between the groups. Finally, the heatmap and clusters were constructed as defined previously.

To find a significant correlation gene networks we applied Weighted Gene Co-expression Network Analysis (WGCNA), which is used to identify clusters (modules) of highly correlated genes[32]. We employed this algorithm to find the set of genes co-expressed in our total merged 58 sample's transcriptome. Total microarray datasets of 58 samples including wild type and OE-DLX5 cells and placenta tissues from individuals with and without PE, provides higher level of variability thus giving stronger power to this study to generate a significantly enriched co-expression networks and moreover, to check whether DLX5 is part of them. WGCNA gave us significant co-expression networks as various modules with ~ 3000 genes. We then used DLX5 as a probe to find out the networks associated with DLX5 from WGCNA output. The metamodel derived from Spearman's correlation threshold of positive (60%) and negative (55%) gave the probable significant networks. We kept this threshold to extract all genes, which were correlated at provided correlation thresholds with DLX5 from extracted networks. Finally, pairwise ranked correlation matrix of identified genes was generated.

Western blot

Equal numbers of cells were lysed with 3X Sample Buffer (cooked in 95^{o}C for 7 min) and denaturized proteins were separated on SDS-PAGE gel. Proteins were then transferred to PDVF membrane using semi-dry blotting. Unspecific binding was blocked with 2.5% w/v

non-fat dry milk in TBS-T. Primary antibodies: rabbit anti-DLX5 (1:1000, Sigma-Aldrich) and mouse anti-β-actin (1:1000, Dianova) were incubated overnight at 4°C. Secondary antibody goat anti-rabbit (1:5000, Thermo Scientific) and goat anti-mouse (1:5000, Thermo Scientific) against the host of primary antibody was used conjugated with horseradish peroxidase and membranes were developed with Amersham ECL Prime Western Blotting Detection Reagent (GE Healthcare). Membranes were then stripped using mild stripping buffer, blocked again and incubated with loading control antibody.

Immunostaining of DLX5 in SGHPL-4 cells

SGHPL-4 wild type and OE-DLX5 cells were seeded in 8-wells chamber slide and incubated in 10% (v/v) FCS Ham´s F10 medium, 1% AA and incubated O/N at 37°C, 5% CO_2. Cells were washed with PBS and 500 µl of ice-cold methanol was added. After 10 min of incubation on ice, methanol was removed. Next, cells were washed 3 times with 500 µl of PBS-T (0.3% TritonX-100). Cells were blocked by incubation with 500 µl of normal donkey serum (10% NDS in PBS) for 30 min at RT. Rabbit anti-DLX5 antibody (Sigma-Aldrich) was diluted (1:50) in PBS with 1% FCS and applied for O/N incubation at 4°C with gentle shaking. After washing, secondary antibody donkey CyTM3 IgG anti-rabbit (1:200; Jackson ImmunoResearch) was added and cells were incubated for 2h at RT in dark. After extensive washes with PBS cell nuclei were stained with DAPI solution (Vectashield with DAPI) for 10 min at RT in dark, washed with PBS and mounted with AquaPoly/Mount solution (Polysciences).

Time-lapse microscopy: proliferation and apoptosis

Proliferation and apoptosis of SGHPL-4 cells was observed over time using an Olympus IX70 inverted microscope equipped with a Hamamatsu C4742-95 digital camera. SGHPL-4 wild type and DLX5-overexpressing cells were preincubated O/N in 0.5% (v/v) FCS in HAM´s

F10 medium (Biochrom) containing 1% P/S. Next control (no treatment), TNFa-treated (30ng/ml) wild type and DLX5-overexpressing cells were followed over 48 h using time-lapse microscopy in HAM´s F10 medium with 0.5% (v/v) FCS, 1% P/S. The microscope and stage were enclosed within a heated (37°C) chamber (Solent Scientific, UK) and cells were cultured in 5% CO_2 in air. Images were captured every 15 min for 48 hours and analyzed using Image Pro Plus software (Media Cybernetics, USA). For analysis, 40 cells in one field of view for each condition were randomly chosen at the beginning of the time-lapse sequence and the distance moved by each cell was recorded. Cell proliferation was assessed as a cumulative number of dividing cells over the 48 h of incubation. Apoptosis was scored according to the time at which clear apoptotic cell morphology was observed during 48 h of incubation (characterized as a membrane blebbing, cytoplasmic shrinkage, nuclear condensation, a phase bright appearance and the formation of blisters)[33].

Cell count using High Throughput Samples (HTS)

SGHPL-4 wild type and DLX5-overexpressing cells were seeded in 12-well plate in 10% (v/v) FCS HAM´s F10 medium (Biochrom), 1% AA and allowed to attach for 3h in 37^{o}C, 5% CO_2. Next control (no treatment) and TNFa-treated (TNFa treatment at conc. 30ng/ml) cells were incubated for 48h in HAM´s F10 medium containing 0.5% (v/v) FCS and 1% Antibiotic-Antimycotic (AA) (Gibco). Cells were trypsinized and collected in a total volume of 100 µl of cell suspension in a 96-well plate and counted using high throughput sampler (HTS) device on BD FACSCanto II system using BD FACSDiva software. Data was analyzed using FlowJo software and is presented as absolute cell number in in SGHPL-4 cell gate.

MTT viability assay

SGHPL-4 cells were seeded in 96-well plate in 100µl of HAM´s F10 medium (Biochrom) with 10% (v/v) FCS, 1% AA and incubated for 48h in 37^{o}C, 5% CO_2. Medium was removed

and replaced with 80µl fresh medium and 20µl of a yellow tetrazole MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) solution (4.15 mg/ml) (Sigma) was added. After 3 h of incubation in 37°C, 5% CO_2 medium was removed and 50µl of DMSO was added to dissolve the insoluble purple formazan. The absorbance was measured at 570 nm by a spectrophotometer.

Mitochondrial assay

To determine the effect of DLX5 overexpression on the mitochondrial respiratory capacity XF Cell Mito Stress Test Kit and XF-24 Extracellular Flux Analyzer (Seahorse Biosciences) was used. Prior to day of mitochondrial assay, sensor cartridge with 1 ml/well XF24 calibrant solution on one utility plate was pre-incubated in a 37°C non-CO_2 incubator overnight. On the day of the assay, SGHPL-4 cells were seeded in the Mito Stress XF24 culture plate in 100µl HAM`s F10 medium with 10% (v/v) FCS and 1% AA and allowed to attach for 4 h (after 1h additional 150µl of medium was added). XF basal medium was supplemented with 6.1 mM D-glucose, 1 mM Na-pyruvate and 0.5% FCS, pH was adjusted to 7.4 and the complete XF medium was sterile filtrated. Next mitochondrial drugs were diluted in the complete XF medium at concentration as follow: 0.75 µM oligomycin, 1 µM FCCP, 1 µM antimycin A + 0.1 µM rotenone and injected into port A, B and C in the pre-incubated sensor cartridge, respectively, and loaded into XF24 Extracellular Flux Analyzer instrument for calibration. In the meantime HAM`s F10 medium in SGHPL-4 cells was replaced with 500 µl of complete XF medium and cells were incubated in a 37°C non-CO_2 incubator for equilibration. After calibration was completed, the utility plate was replaced with pre-incubated cell plate and the mitochondrial activity was measured. Five consecutive measurements were obtained under basal conditions and after the sequential compounds injection as follow:

Command	Time (min)
Calibrate	Fixed
Equilibrate	Fixed

Loop Start Mix Wait Measure Loop End	5x 2 2 3
Inject from port A Loop Start Mix Wait Measure Loop End	 5x 2 2 3
Inject from port B Loop Start Mix Wait Measure Loop End	 5x 2 2 3
Inject from port C Loop Start Mix Wait Measure Loop End	 5x 2 2 3
Program End	

After the run was completed, cells were collected and counted using high throughput sampler (HTS) on BD FACSCanto II system using BD FACSDiva software for normalization. Data was analyzed using FlowJo software. The mitochondrial assay was run in four independent replicates.

ROS production

Wild-type and OE-DLX5 SGHPL-4 cells were seeded in 24-well plate in 1 ml of Ham´s F10 medium with 10% (v/v) FCS, 1% AA and incubated for 24h in 37oC, 5% CO2. Cells were labelled with 0.05mM DCFH-DA for 45 min. After incubation, cells were washed 3 times with PBS, trypsinized and GFP signal was analyzed on FACSCalibur system using CellQuest Software. Data is presented as a mean fluorescent intensity (MFI) of the GFP signal.

ER stress

The degree of ER stress is directly proportional to the degree of oxidative stress. We applied the *in vitro* model of oxidative cellular stress that was developed by Yung et al. using newly derived BeWo cells, BeWo-NG, which grow at physiological concentration of glucose[34]. In short, different degrees of oxidative cellular stress were induced. For the mild and severe oxidative stress, BeWo cells were incubated in serum-free medium (modified DMEM/F12, containing 5.5 mM glucose) under repetitive 6 h cyclic conditions between 5% and 20% O_2 (5/20 H/R) or 1% and 20% O_2 (1/20 H/R), respectively (ExVivo, BioSpherix, USA) for 24 h. Constant 20% O_2 was used as a normoxic control. The severity of ER stress ranks from low to high as follows: 5% O_2 < 1% O_2 < 5%/20% O_2 HR < 1%/20% O_2 HR. After 24h cells were lysed and the DLX5 protein expression was checked by Western blot. To quantify the DLX5 expression in different conditions, ImageJ software was used to measure the intensity of the band.

BeWo cell syncytialization

To study the role of DLX5 in syncytialization, we cultured wild type BeWo cells in 6-well plates in DMEM/F12, 10% (v/v) FCS, 1% AA and incubated for 24h in 37oC, 5% CO_2. Next, we induced syncytium formation with 100μM Forskolin (Sigma). Cells were lysed after 24h and 48h of incubation using with 3X Sample Buffer (cooked in 95^{o}C for 7 min) and denaturized proteins were separated on SDS-PAGE gel. Proteins were then transferred to PDVF membrane using semi-dry blotting. Unspecific binding was blocked with 2.5% w/v non-fat dry milk in TBS-T. Primary antibodies: rabbit anti-DLX5 (1:1000, Sigma-Aldrich), rabbit anti-gCG (1:1000 Abcam), mouse E-cadherin (1:1000, BD Transduction) and mouse anti-β-actin (1:1000, Dianova) were incubated overnight at 4^{o}C. Secondary antibody goat anti-rabbit (1:5000, Thermo Scientific) and goat anti-mouse (1:5000, Thermo Scientific) against the host of primary antibody was used conjugated with horseradish peroxidase and

membranes were developed with Amersham ECL Prime Western Blotting Detection Reagent (GE Healthcare).

Statistics

Data is presented as either mean ± SEM (for normally distributed data) or median with inter-quartile range (for non-normally distributed data). Normality was assessed by Kolmogorov-Smirnov tests. Data sets were compared using the unpaired t test, nonparametric Mann–Whitney test, one-way ANOVA or Kruskal-Wallis test as appropriate. Post hoc testing included Bonferroni and Dunn´s tests for multigroup comparisons. Statistical correlation was performed using nonparametric 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.

Supplementary Tables

Table S1. Patient`s clinical data (provided as an Excel file).

Clinical data of patients subjected for clustering analysis. Table presents clinical and biomarker parameters, n = 36 (in blue are additional parameters used for analysis of PE patients). Table represents the mean ± SEM value of each parameter in Control and PE patients. PE patients were subdivided into EO-PE, LO-PE and PE+IUGR and the p-value was calculated for each group against Controls either using 2-way ANOVA for normally distributed data (Bonferroni multiple comparison test) or Kruskal-Wallis (Dunn`s multiple comparison test) for not normally distributed data.

Table S2. Ingenuity Pathway Analysis of DEGs in three PE clusters (provided as an Excel file).

Table S3. Imprinted gene expression analysis in PE (provided as an Excel file).

Table S4. Primers and probes

	Gene		Primer	Sequence
RT-qPCR	18S		Forward	5'-ACATCCAAGGAAGGCAGCAG-3'
			Reverse	5'-TTTTCGTCACTACCTCCCCG-3'
			Probe	5'-FAM-CGCGCAAATTACCCACTCCCGAC-TAMRA-3'
	hDLX5		Forward	5'-TCCGATCCGGCGACTTC-3'
			Reverse	5'-CCTGAGACGGATGGTGCATA-3'
			Probe	5'-FAM-AAGCTCCGTTCCAGACGTCCGCA-TAMRA-3'
LOI	Genotyping	hDLX5amp (+)	Forward	5`-AGACTGGGAGTCGTGAAGT-3`
			Reverse	5`-ATTGTCCCCAGACATCACAGA-3`
	Amplicon	LOIhDLX5amp	Forward	5`-ACAGATAGACTAAGACCCCTCCCCAC-3`
			Reverse	5`-CGGGGTGGATCTGGTTCTATTGGC-3`
	qASPCR	qAShDLX5-c	Forward	5`-CCACAACAAGCAAAAACACACACAAGC-3`
		qAShDLX5-t	Forward	5`- CCACAACAAGCAAAAACACACACAAGT-3`
		LOIhDLX5amp	Reverse	5`-CGGGGTGGATCTGGTTCTATTGGC-3`

DLX5 cloning	**hDLX5_EcoRV**	Forward	5´-atcgatatcATGACAGGAGTGTTTGACAG-3´
	hDLX5_NotI	Reverse	5´-atcgcggccgcCTAATAGAGTGTCCCGGAG-3`

Table S5. CpG methylation at the DLX5 locus in control and PE placenta samples (provided as an Excel file).

Table S6. Ingenuity Pathway Analysis of DEGs upon DLX5 overexpression (provided as an Excel file).

Table S7. Imprinted gene expression analysis during human embryogenesis (provided as an Excel file).

Supplementary Figures:

Fig. S1. Cluster analysis of microarray gene expression data.

(A) Clustering analysis based on Euclidean distances of microarray data identified three PE groups in placenta: PE_P1 (blue), PE_P2 (yellow), PE_P3 (green). (Control placenta n = 22, PE placenta n = 24). **(B)** Clustering analysis of decidual gene expression (Control placenta n = 22, PE placenta n = 24).

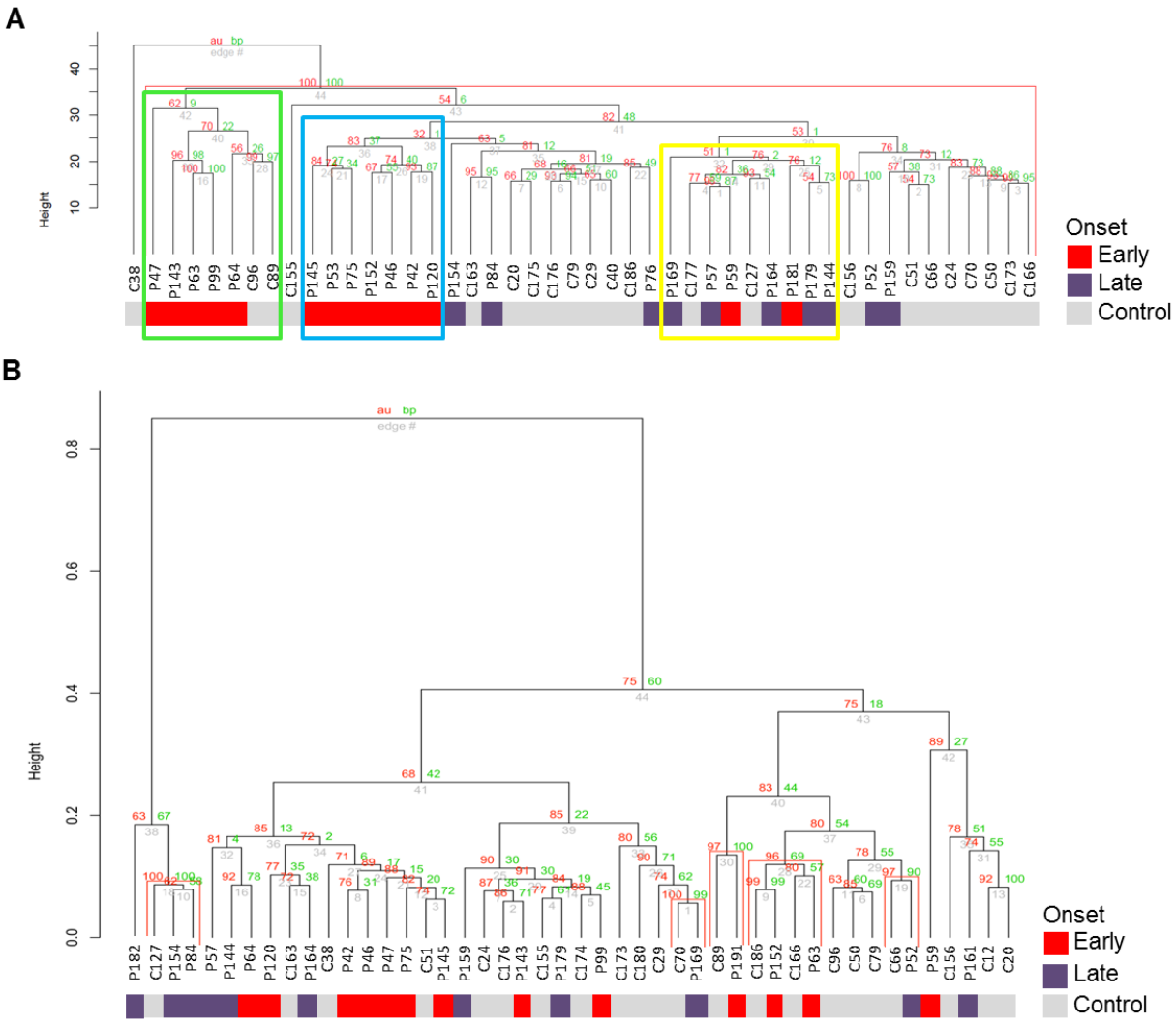
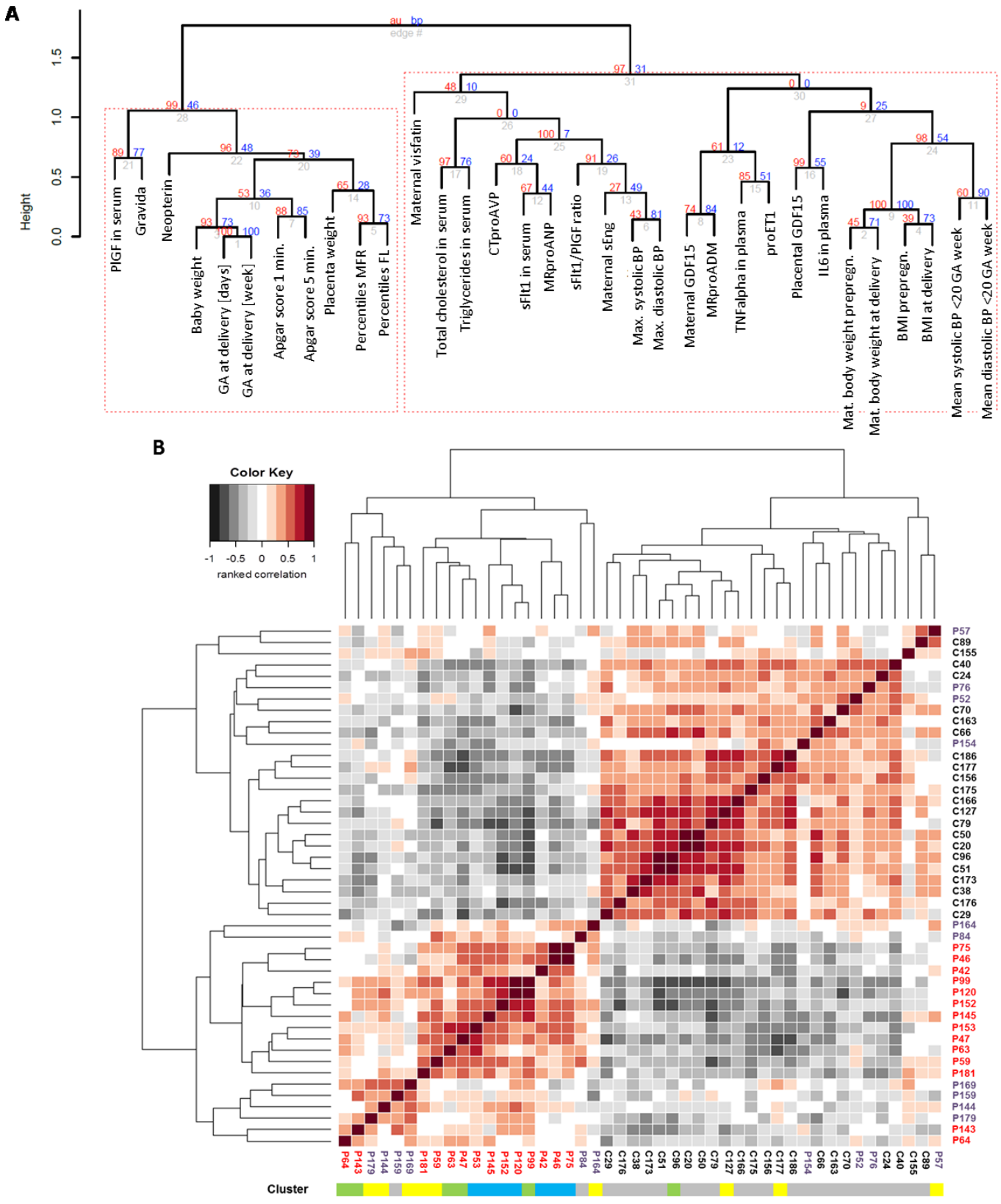


Fig. S2. Clustering analysis of clinical data.

(A) Hierarchical cluster dendrogram calculated using ranked correlation and complete linkage method on relative values from the clinical variables (n = 33). Height of dendrograms represents the Euclidian distance. Numbers in red and blue indicate au and bp values from bootstrapping. **(B)** Heatmap representing ranked correlation between samples based on the clinical data for both Controls and Patients. Most of the Patients and Controls were separated. Only four LO-PE patients were grouped with Controls. Color code as for gene expression clusters: PE_P1 – blue, PE_P2 – yellow, PE_P3 – green. **(C)** Genes associated with the clinical phenotype. Pairwise ranked correlation was calculated for each clinical phenotype with each gene across the samples. Shown pairs are those that passed the correlational threshold of |0.6|.


A
au bp
edge #
Height
0.0
0.5
1.0
1.5
PlGF in serum
Gravida
Neopterin
Baby weight
GA at delivery [days]
GA at delivery [week]
Apgar score 1 min.
Apgar score 5 min.
Placenta weight
Percentiles MFR
Percentiles FL
Maternal visfatin
Total cholesterol in serum
Triglycerides in serum
CTproAVP
sFlt1 in serum
MRproANP
sFlt1/PlGF ratio
Maternal sEng
Max. systolic BP
Max. diastolic BP
Maternal GDF15
MRproADM
TNFalpha in plasma
proET1
Placental GDF15
IL6 in plasma
Mat. body weight prepregn.
Mat. body weight at delivery
BMI prepregn.
BMI at delivery
Mean systolic BP <20 GA week
Mean diastolic BP <20 GA week
B
Color Key
-1 -0.5 0 0.5 1
ranked correlation
Cluster

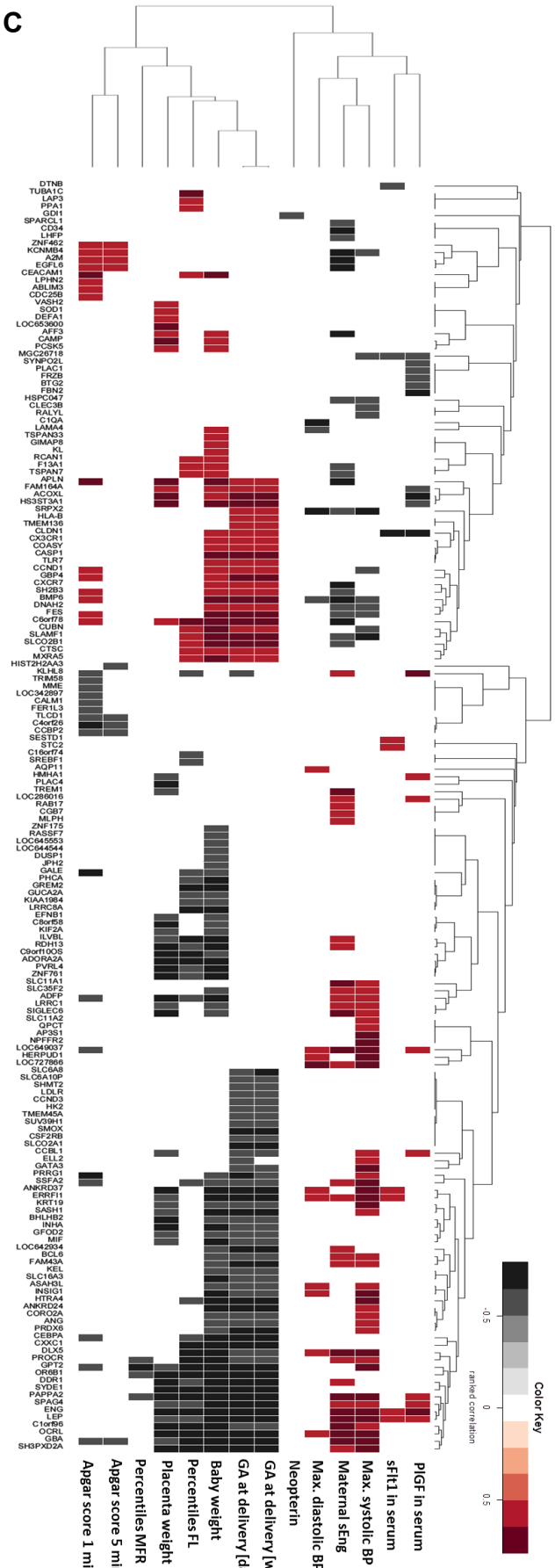

C
PlGF in serum
sFlt1 in serum
Max. systolic BP
Maternal sEng
Max. diastolic BP
Neopterin
GA at delivery [week]
GA at delivery [days]
Baby weight
Percentiles.FL
Placenta weight
Percentiles.MFR
Apgar score 5 min.
Apgar score 1 min.
Color Key
ranked correlation
-0.5
0
0.5

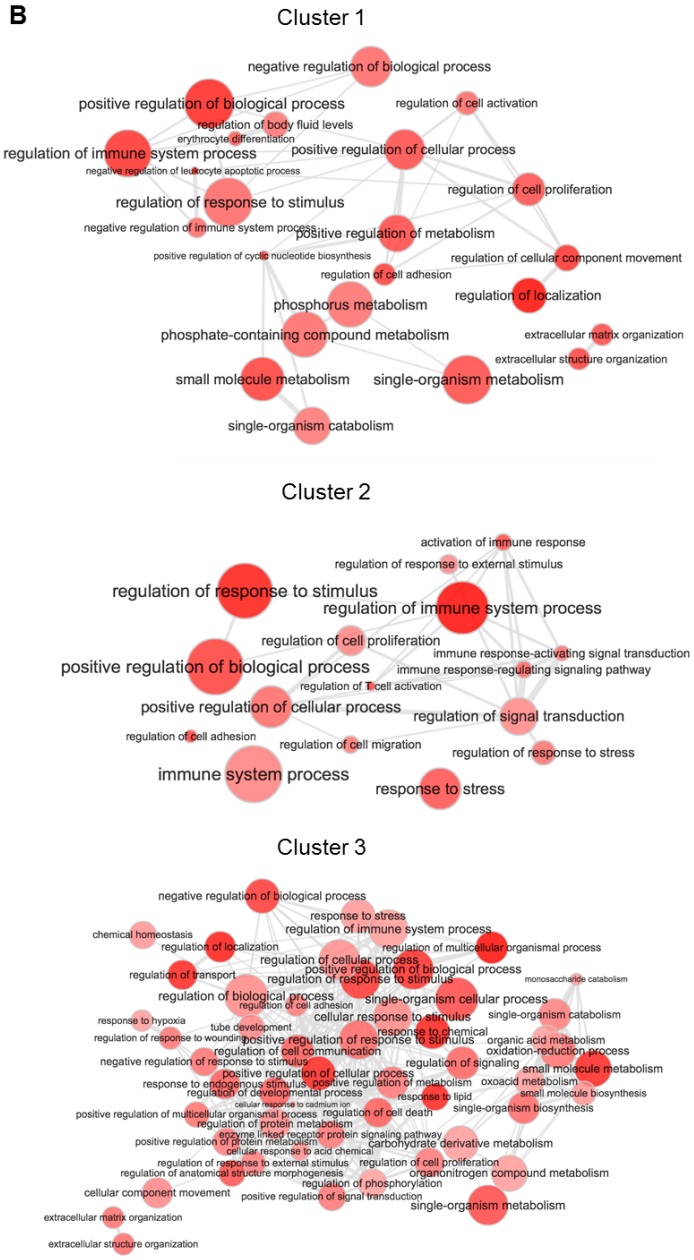
Fig. S3. Molecular characteristics of the three types of PE.

(A) Canonical pathway summary using Ingenuity Pathways Analysis (IPA). The top 5 canonical pathways overrepresented in each cluster of PE placenta. **(B)** Gene ontology enrichment analysis using GOrilla. Presented are gene ontology terms overrepresented in the PE clusters. The size of the circle and intensity of color depends on the enrichment and p-value, respectively.

A

Ingenuity Canonical Pathways	-log(B-H p-value)	Ratio
Cluster 1		
ILK Signaling	3.56	0.285
Hepatic Fibrosis / Hepatic Stellate Cell Activation	3.56	0.279
Axonal Guidance Signaling	2.56	0.222
LXR/RXR Activation	2.33	0.289
p38 MAPK Signaling	2.33	0.291
Cluster 2		
IL-8 Signaling	3.33	0.23
Axonal Guidance Signaling	3.33	0.182
Ephrin Receptor Signaling	3.33	0.23
Breast Cancer Regulation by Stathmin1	3.17	0.22
Androgen Signaling	2.91	0.252
Cluster 3		
NRF2-mediated Oxidative Stress Response	2.51	0.306
G-Protein Coupled Receptor Signaling	2.08	0.27
Axonal Guidance Signaling	2.08	0.245
Gαi Signaling	2.08	0.317
Gαq Signaling	2.08	0.299

B



The analysis revealed that Cluster 1 is enriched in genes involved in ILK signaling, hepatic fibrosis and hepatic stellate cell activation, LXR/RXR activation and p38 MAPK signaling. Among the biological processes regulation of localization and locomotion, negative regulation of leukocyte apoptotic process, regulation of

immune system and regulation of cell motility and cell adhesion were the most overrepresented terms. Cluster 2 is enriched in immune system process and response, regulation of response to stimulus, cell adhesion and T cell activation. Among the top dysregulated pathways, IL-8 signaling, ephrin receptor signaling, androgen signaling and pathway of breast cancer regulation by stathmin-1 are the most significant. Genes identified in Cluster 3 featured dysregulated oxidative stress responses and endothelial dysfunction as the identified enriched genes resided in the NFR2-mediated oxidative stress response and G-protein coupled receptor signaling, as well as Gαi signaling and Gαq signaling. According to the gene ontology analysis the most overrepresented terms are: regulation of multicellular organismal process, regulation of localization and secretion and transport, and response to chemical and lipid. Axonal guidance signaling was the top pathway overrepresented in all three PE clusters.

Fig. S4. DLX5 expression *in vivo*.

(A) Microarray validation of DLX5 overexpression using qPCR. PE samples were subdivided into PE+IUGR and PE without IUGR. Data is presented as median ± IQR (Control 0.765 ± 0.6175-0.9125, n = 26; PE 1.06 ± 0.78-1.25, n = 19; PE+IUGR 1.14 ± 0.88-1.79, n = 7; *p < 0.05, Kruskal-Wallis test, Dunn's multiple comparisons test). **(B)** qPCR validation of DLX5 overexpression in the samples subjected for the microarray analysis according to clusters. Data presented as mean ± SEM (Control Cluster 1 and 2: 0.962 ± 0.045, n = 16; PE Cluster 1: 1.434 ± 0.156, n = 7; PE Cluster 2: 1398 ± 0.186, n = 7; PE Cluster 3: 1.328 ± 0.089, n = 5; *p < 0.05; ANOVA, Bonferroni's multiple comparisons test). **(C)** Western blot analysis of DLX5 expression in Control (n = 10) and PE (n = 10) placenta tissues. Actin was used as a loading control. **(D)** Western blot quantification of DLX5 expression in placenta tissues using ImageJ software. Relative intensity: DLX5/Actin (p=0.0414; Unpaired t-test). **(E)** Dependence between DLX5 expression in the placenta and gestational age in Control and PE samples (Control, n = 28; PE, n = 50, Spearman rank correlation). **(F)** DLX5 expression across different tissues (maternal muscle and fat tissue, decidua, placenta, primary human trophoblasts and fetal macrophages). Data presented as a mean ± SEM; TBs – trophoblasts. **(E)** DLX5 mRNA expression in the decidua of the second cohort shows that DLX5 transcript is increased in late PE. Data shown as mean ± SEM (Control: 1.008 ± 0.115, n = 36; EO-PE+IUGR: 1.338 ± 0.22, n = 8; LO-PE+IUGR: 1.092 ± 0.2, n = 5; EO-PE: 0.916 ± 0.13, n = 13; LO-PE: 1.677 ± 0.28, n = 15) (*p < 0.05; ANOVA, Bonferroni's multiple comparisons test).

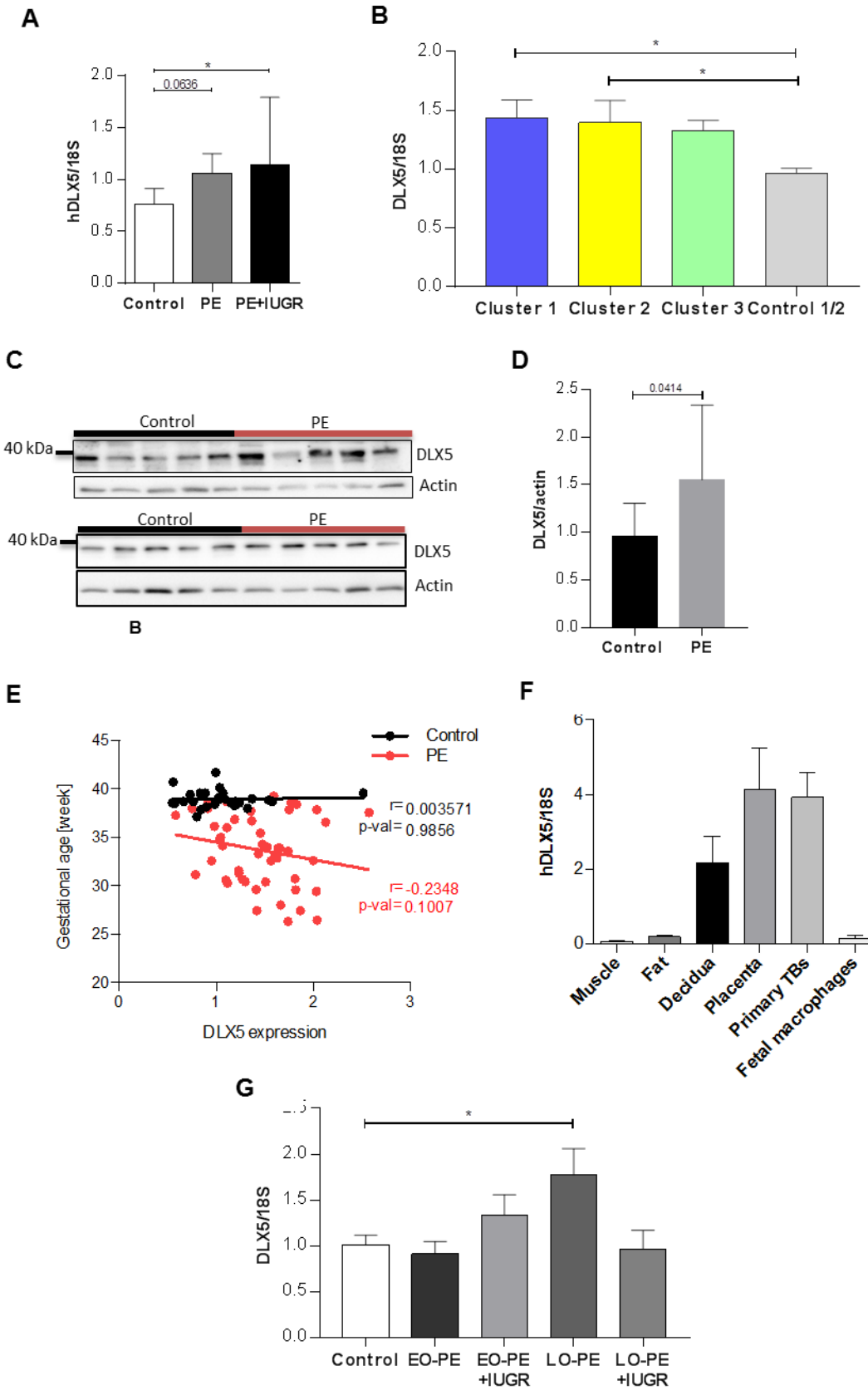

A
hDLX5/18S
2.0
1.5
1.0
0.5
0.0
0.0636
*
Control
PE
PE+IUGR
B
DLX5/18S
2.0
1.5
1.0
0.5
0.0
*
*
Cluster 1
Cluster 2
Cluster 3
Control 1/2
C
Control
PE
40 kDa
DLX5
Actin
Control
PE
40 kDa
DLX5
Actin
B
D
DLX5/actin
2.5
2.0
1.5
1.0
0.5
0.0
0.0414
Control
PE
E
Gestational age [week]
45
40
35
30
25
20
0
1
2
3
DLX5 expression
Control
PE
r=0.003571
p-val=0.9856
r=-0.2348
p-val=0.1007
F
hDLX5/18S
6
4
2
0
Muscle
Fat
Decidua
Placenta
Primary TBs
Fetal macrophages
G
DLX5/18S
2.0
1.5
1.0
0.5
0.0
*
Control
EO-PE
EO-PE +IUGR
LO-PE
LO-PE +IUGR

Fig. S5. Overexpression of human DLX5 protein *in vitro*.

(A) Using *Sleeping Beauty* transposon system human DLX5 gene was stably overexpressed in SGHPL-4 cells. **(B)** DLX5 overexpression was confirmed on mRNA by qPCR (n=6, **p=0.0044, Unpaired t-test) and **(C)** protein level by western blot (n=3). **(D)** Immunofluorescence staining of wild-type cells and OE-DLX5 in SGHPL-4 cells. Predominant nuclear localization of DLX5 is observed.

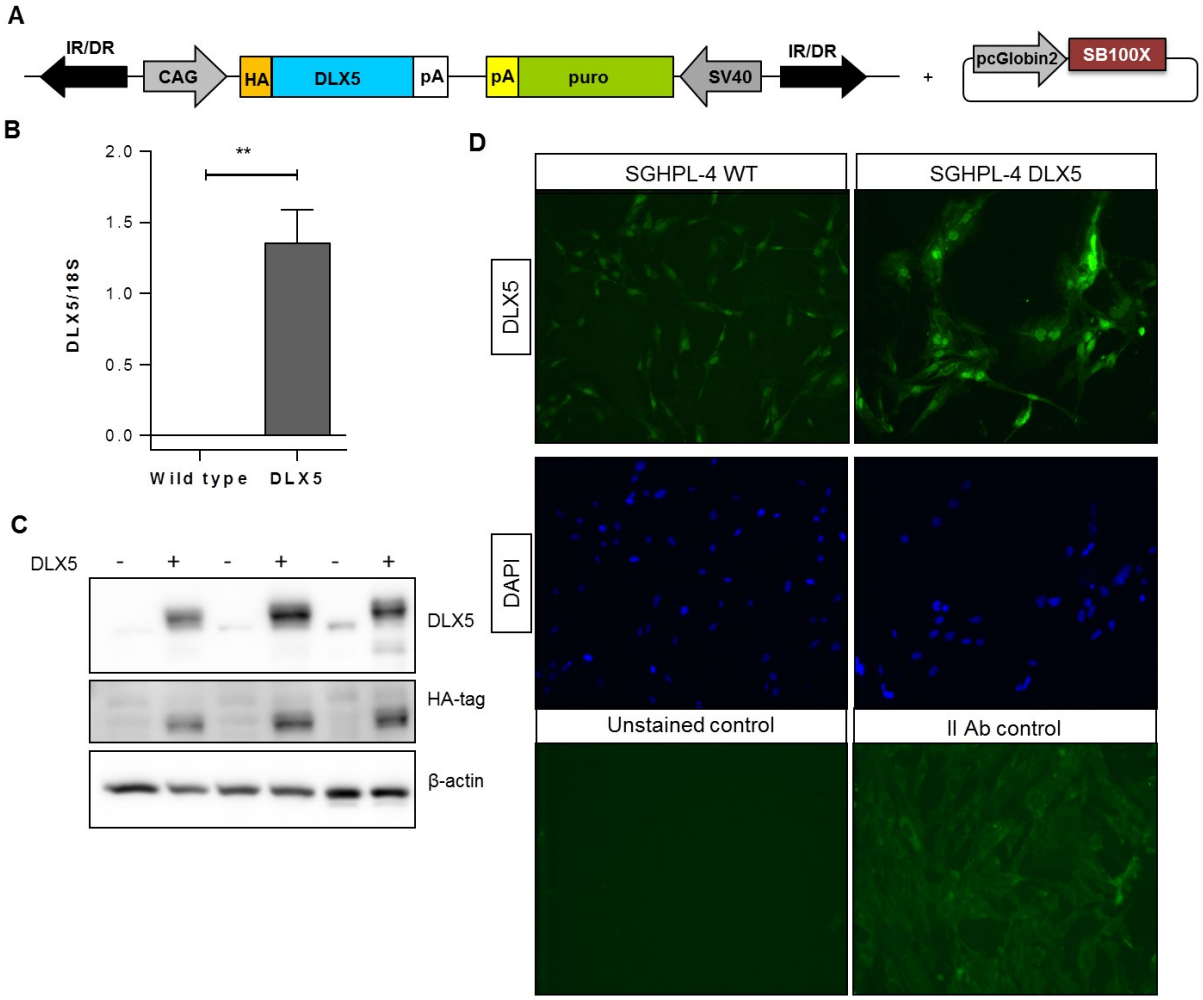
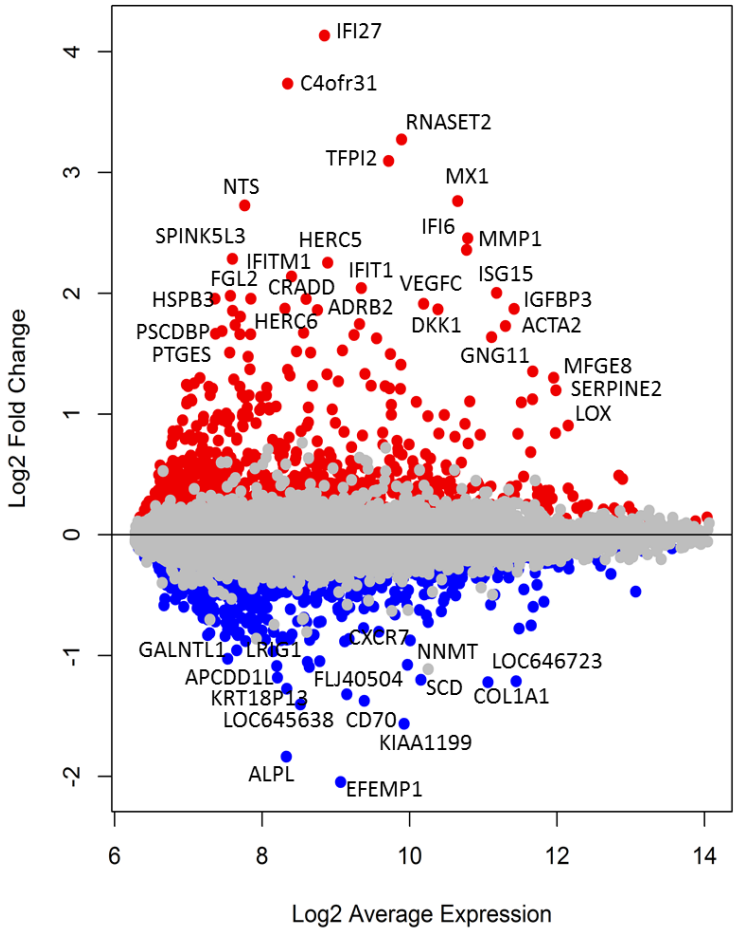


Fig. S6. Transcriptome analysis of SGHPL-4 cells overexpressing human DLX5 protein. **(A)** Fold change of DEGs on the $\log_2$ scale in OE-DLX5 trophoblasts at a significant level of p-value < 0.05; each point represents the $\log_2$ fold change of genes (y-axis) plotted against the $\log_2$ average expression (x-axis); 1897 genes were upregulated (red), 1753 genes were downregulated (blue); n = 6. **(B)** Molecular and cellular functions and physiological activities enriched for genes whose expression is altered in OE-DLX5 cells. Presented are the top five scoring hits in these categories using IPA, which includes the significance scores (p-values) and the number of genes included in each class.

A



B

Name	P-value	Number of molecules
Molecular and cellular functions		
Cellular growth and proliferation	$7.25 \times 10^{-5} - 6.52 \times 10^{-26}$	885
Cell death and survival	$5.39 \times 10^{-5} - 3.28 \times 10^{-19}$	833
Cellular movement	$9.14 \times 10^{-5} - 8.94 \times 10^{-19}$	546
Cellular development	$1.00 \times 10^{-4} - 1.40 \times 10^{-16}$	844
Gene expression	$9.20 \times 10^{-6} - 1.27 \times 10^{-10}$	554
Physiological system development and function		
Organismal survival	$9.48 \times 10^{-6} - 6.72 \times 10^{-17}$	640
Cardiovascular system development and function	$9.30 \times 10^{-5} - 3.14 \times 10^{-13}$	380
Organismal development	$9.75 \times 10^{-5} - 3.14 \times 10^{-13}$	846
Tissue development	$1.00 \times 10^{-4} - 5.25 \times 10^{-12}$	722
Embryonic development	$9.75 \times 10^{-5} - 3.61 \times 10^{-11}$	534

Fig. S7. Intersection of PE transcriptomes with the DLX5^{high} transcriptome. **(A)** DLX5 log_2 relative expression across healthy (blue) and PE (red) placenta samples subjected to the microarray. The placenta samples from patients displaying high (Patients DLX5_high) and low (Patients DLX5_low) DLX5 expression level are indicated. **(B)** Top 6 DLX5-high (Patient_high) and 6 DLX5-low (Patient_low) expressing PE placenta samples were further clustered with WT and OE-DLX5 (DLX5^{high}) SGHLP-4 cells. Heatmap represents Z-score of gene expression across samples (Spearman correlation and distances between observations were calculated using Euclidian distances and average linkage).

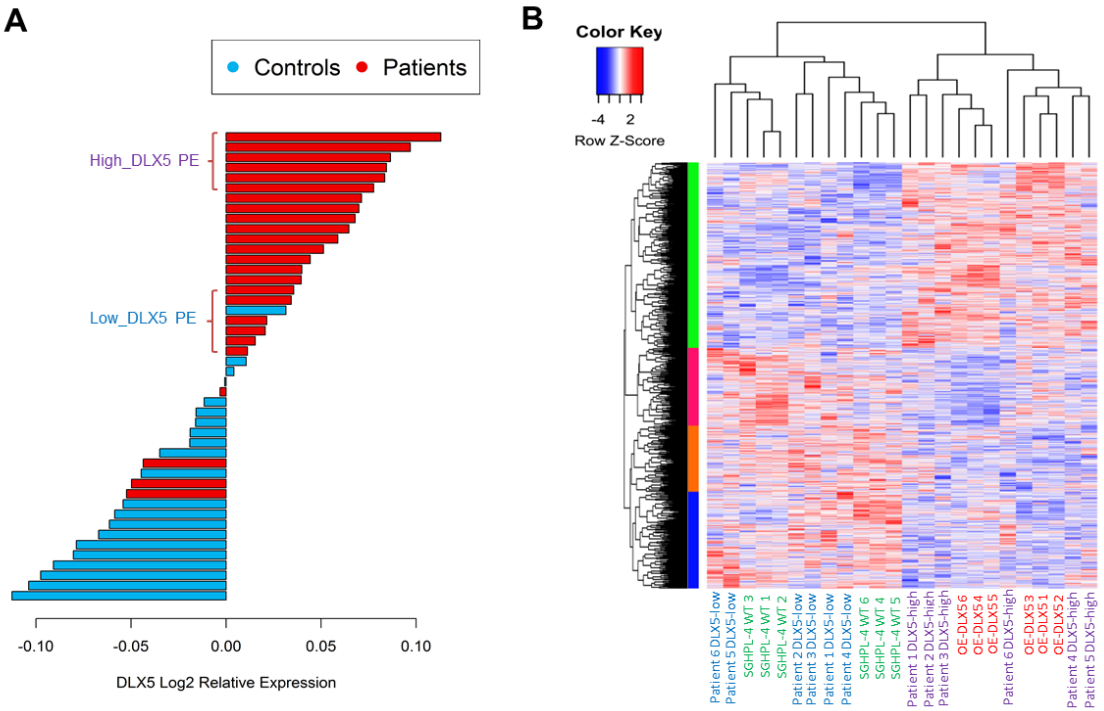


Fig. S8. Effect of DLX5 overexpression on SGHPL-4 cell apoptosis.

(A) DLX5 overexpression slightly increased cell apoptosis as measured by scoring cells undergoing apoptosis over 48h of incubation (not significant, 2-Way ANOVA, Bonferroni's multiple comparisons test). **(B)** After 48h of incubation cell apoptosis is significantly increased after TNFα treatment (conc. 30ng/ml) (DLX5 median 25, IQR: 23.13-31.88 vs. wild type median 12.50, IQR: 6.875-17.50) (*p<0.05; Kruskal-Wallis test, Dunn's multiple comparisons test) but not in untreated cell (DLX5 median 18.75, IQR: 13.75-25 vs. wild-type median 8.75, IQR: 4.375-15.63).

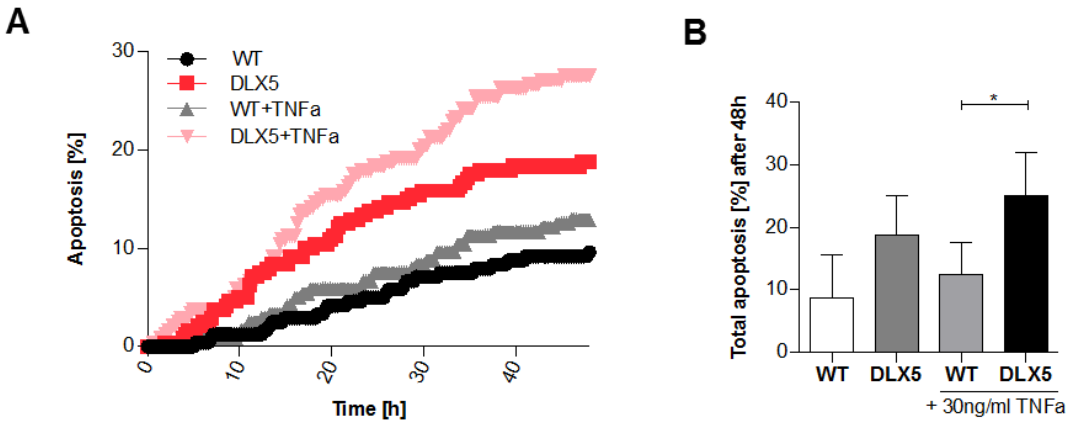


Fig. S9. DLX5 overexpression is associated with ER stress.

(A) DLX5 protein expression level upon induction of ER stress in BeWo cells. 20N – normoxia, 5/20 HR and 1/20 HR - cyclic condition of 6 hour of incubation in 5%/20% O_2 and 1%/20% O_2. **(B)** The transcriptome analysis of OE-DLX5 cells revealed alternation in unfolded protein response pathway (UPR). Shown are upregulated (red) and downregulated (green) UPR genes (at p-value < 0.05). **(C)** Venn diagram representing differentially expressed genes from the UPR pathway in all PE placenta samples and OE-DLX5 cells. 8 dysregulated genes are shared between PE and OE-DLX5 cells. **(D)** Western blot representing expression level of DLX5, E-cadherin, hCG in wild type BeWo cell line treated with 100uM Forskolin to induce syncytialization, either for 24h or 48h. Actin was used as a loading control. DLX5 expression increased upon Forskolin treatment in a time dependent manner.

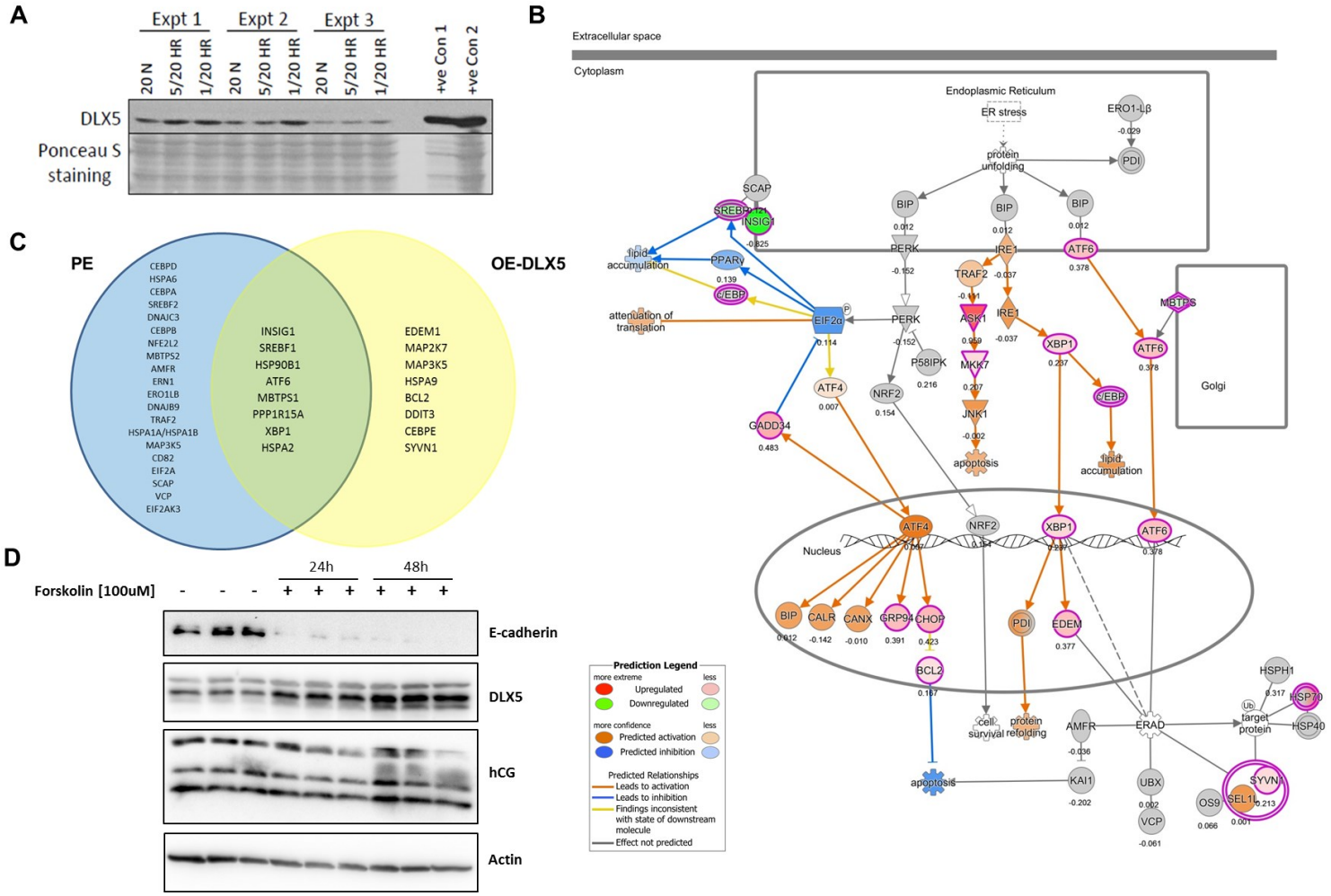
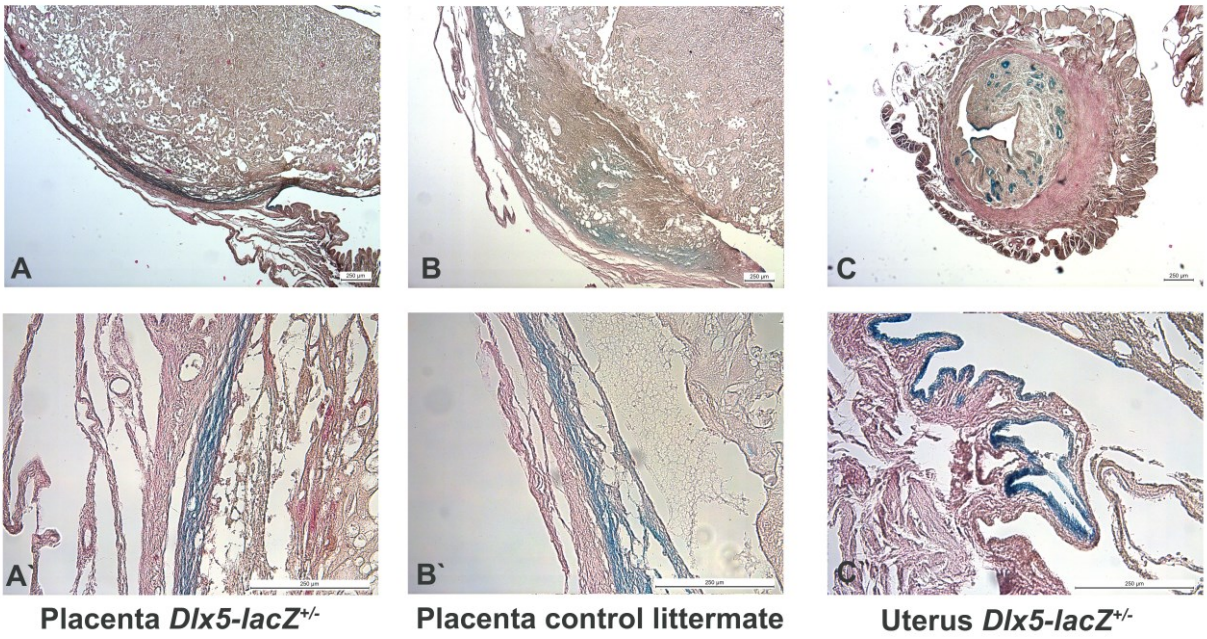


Fig. S10. Dlx5 expression pattern in mouse placenta.

Dlx5-lacZ$^{+/-}$ (A and A`) and control placenta (B and B`) sections were stained for *lacZ* expression. Uterus section of the *Dlx5-lacZ*$^{+/-}$ was used as a positive control. In the placenta weak LacZ staining in the Dlx5-LacZ+/-on the external muscular layer of the placenta, but as similarly, a weak positive staining was also present in controls, we did not consider it. Representative staining are shown.



Supplemental References:

1. Herse F, Lamarca B, Hubel CA, Kaartokallio T, Lokki AI, Ekholm E, Laivuori H, Gauster M, Huppertz B, Sugulle M, Ryan MJ, Novotny S, Brewer J, Park JK, Kacik M, Hoyer J, Verlohren S, Wallukat G, Rothe M, Luft FC, Muller DN, Schunck WH, Staff AC, Dechend R. Cytochrome p450 subfamily 2j polypeptide 2 expression and circulating epoxyeicosatrienoic metabolites in preeclampsia. *Circulation*. 2012;126:2990-2999
2. Sugulle M, Herse F, Seiler M, Dechend R, Staff AC. Cardiovascular risk markers in pregnancies complicated by diabetes mellitus or preeclampsia. *Pregnancy Hypertens*. 2012;2:403-410
3. Braekke K, Holthe MR, Harsem NK, Fagerhol MK, Staff AC. Calprotectin, a marker of inflammation, is elevated in the maternal but not in the fetal circulation in preeclampsia. *American journal of obstetrics and gynecology*. 2005;193:227-233
4. Przybyl L, Haase N, Golic M, Rugor J, Solano ME, Arck PC, Gauster M, Huppertz B, Emontzpohl C, Stoppe C, Bernhagen J, Leng L, Bucala R, Schulz H, Heuser A, Weedon-Fekjaer MS, Johnsen GM, Peetz D, Luft FC, Staff AC, Muller DN, Dechend R, Herse F. Cd74-downregulation of placental macrophage-trophoblastic interactions in preeclampsia. *Circulation research*. 2016;119:55-68
5. Burton GJ, Sebire NJ, Myatt L, Tannetta D, Wang YL, Sadovsky Y, Staff AC, Redman CW. Optimising sample collection for placental research. *Placenta*. 2014;35:9-22
6. Staff AC, Ranheim T, Khoury J, Henriksen T. Increased contents of phospholipids, cholesterol, and lipid peroxides in decidua basalis in women with preeclampsia. *American journal of obstetrics and gynecology*. 1999;180:587-592
7. Du P, Kibbe WA, Lin SM. Lumi: A pipeline for processing illumina microarray. *Bioinformatics (Oxford, England)*. 2008;24:1547-1548
8. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. Limma powers differential expression analyses for rna-sequencing and microarray studies. *Nucleic acids research*. 2015;43:e47
9. Suzuki R, Shimodaira H. Pvclust: An r package for assessing the uncertainty in hierarchical clustering. *Bioinformatics (Oxford, England)*. 2006;22:1540-1542
10. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol*. 2004;3:Article3
11. Hamada H, Okae H, Toh H, Chiba H, Hiura H, Shirane K, Sato T, Suyama M, Yaegashi N, Sasaki H, Arima T. Allele-specific methylome and transcriptome analysis reveals widespread imprinting in the human placenta. *The American Journal of Human Genetics*. 2016;99:1045-1058
12. Lambertini L, Diplas AI, Lee MJ, Sperling R, Chen J, Wetmur J. A sensitive functional assay reveals frequent loss of genomic imprinting in human placenta. *Epigenetics*. 2008;3:261-269
13. Blair JD, Yuen RK, Lim BK, McFadden DE, von Dadelszen P, Robinson WP. Widespread DNA hypomethylation at gene enhancer regions in placentas associated with early-onset pre-eclampsia. *Molecular human reproduction*. 2013;19:697-708
14. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA. Minfi: A flexible and comprehensive bioconductor package for the analysis of infinium DNA methylation microarrays. *Bioinformatics (Oxford, England)*. 2014;30:1363-1369

15. Petropoulos S, Edsgard D, Reinius B, Deng Q, Panula SP, Codeluppi S, Plaza Reyes A, Linnarsson S, Sandberg R, Lanner F. Single-cell rna-seq reveals lineage and x chromosome dynamics in human preimplantation embryos. *Cell*. 2016;165:1012-1026
16. Deng Q, Ramsköld D, Reinius B, Sandberg R. Single-cell rna-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science (New York, N.Y.)*. 2014;343:193-196
17. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. Star: Ultrafast universal rna-seq aligner. *Bioinformatics (Oxford, England)*. 2013;29:15-21
18. Liao Y, Smyth GK, Shi W. Featurecounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics (Oxford, England)*. 2014;30:923-930
19. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical bayes methods. *Biostatistics (Oxford, England)*. 2007;8:118-127
20. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics (Oxford, England)*. 2012;28:882-883
21. Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. *Nature biotechnology*. 2015;33:495-502
22. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, Trombetta JJ, Weitz DA, Sanes JR, Shalek AK, Regev A, McCarroll SA. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell*. 2015;161:1202-1214
23. Kharchenko PV, Silberstein L, Scadden DT. Bayesian approach to single-cell differential expression analysis. *Nature methods*. 2014;11:740-742
24. Blakeley P, Fogarty NME, del Valle I, Wamaitha SE, Hu TX, Elder K, Snell P, Christie L, Robson P, Niakan KK. Defining the three cell lineages of the human blastocyst by single-cell rna-seq. *Development*. 2015;142:3151-3165
25. Biase FH, Cao X, Zhong S. Cell fate inclination within 2-cell and 4-cell mouse embryos revealed by single-cell rna sequencing. *Genome research*. 2014;24:1787-1796
26. Nakamura T, Okamoto I, Sasaki K, Yabuta Y, Iwatani C, Tsuchiya H, Seita Y, Nakamura S, Yamamoto T, Saitou M. A developmental coordinate of pluripotency among mice, monkeys and humans. *Nature*. 2016;537:57-62
27. Yan L, Yang M, Guo H, Yang L, Wu J, Li R, Liu P, Lian Y, Zheng X, Yan J, Huang J, Li M, Wu X, Wen L, Lao K, Li R, Qiao J, Tang F. Single-cell rna-seq profiling of human preimplantation embryos and embryonic stem cells. *Nature structural & molecular biology*. 2013;20:1131-1139
28. Andrews S. Fastqc a quality control tool for high throughput sequence data. *Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc*. 2010
29. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for rna-seq data with deseq2. *Genome biology*. 2014;15:550
30. Feng J, Meyer CA, Wang Q, Liu JS, Shirley Liu X, Zhang Y. Gfold: A generalized fold change for ranking differentially expressed genes from rna-seq data. *Bioinformatics (Oxford, England)*. 2012;28:2782-2788
31. Cartwright JE, Holden DP, Whitley GS. Hepatocyte growth factor regulates human trophoblast motility and invasion: A role for nitric oxide. *British journal of pharmacology*. 1999;128:181-189
32. Langfelder P, Horvath S. Wgcna: An r package for weighted correlation network analysis. *BMC bioinformatics*. 2008;9:559

33. Dash PR, Cartwright JE, Whitley GS. Nitric oxide inhibits polyamine-induced apoptosis in the human extravillous trophoblast cell line sghpl-4. *Hum Reprod*. 2003;18:959-968
34. Yung HW, Atkinson D, Campion-Smith T, Olovsson M, Charnock-Jones DS, Burton GJ. Differential activation of placental unfolded protein response pathways implies heterogeneity in causation of early- and late-onset pre-eclampsia. *The Journal of pathology*. 2014;234:262-276