# Predicting the Risk to Develop Preeclampsia in the First Trimester Combining Promoter Variant -98A/C of *LGALS13* (Placental Protein 13), Black Ethnicity, Previous PE, Obesity and Maternal Age

**Liora Madar-Shapiro[1], Ido Karady[1], Alla Trahtenherts[1], Argryo Syngelaki[2], Ranjit Akolekar[2], Ruth Cohen[2], Adi Sharabi-Nov[3], Berthold Huppertz[4], Marei Sammar[5], Kata Juhasz[6], Nandor Gabor Than[6,7], Zoltan Papp[7], Roberto Romero[8], Kypros H Nicolaides[2], Hamutal Meiri[1]**

**Short title**: Predicting preeclampsia by promoter polymorphism of *LGALS13* (PP13)

[1] Hy Laboratories, Rehovot, Israel

[2] Harris Birth Rights, King's College Hospital and Fetal Medicine Foundation, London, UK

[3]Ziv Medical Center, Safed and Tel-Hai Academic College, Israel

[4] Biobank Graz and Institute of Cell Biology, Histology and Embryology, Medical University of Graz, Graz, Austria

[5] *Prof. Ephraim Katzir* Department of Biotechnology Engineering, ORT Braude College, Karmiel, Israel

[6] Systems Biology of Reproduction Momentum Research Group, Institute of Enzymology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Budapest, Hungary

[7] Maternity Private Department, Kutvolgyi Clinical Block, Semmelweis University, Budapest, Hungary

[8] Perinatology Research Branch, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, U.S. Department of Health and Human Services (NICHD/NIH/DHHS)

**Corresponding Author**: Dr. Hamutal Meiri, Hy Laboratories, 6 Plaut St., 7670606, Rehovot, Israel, Tel: +972-45-7774762, Mail: hamutal62@hotmail.com

# Abstract (199))

BACKGROUND: We studied *LGALS13* [Placental Protein 13 (PP13)] promoter DNA polymorphisms in preeclampsia (PE) prediction, given PP13's effects on hypotension, angiogenesis and immunotolerance.

METHODS: We retrieved 67 PE (49 term, 18 preterm) cases and 196 matched controls from first trimester plasma samples prospectively collected at King's College Hospital, London. Cell-free DNA was extracted and the four *LGALS13* exons were sequenced after PCR amplification. Expression of *LGALS13* promoter reporter constructs were determined in BeWo trophoblast-like cells with luciferase assays.

RESULTS: A/C genotype in –98 position was the lowest in term PE compared to controls ($p<0.032$), similar to a South African cohort. Control but not all PE allele frequencies were in Hardy-Weinberg equilibrium ($p=0.036$). The Odds ratio for term PE calculated from prior risk, the A/A genotype and black ethnicity was 14 ($p<0.001$). In luciferase assays, the *LGALS13* promoter "-98A" variant had 13% ($p=0.04$) and 26% ($p<0.001$) lower expression than the "-98C" variant in non-differentiated and differentiated BeWo cells, respectively. After 48-hour differentiation, there was 4.55-fold increase in expression of "-98C" variant versus 3.85-fold of "-98A" variant ($p<0.001$).

CONCLUSION: Lower *LGALS13* (PP13) expression by the "-98A/A" genotype appears to impose higher risk to develop PE and could aid in PE prediction.

# Introduction

Preeclampsia (PE) is a pregnancy-specific disorder characterized by hypertension and proteinuria that occurs after 20 weeks of gestation [1,2]. According to the World Health Organization, PE is associated with significant maternal and perinatal mortality, and affects approximately 2-7% of all pregnant women [3]. Secondary PE symptoms include liver, kidneys, brain and the clotting system complication [4,5]. The disorder can exacerbate to stroke and convulsion (eclampsia) associated with severe morbidity and mortality [1-5].

Epidemiological studies have indicated that PE has strong familial pre-disposition across ethnic groups, socio-economic, and geographic origins [6]. According to the South African (SA) Department of Health, the frequency of PE in 2015 was higher in SA than in England, particularly among the black and colored ethnic groups compared to the white ones [7]. At King's College Hospital in London the incidence of PE is higher among black women attending prenatal screening and follow up compared to Caucasians [8]. According to Nakimuli et al. [6], there are different PE frequencies and sub-patterns between blacks from sub-Saharan or Caribbean origin in England and their counterparts of a Zulu origin in SA.

Studies have implied an association between carriers of certain gene mutations and the development of PE [9]. Potential association of PE frequency was found among blacks and other ethnic groups of carriers of certain gene mutations, chromosomal changes and single nucleotide polymorphisms (SNPs). Such nucleic acid variations are particularly found in the context of blood pressure regulation [9], vasodilatation [10], angiogenesis regulation [11], and inflammatory genes, especially those encoding for interleukins [12]. Specific gene polymorphisms are involved in the

development of human placental diseases [13] but their role is not fully elucidated yet [14].

Impaired placentation [15], loss of acquired maternal immunotolerance to the fetus [16], an intensified anti-angiogenic state [17], and inability of the cardiovascular system to adapt to the pregnancy burden [18-19] were listed among the leading causes that trigger PE. The search for genetic correlations has yielded examinations of polymorphisms among gene products that were implicated in the above functional pathways. One of them is Placental Protein 13 (PP13) and its encoding gene [*LGALS13*, NM_013268.2] [16, 20]. *LGALS13* is located on chromosome 19q13 in close vicinity to additional genes of the galectin protein family [16, 20]. As a member of such galectin family, the protein product of *LGALS13,* PP13 has high affinity for sugar residues of glycoproteins, mainly annexin-II and actin-β and -γ [16, 20-21]. PP13 is predominantly expressed at the outer layer of placental villi, on the apical brush border membrane of the syncytiotrophoblast facing maternal blood [16, 21-23]. The role of PP13 in the human placenta appears to be related to inflammation and immune defense [16, 23-24]. In recent years it was also discovered that when applied to gravid rodents PP13 decreases blood pressure and causes blood vessel expansion [25]. From the $5^{th}$ week of pregnancy, PP13 can be identified in maternal blood, where its mRNA levels can also be determined [26-27]. A meta-analysis of 18 studies has reported reduced first trimester levels of PP13 protein and its respective mRNA in women developing PE later during pregnancy [22].

The potential relations between the polymorphic variants of *LGALS13* and the involvement of the encoded PP13 protein were first described by Bruins et al. [28], showing multiple SNP variants and emphasizing the importance of a particular one at the -98 position in the promoter region of *LGALS13* [rs: 3764843]. They discovered

higher frequency of the A/A genotype versus A/C and C/C among women who developed term PE. Gebhardt et al. [14] have reported additional variants. They discovered a thymidine deletion at position 221 in exon three of *LGALS13* ($DelT_{221}$). Having the heterozygous deletion among black women in SA provided an 89% positive predictive value for the development of early-onset severe PE. Homozygous $DelT_{221}$ variant was not associated with a viable pregnancy.

The purpose of this study was to investigate polymorphisms in the *LGALS13* gene in a cohort enrolled for first trimester pregnancy evaluation at King's College Hospital, London. Our aim was to compare *LGALS13* polymorphisms in predicting PE. We also compared our data to the results published by the group from SA [14, 28]. To evaluate the effect of this polymorphism on LGALS expression we have constructed *LGALS13* promoter reporter clones with "A" or "C" in the -98 position and transfected them into BeWo human trophoblast-like cells. The expression was detected with the luciferase assays in order to assess the impact of the polymorphism on trophoblastic *LGALS13* expression *in vitro*.

# Material and Methods

**Study population:**

We used a prospective cohort of women attending their routine first hospital visit at King's College Hospital, London, at gestational weeks 11+0 to 13+6 between March 2006 and September 2009. The study was approved by the ethics committee of King's College Hospital. Women agreeing to participate provided written informed consent. Pregnancy age was determined by measurement of the fetal crown-rump length (CRL) [29].

We included pregnant women with viable singleton pregnancies who delivered live or a phenotypically normal stillbirth at or after 24 weeks of gestation. We excluded pregnancies with major fetal abnormalities and those ending in termination, miscarriage or fetal death before 24 weeks.

Samples of serum and plasma were drawn in the first trimester and stored at -80°C for subsequent analysis. The samples were tested for a large diversity of biochemical markers as detailed by Akolekar et al. [8], including serum PP13.

**Maternal history and characteristics:**

Patients were asked to complete a questionnaire on maternal age, racial origin (Caucasian, African, South Asian, East Asian and mixed), method of conception (spontaneous or assisted conception requiring the use of ovulation drugs), cigarette smoking during pregnancy (yes/no), substance abuse during pregnancy (yes/no), history of chronic hypertension (yes/no), history of type 1 or 2 diabetes mellitus (yes/no), family history of PE in the mother of the patient (yes/no) and obstetric history including parity (parous/nulliparous if no previous pregnancies at or after 24 weeks) and previous pregnancy with PE (yes/no). The questionnaire was then

reviewed by a doctor together with the patient, and maternal weight and height were measured [8]. Maternal mean arterial blood pressure (MAP) was measured by automated devices [30]. Trans-abdominal color Doppler ultrasound was used to visualize the left and right uterine artery and to measure the pulsatility index (PI) in each vessel and calculate the mean PI [31-32].

**Outcome measures:**

The definition of PE was according to the International Society for the Study of Hypertension in Pregnancy [33]. The systolic blood pressure should be 140 mm Hg or more and/or the diastolic blood pressure should be 90 mm Hg or more on at least two occasions 4 h apart, developing after 20 weeks of gestation in previously normotensive women. There should also be proteinuria of 300 mg or more in 24 h or two readings of at least $2^{+}$ on dipstick analysis of midstream or catheter urine specimens, if no 24h urine collection is available. PE cases superimposed on chronic hypertension were excluded [33]. Data on pregnancy outcome were collected from the hospital maternity records or the women's general medical practitioners. The obstetric records of all women with hypertension were examined to differentiate between gestational hypertension and chronic hypertension. For this study we only used cases of term PE (n=49; delivery at >37 weeks) and preterm PE (n=18; delivery at 34-37 weeks), but not early PE (delivery at <34 weeks) cases due to shortage of cases.

**Nested case-control study for biochemical markers:**

In the nested case-control study the cases were drawn from the study population as described above on the basis of availability of stored samples. The controls (n=196) were selected from pregnancies with no complications and normal outcome, and were

matched to the cases according to gestational week and storage time. None of the samples were previously thawed and refrozen.

**Serum PP13 testing:**

Serum PP13 was measured by DELFIA (Dissociation-Enhanced Lanthanide Fluorescent Immunoassay) using research reagents (PerkinElmer Life and Analytical Sciences, Turku, Finland) [8].

**Blood processing and DNA extraction:**

To study PP13 DNA polymorphisms, we used maternal plasma that was drawn from first trimester (gestational weeks 10-13) maternal peripheral veins and stored in ethylene-diamine-tetra-acetic acid (EDTA) tubes. Blood was centrifuged at 1,600 × g for 10 min at 4°C, and the supernatant plasma was aspired and re-centrifuged again at 13,000 × g for 10 min at 4°C to remove residual cells. Cell-free DNA was extracted from 0.5 ml of plasma using AccuPrep Genomic DNA extraction kit (Bioneer Corporation, Daejeon, South Korea) according to the manufacturer's instructions. DNA was eluted in 50µl double-distilled water (DDW).

We used Hylabs' recombinant clone that was constructed according to the website of the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/) to sequence the *LGALS13* (NM 013268.2) gene. Primer pairs for PCR amplification were designed to encompass the promoter and intronic sequences flanking the PP13 encoding regions of each *LGALS13* exon under investigation.

Primers were synthesized by the Hylabs' company oligo service unit using BH5 oligo making devise (Metabion, Planegg, Germany) and the Solid-Phase Oligo nucleotide Synthesis Method (ATDBio, University of South Hampton, the UK).

Primer sequences are shown in Figure 1A and were verified by the Multiple Primer Analyzer (ThermoFisher, Waltham, MA USA) and sequence specificity using NCBI Basic Local Alignment Search Tool (http://www.ncbi.nlm.nih.gov/BLAST/).

**[Place here Fig 1: Primer details and PCR products ]**

**PCR Reaction:**

PCR products were amplified according to *LGALS13* exons and the flanking introns using Takara Ex Taq Hot Start kit (Cat. #R006A, Takara Bio, Kusatsu, Shiga, Japan). Each reaction was prepared to a final volume of 50µl containing 10µl of pregnant women's plasma extracted genomic DNA (gDNA), and the Takara kit reagents including: 5µl of 10 × Taq buffer, 4µl of dNTPs mixture (2.5 mM each), 1µl primer pairs at 10 mM concentration, 0.25µl of Taq DNA polymerase (5 U/µl) and 30µl DDW. Amplifications were performed with the following sequential steps: denaturing at 98°C for 10 sec, followed by 40 cycles of amplification with initial denaturing at 94°C for 30 seconds, annealing of primers at corresponding temperature for 30 seconds and extension at 72°C for 60 seconds, and a final extension at 72°C for 10 minutes.

The PCR products (9 µl each) were then mixed with 1 µl of 10× loading buffer (Takara, Japan) and resolved by electrophoresis on 2% agarose gels in 1× TAE buffer (90 mM Tris–HCl, 90 mM boric acid and 1mM EDTA, pH 8.0) for 20 min to ensure a single PCR product (Figure 1B). The gels were then photographed under ultraviolet

light (260 nm) using a Bio-Rad Universal Hood and Bio-Rad Quantity One software (Bio-Rad, Hercules, CA, USA).

Sequencing reactions were carried out using the BigDye® Terminator V1.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) and the products were analyzed using an ABI 3730 × l Genetic Analyzer (Applied Biosystems, USA). Sequencing was bi-directional to verify accuracy.

Mutational analysis of the *LGALS13* gene was performed using the BioEdit Sequence Alignment Editor version 7.2.5 (Isis Pharmaceutical, Carlsbad, CA, USA) (Figure 2A). SNPs were verified by repeating the sequence and PCR procedures (Figure 2B).

**[Place Here Figure 2: Sequencing of the -98 Promoter Region Site ]**

**Transcription factor binding site analysis:**

The Transfac Database of the BIOBASE Biological Databases (www.biobase-international.com) was used to predict putative transcription factor binding sites in the promoter of *LGALS13* gene. The Positional Probability Matrix (PPM) and the Positional Weight Matrix (PWM) of the canonical TFAP2A (transcription factor AP-2 alpha, activating enhancer binding protein 2 alpha) binding site (MA0003.1) were downloaded from the Jaspar database (http://jaspar.genereg.net/).

**Promoter luciferase assays:**

To test the effect of the A/C variation in -98 position of the *LGALS13* promoter on gene expression, luciferase assays coupled with trophoblast differentiation experiments were completed following published protocols [34]. Briefly, the

*LGALS13* promoter reporter construct, which contains Gaussia luciferase reporter gene linked to the *LGALS13* promoter, was designed at the NIH Perinatology Research Branch and generated by GeneCopoeia (Rockville, MD, USA). This “-98C” clone contains a cytosine in the -98 position. An additional reporter construct (“-98A”) was generated at the Hungarian Academy of Sciences Research Centre for Natural Sciences by replacing cytosine for adenine in -98 position by site directed mutagenesis using the QuikChange Lightning kit (Agilent Technologies, Santa Clara, CA, USA) following the manufacturer´s protocol.

The "98C“ and "-98A“ reporter constructs (1 μg/well in a 24-well plate) were transfected into BeWo cells (American Type Culture Collection, Manassas, VA, USA) cultured in F12K medium (Sigma-Aldrich, St. Louis, MO, USA) supplemented with 10% fetal bovine serum (Gibco-ThermoFisher Scientific, Waltham, MA, USA) using Lipofectamine-2000 reagent (Invitrogen-ThermoFisher Scientific) according to the manufacturer´s protocol. Twenty four hours after transfection, BeWo cells were treated with either Forskolin (25 μM in DMSO; Sigma-Aldrich) or DMSO, and incubated at 37°C for 48h. In all experiments, supernatants were collected and secreted Gaussia luciferase activity was determined by Secrete-Pair Gaussia Luciferase Assay (Genecopoeia) according to the manufacturer’s protocol. This assay measured the Gaussia luciferase reporter gene's expression linked to the *LGALS13* promoter by measuring the luciferase activity of the secreted luciferase reporter protein, since secreted luciferase activity correlates with *LGALS13* promoter activity and luciferase gene expression. The luminescence was immediately measured with a Victor X3 microplate reader (PerkinElmer, Inc., Waltham, MA, USA).

**Statistical analysis:**

General - For categorical variables comparisons between each outcome group and unaffected controls were made by Fisher's exact test. Kruskal-Wallis or Mann Whitney non-parametric tests were used for continuous variables.

Hardy-Weinberg equilibrium and $\chi^2$ tests were used to compare the genotype and allelic frequency distribution in the study groups. Genotypes and alleles were considered to be in Hardy-Weinberg equilibrium if the observed frequencies did not differ significantly from the expected ($p>0.05$). The 95% confidence intervals (CI) were determined to verify the pattern of population distribution and overlap.

Luciferase assays - Data generated by luciferase assays were analyzed using the t-test for comparison between the constructs.

P-values of 5% or less were considered as significantly different. The data were analyzed using the SPSS version 24 (SPSS Inc., Chicago, IL, USA).

Odds Ratios (OR) for individual maternal risk to develop PE were calculated first on their own ad then adjusted given having the other ones. Multiple regression was then calculated according to all adjusted risk factors from the observed measures.

Likelihood ratio (LR)- Positive LR was calculated according to [sensitivity/(1-specificity)] and negative LR was calculated according to: [(1-sensitivity/Specificity]. Overall LR was the division of the positive and the negative LR.

Power analysis was conducted assuming the detection accuracy of promoter genotype polymorphism is 100%. Accordingly, we evaluated the frequency of the "A" allele in the control and PE groups. To yield the power of the study given 5% type 1 erro we entered these values into the PS Power and Sample Size Program (version 3.1.2) (http://ps-power-and-sample-size-calculation.software.informer.com)

# Results

We have identified 20 preterm PE cases (delivered at gestational age between 34 to 36 weeks and 6 days) and 50 cases of late PE (delivery >37 weeks). These were matched to 200 unaffected controls, according to the time of enrolment (± 1 week) and gestational age at enrolment (± 1 week). The amount of cell-free DNA extracted from the samples was sufficient to run the analysis for 196 controls, 18 preterm PE cases and 49 term PE cases. In case of other samples, we could not isolate a suitable amount of cell-free DNA to conduct the analysis.

Cases and controls' demographic and pregnancy information are summarized in Table 1, showing that gestational week (CRL) at enrolment were not different between the groups. Cases that developed PE had higher maternal body mass index (BMI) at enrolment, and lower frequency of spontaneous conception (e.g. – more conceived by assisted reproduction technology (ART). In the PE group, there were more women with a history of previous PE and black ethnicity.

**[Place here Table 1. Enrolment and delivery information in pregnancies with and without PE].**

**The PP13 serum level**:

The PP13 serum level was previously reported by Akolekar et al. [8]. Accordingly, the PP13 first trimester level expressed as multiple of the median in un-affected patients was 1.00 [0.76-1.33] [median and 95% CI], compared to 0.93 [0.7-1.3] and 1.11 [0.89-1.49] for preterm PE and term PE, respectively. The detection rate for 10% false positive rate was 41.4 [32.2-51.2] and 37.8 [28.9-47.6] for preterm PE and term PE, respectively.

Combining maternal prior risk factors, biochemical and biophysical markers the prediction of PE was 79.5% and 64.2% for preterm PE and term PE, respectively. Adding PP13 to the prediction of PE increased the prediction accuracy to 85.9% and 71.2%, respectively.

**Polymorphisms:**

As depicted in Table 2, the respective primers revealed the presence of several of the previously reported polymorphic variants reported by Gebhardt et al. [14] and a few new ones, not previously listed in the NCBI SNP database of *LGALS13*. In this cohort, the $DelT_{221}$ variant of Gebhardt et al. [14] was not found. This is most likely related to the relative small study size. For o,8 power, a 1:4 ratio of cases and controls, a sample size of 168 cases of early severe PE versus 672 controls is required to identify the DelT221 mutation. Such large number of early cases is not available to us today.

There were additional SNPs that were reported in SA and not found in the London cohort (Table 2). Those that were detected (new and old ones) did not appear in a significant number of samples to run analysis for elucidating any specific conjunction with PE except -98A/C [28].

**[Place here Table 2. SNP variants identified in the *LGALS13* gene.**]

**-98A/C genotype variation (dbSNP: rs3764843):**

The -98 A/C genotype polymorphism was detected using the pair of the first two primers (Figs. 2 A and B). Table 3 depicts the genotype distribution in PE and unaffected controls. The A/A variant at the -98 position [28] was found to be the dominant genotype in all groups. The C/C genotype has the lowest frequency. In the

comparison between groups, the frequency of the A/C genotype was higher in term PE, less in preterm PE, and the lowest in the unaffected group.

The pattern of genotype distribution was more significantly different when comparing term PE cases to controls (p=0.032) than when all PE cases were compared to controls (p=0.068). There was no difference (p=0.730) in genotype pattern distribution for preterm PE compared to controls (Table 3), presumably due to the very low number of preterm PE cases.

**[Place here Table 3. Correlation between patient outcome groups and genotype.]**

Altogether, our sample had  controls to 1 term PE case . Thus, among the 196 unaffecred controls the "A" allele appears 312 times, corresponding a probability of 0.8. Among the 49 term PE cases the "A" allele appears 80 times, corresponding to probability of 0.92 . Accordingly, the study power for this polymorphism was 0.9 given $\alpha$=0.05. The sample size of preterm PE was underpower for an accurate probability assessment in this study.

**Cohort comparison between London and South Africa (SA):**

- A/A genotype - In both the SA and the London cohorts the A/A genotype had a higher allele frequency overall. However, the A/A majority was more dominant for PE cases compared to control in both London and SA (p<0.001) (Table 4). For the A/A genotype, the odds ratio for term PE was OR=2.91 in London and OR=1.84 for the SA population.
- A/C genotype - The A/C genotype was higher in controls than in PE in both SA and London cohorts.

- C/C genotype - The C/C genotype was carried by too few subjects and thus statistical evaluation was not valuable.

**[Place here Table 4. $\chi^2$ for good of fit to measure the distribution of the genotype in London and SA by outcome].**

In summary, while there were some differences in the genotype distributions between SA and London, the presence of cytosine nucleotide in -98 promoter position (either as a heterozygous (A/C) or homozygous (C/C) genotype) was higher in unaffected controls versus all PE cases and particularly in term PE cases ($p<0.001$).

The comparison of genotype distribution according to racial origin is depicted in Table 5. For the population at large, both black and non-black had a higher frequency of the A/A genotype and a lower frequency of the A/C and C/C genotypes ($p=0.813$). However, there was significant correlation between black ethnicity and PE. Term PE was more frequent in black compared to the non-blacks (31.5% vs. 12.1%, $p<0.001$).

**[Place Table 5. Correlation between patient's race and genotype or patient race and outcome.**]

**Genotype, ethnicity and outcome:**

The C allele (in either C/C or A/C genotype) appears to confer protection from developing term PE for all ethnic groups (Table 6). The A/A genotype was more common in black women with term PE. Accordingly, there was high correlation between term PE outcome, A/A genotype and black ethnicity.

**[Place here Table 6. Correlations between patient's genotype and race according to the outcome group].**

**Hardy-Weinberg analysis:**

Applying the Hardy-Weinberg equation to the genotype distribution as summarized in Table 3 yielded the followings:

- The A/A genotype was detected in 82.1% of cases and 67.3% of controls.
- Heterozygosity (A/C) was detected in 13.4% of cases and 24.5% of controls.
- The C/C genotype was detected in 4.5% of the cases and 8.2% of controls.

Accordingly, the genotype frequencies in the control individuals were in Hardy-Weinberg equilibrium. However, All PE cases deviated significantly from equilibrium ($p=0.036$).

**Odds ratios:**

Calculation of the OR based on a recessive homozygosity model for PE development was conducted by comparing the A/A genotypes to other genotypes. It yielded an OR=1.26 for preterm PE, OR=2.91 for term PE and OR=2.22 for all PE (values not shown). This was consistent with the previous report from SA [28] that the to be protective against PE development. However, on its own the promoter variant genotypes were not strong enough as stand-alone markers for predicting PE.

A Multi-Variate analysis was conducted including: having A/A genotype in -98 position, being of black ethnicity and experiencing previous PE as major confounders, and high body mass index (BMI) and maternal age as weaker contributors. The individual odds ratios (ORs) for developing PE were then recalculated to account for

each risk factor considering the presence of the other risk factors to yield the adjusted OR (Table 7).

**[Place here Table 7. Genotype and other risk factors – adjusted odds ratios.]**

- The accuracy of the observed overall multiple marker assessment was 77.9% for all PE and 83.7% for term PE and 6% and 3% false positive rates, respectively.
- The combined OR for All PE had $R^2$=0,24, $\chi^2_{(5)}$= 47.35, p<0.001, and an Adjusted OR=7.
- Combined OR for term PE has $R^2$=0.27, $\chi^2_{(5)}$=44.99, p<0.001 and the adjusted OR=14
- The Positive Likelihood ratio (LR)=[Sensitivity/(1-Specificity)] was 10.2 and 5.21 for term PE and All PE, respectively.
- The Negative LR-[(1-sensitivity)/ Specificity] was 0.72 and O.73 for term PE and All PE, respectively.
- The Overall LR=Positive LR/Negative LR were 14.17 and 7.13 for term PE and All PE, respectively.

**The effect of -98 polymorphism on *LGALS13* expression:**

In order to study the functional effect of this polymorphism, first we performed a transcription factor binding site search using the Transfac database, which revealed a TFAP2A binding site in the *LGALS13* gene promoter at -101 to -93 positions (Figure 3A). Of interest, the comparison of this binding site nucleotide sequence with the PPM and PWM of the canonical TFAP2A binding site (MA0003.1) in the Jaspar database showed that: 1) 7 out of the 9 bases in the *LGALS13* promoter binding site match the most frequent bases in the canonical binding site; and 2) the „C" in -98 position has a higher occurrence, and thus binding affinity, than "A" in the canonical binding site (Figure 3B). This suggested that having "A" in -98 position may lead to

weaker binding of TFAP2A than having “C” in the same position in the *LGALS13* promoter and a consequent lower *LGALS13* gene expression.

[**Place here Figure 3: <u>Promoter luciferase assays</u>**]

To investigate this hypothesis, next we generated “-98C” and “-98A” *LGALS13* promoter clones and examined their expression in non-differentiated and differentiated BeWo cells. BeWo cells have TFAP2A expression in the non-differentiated state, which increases during trophoblast differentiation. Therefore, we expected to find an increasing difference in luciferase activity with the progress of differentiation between the two clones. Indeed, promoter “-98A" variant had 13% lower expression in non-differentiated BeWo cells ($p=0.04$) while it had 26% lower expression in Forskolin-induced BeWo cells after 48 hours of differentiation ($p<0.001$) compared to the "-98C" variant. The expression of both promoter variants increased during the 48 hours of differentiation. However, the increased expression of "-98C" variant was by 4.55-fold while the increased expression of "-98A" variant was only by 3.85-fold increase (-15%, $p<0.001$).

# Discussion

**Major findings of the study:**

1) *LGALS13* -98 A/C genotype was the lowest in term PE compared to controls ($p<0.032$), similar to a South African cohort; 2) Control but not all PE allele frequencies were in Hardy-Weinberg equilibrium ($p=0.036$); 3) The Odds ratio for term PE calculated from prior risk, the A/A genotype and black ethnicity was 14 ($p<0.001$); 4) In luciferase assays, the *LGALS13* promoter “-98A" variant had 13% ($p=0.04$) and 26% ($p<0.001$) lower expression than the "-98C" variant in non-differentiated and differentiated BeWo cells, respectively. After 48 hours of differentiation, there was 4.55-fold increase in expression of "-98C" variant versus 3.85-fold of "-98A" variant ($p<0.001$).

**The relevance of -98A/C *LGALS13* polymorphism for preeclampsia prediction**

In the present study, we searched for first trimester SNPs of the *LGALS13* gene encoding for PP13 and the potential relevance of these SNPs to PE prediction. An A/C polymorphism previously reported for the -98 promoter position of *LGALS13* gene in a SA cohort collected at term. Our study revealed it for a cohort from London already in the first trimester samples.

The presence of cytosine ("C") in the -98 position (either in its heterozygous A/C genotype variant or the homozygous C/C variant) seems to convey protection from the development of PE, as was previously reported for the cohort in SA [28]. This protection from PE is probably related to the higher expression of *LGALS13* when "C" is in the -98 position compared to "A" in this position in the *LGALS13* promoter, as was detected in BeWo trophoblast-like cells. Our study revealed that while the unaffected control group was in Hardy-Weinberg equilibrium, all PE cases

deviated from it. This was true for both the London and the SA cohorts, and strongly emphasized that having the "C" variant in the -98 position is protecting from developing PE.

PP13 was implicated in several functions of placental development, including immune tolerance to the paternal genes of the migrating trophoblasts [16, 23] as well as blood pressure and utero-placental vasculature control [25]. Therefore, knowledge of polymorphisms in the promoter for this gene might facilitate our understanding of its potential role in regulating the normal progress of pregnancy. The presence of cytosine either in the C/C or A/C format appeared to convey protection from PE, regardless of the differences in the genotype distributions between SA and London.

We have found that genotype variation in the -98 position and black ethnic origin were major contributors in predicting high risk to develop PE. Some differences were found between the London and the SA cohorts in the proportion of A/A, C/C and A/C. Regardless of those differences, in both cohorts having the "C" variant was protecting from PE. The differences in the distributions may be related to tribe origin: Sub-Saharan and Caribbean origin in London compared to Zulu and additional South- and Central African origin in SA [8, 15]. In this context on may address the issue of ancestry information markers (AIMS). To the best of our knowledge there are no AIMS for preeclampsia that can be used in PE predicting according to ethnic origin, although such AIMS were described for certain different geographical regions in the context of other diseases [6-7]. Given the different frequency of PE even among the black ethnic group between South African, Caribbean African and African-American, in the future the polymorphism of PP13 may be developed as a potential future AIM for preeclampsia.

**The relevance of -98A/C polymorphism for *LGALS13* expression**

How the cytosine nucleotide in the A/C or C/C configuration (versus A/A genotype) in -98 position in the promoter region of *LGALS13* may convey resistance to PE? The transcription factor binding site analysis has indicated that having “C” in -98 position in the *LGALS13* promoter compared to having “A” in this position predicts a higher binding of TFAP2A. This transcription factor is critical for the trophoblastic expression of several placental genes including *LGALS13* and other galectin genes in its close vicinity on Chr19 [34]. Thus, it was anticipated that having “C” in the -98 position yields a higher *LGALS13* gene expression and consequently more PP13 protein synthesis.

Indeed, the *in vitro* expression studies confirmed this *in silico* analysis. We performed luciferase assays on the "-98C" and the "-98A" *LGALS13* promoter clones. It was found that the "98A" variant had lower expression in both non-differentiated and differentiated BeWo cells compared to the "-98C" variant. In addition, during differentiation, having the "-98C" variant in the promoter construct was accompanied by a higher fold increase of expression compared to having the "-98A. Since luciferase assays measured only the expression of the promoter constructs transfected into BeWo cells but not the internal expression of *LGALS13* gene copies in BeWo cells, the genotype of BeWo cells for the -98A/C *LGALS13* polymorhisms as well as their aneuploidity [35] did not interfere with our assay.

The expression studies with the -98 "A" and "C" promoter reporter variants emphasized the effect of this polymorphism on *LGALS13* expression in the context of trophoblast differentiation. It is important, since *there is an elevated induction of LGALS13* expression during the differentiation of the trophoblasts to generate the placental villi, in correlation with the increased expression of *TFAP2A* among other

transcription regulatory genes [34]. Thus, this functional observation validates in the *in-silico* prediction of the role of TFAP2A binding to the polymorphism containing binding site. Moreover, Kliman et al. [23] have shown that PP13 is involved in the remodeling of the utero-placental vasculature. Furthermore, both Kliman et al [23] and Then et al. [16] have shown that PP13 induces white blood cell apoptosis and confers maternal immunotolerance to the pregnancies. Thus, the results of this study point to the potential link between placental development, trophoblast differentiation, *LGALS13* expression, maternal blood PP13 concentrations and normal pregnancy maintenance, and prompts additional studies – functional and clinical - to uncover this unexplored area in details.

**The relevance of *LGALS13* polymorphisms for the development of preeclampsia**

We determined *LGALS13* on cell free DNA of maternal blood. In recent years, several tools were developed to investigate how impairments in *LGALS13* nucleotide sequence and its RNA/protein expression could affect the risk to develop PE [14, 22, 24, 28, 36]. In addition to the - 98A/C promoter polymorphism, there is an additional polymorphism of having deletion of the nucleotide thymidine in position 221 ($DelT_{221}$). Identified among black women in SA by Gebhardt et al. [14], the $DelT_{221}$ variant conveyed an 89% positive predictive value for developing early-onset severe PE when present in its heterozygous form. The homozygous form was not associated with viable pregnancy. Sammar et al. [37] constructed the $DelT_{221}$ variant and expressed both the wild type and the $DelT_{221}$ variants of *LGALS13* in *E. coli*. The polypeptides were subsequently purified as recombinant proteins. The $DelT_{221}$ variant is a shorter polypeptide caused by a premature stop codon in the open reading frame compared to the wild type PP13. This wild type PP13 and its truncated PP13 variant

were both found to reduce blood pressure when applied to gravid rats, and this effect lasted for the entire period of PP13 delivery from a slow releasing inter peritoneal pump application [25]. However, unlike the wild-type, the truncated protein failed to cause expansion of the utero-placental vasculature and could not increase placental size and pup weight [25, 37]. These findings emphasize the role of the sugar binding residues of PP13 molecule that are partially missing in the truncated variant as was demonstrated previously [16, 34, 36]. However, as indicated above, the sample size of early severe PE cases and overall in this study was too small to identify the $DelT_{221}$ mutation.

The very rare $DelT_{221}$ mutation and the -98A/C polymorphism demonstrate the importance of PP13 in PE research. In the case of the $DelT_{221}$ mutation, a strong risk to develop early onset and severe PE was discovered [14]. In the case of the -98 A/C promoter variants, the impact is increased risk to develop term and all cases of PE. Accordingly, each variant conveyed a different risk level to develop PE. Further studies in PE and animal models may reveal the physiological and morphological mediators involved in the risk to develop PE derived from each variant and for the two combined.

**The relevance of gene polymorphisms for the development of preeclampsia**

Preeclampsia is a multifactorial disorder. Our study emphasized the increase in our tool box for identifying patients at risk to develop PE, which may be derived from impaired sequence and/or expression of *LGALS13* gene and the encoded PP13 protein. Increasing this tool box may improve our ability to develop *novel drug-target composite* through identifying the sub-group of patients whose risk to develop PE is associated with impaired PP13 sequence / expression / function. PE research signals

the need to link marker discovery, risk stratification, and identify subgroup candidates for evaluating potential new methods of therapy. This approach may open the road to facilitate the delivery of new monitoring and management [38].

Polymorphisms were already identified in various genes of proteins that are involved in different functions related to the PE disorder [9-14]. Among them are the polymorphisms in genes that are involved in blood pressure regulation through the renin-angiotensin-aldosterone system. Polymorphisms in this system are associated with hypertensive disorders in pregnancy as was reported for the white population of Poland [39], and the black Afro-Caribbean, but not the white Europeans in South and Central London [9]. However, this was not the case among Asian women of Taiwan [9]. More than one type of polymorphism in this system, including the variant related to hypertensive disorders in pregnancy, was identified among Tunisian Arabs [40]. At the same time, these two variants of polymorphisms have been shown to be tightly related to other cardiovascular disorders. The latter includes congestive cardiomyopathy, supraventricular arrhythmias, reduced exercise capacity and an increase in left ventricular systolic dimension. They all underline the higher frequency of cardiovascular morbidity among women who experience hypertensive disorders in pregnancy [41-42].

Polymorphisms were also identified in the endothelial nitric oxide synthase (eNOS) system that regulates vasodilation, which is one of the hallmarks of early stages of pregnancy development. Linkage studies in affected sibling pairs have implicated the *NOS3* gene (encoding eNOS) locus on chromosome 7q35 [10, 43]. However, a recent meta-analysis failed to show a repeated correlation between this polymorphism and PE, either because the studies were not powered or because this candidate gene might be different according to ethnic background [9-10].

Analysis of the polymorphisms of *IL27* in PE among Han Chinese women revealed a significantly reduced risk of PE in one genotype variant compared to the others in a dominant allele model. This was particularly true for the severe PE subgroup, implying that SNPs in *IL27* may have an effect on individual susceptibility to PE [12].

These findings may lead to the development of new tools for PE prediction based on SNPs. A new diagnostic approach can be adapted based on the use of a battery of several SNPs. Tools like quantitative real time multiplex PCR or DNA chips could be developed to identify *pre-dispositions to develop PE*.

PE has a significant inherited component and it is likely that many genes are involved [44]. Considering that the etiology of PE remains complex, identification of candidate genes for PE pre-disposition could be a substantial aid in the understanding of this important public health problem and provide clues for prevention and treatment. It might also have wide relevance since women with a history of PE are at increased risk of cardiovascular disease in later life [45].

**Study Limitation**

While our study was powered enough to predict term PE according to -98A/C polymorphism, the study was underpowered to evaluate the value of this polymorphism for predicted preterm PE. This is only the second study that explore the polymorphism of PP13 following a previous one conducted in South Africa [28]. While both studies implies the importance of PP13 promotor -98 A/C polymorphism, more studies are required with additional cohorts of other ethnic origins to verify the use of this polymorphism in predicting PE. For identifying carriers of the $delT_{221}$ much larger sample size is required in view of the rare mutation frequency.

**Is the *LGALS13* DNA detected here in this study of maternal or fetal origin?**

Studies have shown that approximately 5-10% of the cell-free DNA in first trimester maternal blood is of fetal/placental origin while 90-95% is of maternal origin [46]. Therefore, 90-95% of cell-free genomic DNA fragments carrying the *LGALS13* gene are supposed to be originated from maternal tissues. Since *LGALS13* is 5,000 base-pairs long compared to the 3 billion base-pair length of the human genome, the isolated cell-free genomic DNA contained approximately $1.7x10^{-4}$ % of maternal and fetal *LGALS13* sequences combined and $1.7x10^{-5}$ % of fetal *LGALS13* sequences. Accordingly, we estimate that the fetal fraction of *LGALS13* in this study was too low for detection.

# Conclusion

This is the first study conducted to identify *LGALS13* DNA sequence variants in maternal blood in the first trimester. Our findings suggest that the -98A/C gene polymorphisms in the promoter region of the *LGALS13* gene may be associated with PE development with strong correlation to black ethnicity. The -98 A/C genotype appears to provide protection from PE development. More studies are required to verify the use of the -98A/C promoter variant for aiding PE prediction alone or in combination with genetic variants of other genes with possible diagnostic significance [47-49]. Further studies are required in PE models to examine the impact of the polymorphism on the disease development and in finding novel therapies.

# Competing Interest

LMD, RC and HM are employees of Hy Laboratories, which owns the rights for the PP13 technology. All other authors declare no conflict of interest.

## Funding Sources

This research was supported, in part, by ASPRE FP7 # 601852 grant (to LML, RC, KHN, and HM), by the Perinatology Research Branch, Division of Intramural Research, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), National Institutes of Health (NIH), Department of Health and Human Services (DHHS), by Federal funds from the NICHD under Contract No. HSN275201300006C, and by the Hungarian Academy of Sciences Momentum Grant "LP2014-7/2014" (to NGT).

## Acknowledgements

We thank Zsolt Gelencser (Hungarian Academy of Sciences) for technical assistance.

# REFERENCES


1. Walker JJ. Pre-eclampsia. Lancet 2000; 356(9237):1260-1265.
2. Roberts JM, Cooper DW. Pathogenesis and genetics of pre-eclampsia. Lancet 2001; 357(9249):53-56.
3. Khan KS, Wojdyla D, Say L, Metin-Gülmezoglu AM, Van Look PA**.** WHO analysis of causes of maternal death: a systematic review. Lancet 2006; 367(9516):1066–1074.
4. Duley L. The global impact of pre-eclampsia and eclampsia. Semin Perinatol 2009; 33:130-137.
5. Redman CW, Sacks GP, Sargent IL. Preeclampsia: an excessive maternal inflammatory response to pregnancy. Am J Obstet Gynecol 1999; 180(2 Pt 1):499-506.
6. Nakimuli A., Chazara O, Byamugisha J, Elliott AM, Kaleebu P, Mirembe F, et al. Pregnancy, parturition and preeclampsia in women of African ancestry. American J Obstet Gynecol 2013; 210:510–20.e1.
7. Moodley J, Chair of NCCEMD. Saving Mothers 2011-2013: Sixth report on the Confidential Enquiries into Maternal Deaths in South Africa (NCCEMD) Department of Health, South Africa. Pattinson R (Editor) 1-93. http://www.kznhealth.gov.za/mcwh/Maternal/Saving-Mothers-2011-2013-short-report.pdf ; 2015
8. Akolekar R, Syngelaki A, Sarquis R, Zvanca M, Nicolaides KH. Prediction of early, intermediate and late pre-eclampsia from maternal factors, biophysical and biochemical markers at 11-13 weeks. Prenat Diagn 2011; 31:66-74.

9. Yang J, Shang J, Zhang S, Li H, Huiong L. The role of the rennin-angiotensin-aldosterone system in preeclampsia: genetic polymorphism and micro RNA. J Molec Endoc 2013; 50:R53-R66.

10. Yu CK, Casas JP, Savvidou MD, Sahemey MK, Nicolaides KH, Hingorani AD. Endothelial nitric oxide synthase gene polymorphism (Glu298Asp) and development of pre-eclampsia: a case-control study and a meta-analysis. BMC Pregnancy Childbirth 2006; 16:1-9.

11. Papazoglou D, Galazios G, Koukourakis MI, Panagopoulos I, Kontomanolis EN, Papatheodorou K, et al. Vascular endothelial growth factor gene polymorphism and pre-eclampsia. Mol Hum Reprod 2004; 10:321-324.

12. Chen P, Gong Y, Pu Y, Wang Y, Zhou B, Song Y, et al. (2016) Association between polymorphisms in IL-27 gene and pre-eclampsia. Placenta 37:61-64.

13. Arias F, Rodriquez L, Rayne SC, Kraus FT. Maternal placental vasculopathy and infection: two distinct subgroups among patients with preterm labor and preterm ruptured membranes. Am J Obstet Gynecol 1993; 168; 585-591.

14. Gebhardt S, Bruiners N, Hillerman R. A novel exonic variant (221delT) in the LGALS13 gene encoding placental protein 13 (PP13) is associated with preterm labour in a low risk population. J Reprod Immunol 2009; 82:166-173.

15. Huppertz B. Placental origins of preeclampsia: challenging the current hypothesis. Hypertension 2008; 51:970-975.

16. Than NG, Romero R, Goodman M, Weckle A, Xing J, Dong Z, et al. A primate subfamily of galectins expressed at the maternal-fetal interface that promote immune cell death. Proc Natl Acad Sci USA 2009; 16:9731-9736.

17. Romero R, Nien JK, Espinoza J, Todem D, Fu W, Chung H, et al. A longitudinal study of angiogenic (placental growth factor) and anti-angiogenic

(soluble endoglin and soluble vascular endothelial growth factor receptor-1) factors in normal pregnancy and patients destined to develop preeclampsia and deliver a small for gestational age neonate. J Matern Fetal Neonatal Med 2008; 21:9-23.

18. Thilaganathan B. Placental syndromes: getting to the heart of the matter. Ultrasound Obstet Gynecol. 2017; 49(1):7-9.
19. Osol G, Bernstein I. Preeclampsia and maternal cardiovascular disease: consequence or predisposition? J Vasc Res 2014; 51:290-304.
20. Than NG, Sumegi B, Than GN, Berente Z, Bohn H. Isolation and, sequence analysis of a cDNA encoding human placental tissue protein 13 (PP-13), a new lysophospholipase, homologue of human eosinophil Charcot-Leyden Crystal protein. Placenta 1999; 20:703-710.
21. Balogh A, Pozsgay J, Matkó J, Dong Z, Kim CJ, Várkonyi T, et al. Placental protein 13 (PP13/Galectin-13) undergoes lipid raft-associated subcellular redistribution in the syncytiotrophoblast in preterm preeclampsia and HELLP syndrome. Am J Obstet Gynecol 2011; 205:156.e1-14.
22. Huppertz B, Meiri H, Gizurarson S, Osol G, Sammar M. Placental protein 13 (PP13): a new biological target shifting individualized risk assessment to personalized drug design combating pre-eclampsia. Hum Reprod Update 2013; 19:391-405.
23. Kliman HJ, Sammar M, Grimpel Y-I, Lynch SK, Milano KM, Pick E, et al. Placental Protein 13 and decidual zones of necrosis: an immunologic diversion that may be linked to preeclampsia. Reprod Sci 2012; 19:16-30.
24. Than NG, Romero R, Balogh A, Karpati E, Mastrolia SA, Staretz-Chacham O, et al. Galectins: double edged swords in the cross-roads of pregnancy

complications and female reproductive tract inflammation and neoplasia. J Pathol Transla Med 2015; 49:181-208.

25. Gizurarson S, Sigurdardottir ER, Meiri H, Huppertz B, Sammar M, Sharabi-Nov A, et al. Placental Protein 13 administration to pregnant rats lowers blood pressure and augments fetal growth and venous remodeling. Fetal Diagn Ther 2016; 39:56-63.

26. Sammar M, Nisemblat S, Fleischfarb Z, Golan A, Sadan O, Meiri H, et al. Placenta-bound and body fluid PP13 and its mRNA in normal pregnancy compared to preeclampsia, HELLP and preterm delivery. Placenta 2011; 32 Suppl:S30-S36.

27. Farina A, Zucchini C, Sekizawa A, Purwosunu Y, de Sanctis P, Santarsiero G, et al. Performance of messenger RNAs circulating in maternal blood in the prediction of preeclampsia at 10-14 weeks. Am J Obstet Gynecol 2010; 203:e1-e6.

28. Bruiners N, Bosman M, Postma A, Gebhardt S, Rebello G, Sammar M, et al. Promoter variant -98A-C of the LGALS13 gene and pre-eclampsia. 8th World Congress of Prenatal Medicine and Fetal Development. Florence, September; 2008.

29. Hadlock FP, Shah YP, Kanon DJ, Lindsey JV. Fetal crown-rump length: re-evaluation of relation to menstrual age (5–18 weeks] with high-resolution real-time US. Radiology 1992; 182:501-505.

30. Poon LC, Stratieva V, Piras S, Piri S, Nicolaides KH. Hypertensive disorders in pregnancy: combined screening by uterine artery Doppler, blood pressure and serum PAPP-A at 11-13 weeks. Prenat Diagn 2010; 30:216-223.

31. Martin AM, Bindra R, Curcio P, Cicero S, Nicolaides KH. Screening for pre-eclampsia and fetal growth restriction by uterine artery Doppler at 11–14 weeks of gestation. Ultrasound Obstet Gynecol 2001; 18:583-586.

32. Poon LCY, Staboulidou I, Maiz N, Plasencia W, Nicolaides KH. Hypertensive disorders in pregnancy: screening by uterine artery Doppler at 11–13 weeks. Ultrasound Obstet Gynecol 2009; 34:142-148.

33. Lindheimer MD, Taler SJ, Cunningham FG. American Society of Hypertension position paper: hypertension in pregnancy. Clin Hypertens (Greenwich) 2009; 11:214-225.

34. Than NG, Romero R, Xu Y, Erez O, Xu Z, Bhatti G, et al. (2014) Evolutionary origins of the placental expression of Chromosome 19 cluster galectins and their complex dysregulation in preeclampsia. Placenta 35:855-865.

35. Long CA, Bauer GS, Lowe ME, Strauss AW, Gast MJ. Isolation and characterization of the gene from a human genome encoding 17β-estradiol dehydrogenase: A comparison of Jar and BeWo choriocarcinoma cell lines. Am J Obstet Gynecol 2010; 163 (6):1976–1981

36. Than NG, Balogh A, Romero R, Karpati E, Erez O, Szilagyi A, et al. Placental Protein 13 (PP13) – a placental immunoregulatory galectin protecting pregnancy. Front Immunol 2014; 5(348):1-25

37. Sammar M, Nisamblatt S, Gonen R, Huppertz B, Gizurarson S, Osol G,et al. The role of the carbohydrate recognition domain of placental protein 13 (PP13) in pregnancy evaluated with recombinant PP13 and the DelT221 PP13 variant. PLoS One 2014; 9(7):e102832.

38. Hahn S. Preeclampsia - will orphan drug status facilitate innovative biological therapies? Front Surg 2015; 2:1-4.

39. Bogacz A, Bartkowiak-Wieczorek J, Procyk D, Seremak-Mrozikiewicz A, Majchrzycki M, Dziekan K, et al. Analysis of the gene polymorphism of aldosterone synthase (CYP11B2) and atrial natriuretic peptide (ANP) in women with preeclampsia. Eur J Obstet Gynecol Reprod Biol 2016; 197:11-15.

40. Saidi S, Mahjoub T, Almawi WY. Aldosterone synthase gene (CYP11B2) promoter polymorphism as a risk factor for ischaemic stroke on Tunisian Arabs. J Renin Angiotensin Aldosteron Syst 2010; 11:180-186.

41. Lo HM, Lin FY, Lin JL, Tseng CD, Hsu KL, Chiang FT, et al. Electrophysiological properties in patients undergoing atrial compartment operation for chronic atrial fibrillation with mitral valve disease. Eur Heart J 1997; 18:1805-1815.

42. Ardizzone N, Cappello F, Di Felice V, Rappa F, Minervini F, Marasà S, et al. Atrial natriuretic peptide and CD34 overexpression in human idiopathic dilatedcardiomyopathies. APMIS 2007; 115: 1227-1233.

43. Lade JA, Moses EK, Guo G, Wilton AN, Grehan M, Cooper DW, et al. The eNOS gene: a candidate for the preeclampsia susceptibility locus? Hypertens Pregnancy 1999; 18: 81-93.

44. Pridjian G, Puschett JB. Preeclampsia: Part 2: experimental and genetic considerations. Obstet Gynecol Surv 2002; 57: 619-40.

45. Smith GC, Pell JP, Walsh D. Pregnancy complications and maternal risk of ischaemic heart disease: a retrospective cohort study of 129,290 births. Lancet 2001; 357: 2002-2006.

46. Lo YMD, Chan KCA, Dun H, Chen EZ, Jiang P, Lun FMF, Zheng YW, Leung TY, Lau TK, Cantor CR, Chiu RWK. Maternal Plasma DNA Sequencing

Reveals the Genome-Wide Genetic and Mutational Profile of the Fetus. *Science Translational Medicine* 2010;. 2(61):61ra91

47. Rigourd V, Chelbi S, Chauvet C, Rebourcet R, Barbaux S, Bessières B, et al. Re-evaluation of the role of STOX1 transcription factor in placental development and preeclampsia. J Reprod Immunol 2009; 82: 174-181.

48. Fong FM, Sahemey MK, Hamedi G, Eyitayo R, Yates D, Kuan V, et al. Maternal genotype and severe preeclampsia: a HuGE review. Am J Epidemiol 2014; 180: 335-345.

49. Haram K, Mortensen JH, Nagy B**.** Genetic aspects of preeclampsia and the HELLP Syndrome. J Pregnancy 2015; 2014: 1-13.

# Tables

**Table 1. Enrolment and delivery information in pregnancies with and without PE.**

| Parameter | Controls (n=196) | Term PE (n=49) | Preterm PE (18) | p |
|---|---|---|---|---|
| **Enrolment** | | | | |
| **BMI (kg/h$^2$) (Mean [95% CI])** | 25.2[b] [24.5-25.9] | 29.5[a] [27.5-31.5] | 29.9[a] [26.7-33.1] | <0.001 |
| **Maternal age (yrs) (Mean [95% CI])** | 31.9 [31.0-32.7] | 31.0 [29.4-32.5] | 33.2 [29.6-36.7] | 0.178 |
| **Smoking (%)** | 7.7 | 6.1 | 0.0 | 0.479 |
| **Conception Spontaneous (%)** | 95.9 | 95.9 | 88.9 | 0.415 |
| **Previous PE or GH (%)** | 4.6[b] | 20.4[a] | 16.7 [a] | 0.002 |
| **Ethnicity Black (%)** | 27.6[b] | 57.1[a] | 38.9[a] | <0.001 |
| **Delivery** | | | | |
| **Gestation age (days) (mean [95% CI])** | 39.9[a] [39.7-40.1] | 39.3[a] [38.9-39.6] | 35.8[b] [35.4-36.2] | <0.001 |
| **Baby weight (gr) (Mean[95% CI])** | 3399[a] [3352-3447] | 3145[b] [3010-3281] | 2244[c] [2054-2435] | <0.001 |

PE - pre-eclampsia, GH - gestational hypertension, BMI - body mass index

The letter indicate significantly higher (a), lower (b) or the lowest (c)

**Table 2. SNP variants identified in the *LGALS13* gene.**

| Variant name | Code | Region | a.a substitution | No. of individuals |
|---|---|---|---|---|
| **-98 A/C** | rs 3764843 | Promoter | - | 244 |
| **15 C to T** | Novel | Exon 1 | P -> P | 1 |
| **Ex1+104 C/T** | rs 3764844 | Intron 1 | - | 6 |
| **Ex1+116 C/T** | Novel | Intron 1 | - | 13 |
| **Ex3-15 G/A** | rs 116537355 | Intron 2 | - | 2 |
| **Ex3-22 A/G** | rs 2233706 | Intron 2 | - | 14 |
| **Ex3-36 A/G** | rs 201378711 | Intron 2 | - | 0 |
| **130 A to G** | rs 774287801 | Exon3 | M -> V | 0 |
| **221delT** | rs 371049824 | Exon 3 | - | 0 |
| **229-237HotSpot** | - | Exon 3 | T->N, T->L, D->H | 0 |
| **260 A to G** | rs 74400185 | Exon3 | K -> R | 0 |
| **330 C to T** | Novel | Exon 4 | Y -> Y | 1 |

The list of variant names includes all those already listed by Gebhardt et al. [14] for the SA cohort, even if not found here and the three novel variants discovered for the London cohort and not yet included in the NCBI library.

1. **Novel Exonic variant 15 C to T** (30 nucleotides before and after)

CAAGAAGGAGAGAACAATGTCTTCTTTACC[C/T]

GTGAGTTGAAAAGGCACAGCCTTCAAAAAT

2. **Novel Intronic variant EX1+116** (30 nucleotides before and after)

GCATTTTTGCTGTGAATGCTTTACTTAGAG[C/T]

TATTGAGGTGTGGAATAGAAACCCTGAGGC

3. **Novel Exonic 330 C to T** (30 nucleotides before and after)

GTAGATAAAGGTCAATGGCATACGCATTTA[C/T]

GGCTTTGTCCATCGAATCCCGCCATCATTT

**Table 3. Correlation between patient outcome groups and genotype.**

| | Unaffected Control (n=196) | | All PE (n=67) | | Preterm PE (PE 34+0 - 36+6) (n=18) | | Term PE (PE>37 weeks (n=49) | |
|---|---|---|---|---|---|---|---|---|
| **Genotype** | **N** | **%** | **N** | **%** | **N** | **%** | **N** | **%** |
| **A/A** | 132 | 67.3 | 55 | 82.1 | 13 | 72.2 | 42 | 85.5 |
| **A/C** | 48 | 24.5 | 9 | 13.4 | 3 | 16.7 | 6 | 12.2 |
| **C/C** | 16 | 8.2 | 3 | 4.5 | 2 | 11.1 | 1 | 2 |
| **p** | | | p=0.068 | | p=0.730 | | p=0.032 | |

*LGALS13* gene was sequenced from maternal plasma cfDNA of PE and non-PE using the primers flanking exon 1 and the flanking introns.

**Table 4. $\chi^2$ for good of fit to measure the distribution of the genotype in London and SA by outcome.**

| Genotype | Control London N=196 | | Control SA N=282 | | Term PE London N=49 | | Term PE SA N=34 | |
|---|---|---|---|---|---|---|---|---|
| | **N** | **%** | **N** | **%** | **N** | **%** | **N** | **%** |
| A/A | 132 | 67.3 | 150 | 53.0 | 42 | 85.8 | 23 | 67.7 |
| A/C | 48 | 24.5 | 114 | 40.5 | 6 | 12.2 | 6 | 17.6 |
| C/C | 16 | 8.2 | 18 | 6.5 | 1 | 2.0 | 5 | 14.7 |
| p | p<0.001 | | | | p=0.014 | | | |

The results of SA were extracted from Bruiners et al. [28].

The statistical significance compared the patients according to the outcome group

**Table 5. Correlation between patient's race and genotype or patient race and outcome.**

| | Black (n=89) | | Non Black (n=140) | | |
|---|---|---|---|---|---|
| **Genotype** | **N** | **%** | **N** | **%** | |
| A/A | 63 | 70.8 | 101 | 72.2 | |
| A/C | 19 | 21.3 | 31 | 22.1 | p=0.813 |
| C/C | 7 | 7.9 | 8 | 5.7 | |
| **Group** | | | | | |
| Control | 54 | 60.7 | 114 | 81.5 | |
| Preterm PE | 7 | 7.9 | 9 | 6.4 | p<0.001 |
| Term PE | 28 | 31.5 | 17 | 12.1 | |

Top: Genotype distribution according to ethnic origin. Comparison was made between black and non-black according to having the detected all three genotypes and not for comparing each genotype on its own, and indicated no significance difference in the genotype distribution.

Bottom: Outcome distribution according to ethnic origin. Comparison was made between black and non-black according to the distribution of all three outcome groups (and not for comparing each outcome group on its own), and indicated significant higher frequency of PE in the black ethnic group.

**Table 6. Correlations between patient's genotype and race according to the outcome group.**

| | Black (n=54) | | Non Black (n=114) | | |
|---|---|---|---|---|---|
| **Control** | **N** | **%** | **N** | **%** | |
| A/A | 35 | 64.8 | 79 | 69.3 | |
| A/C | 13 | 24.1 | 28 | 24.6 | p=0.527 |
| C/C | 6 | 11.1 | 7 | 6.1 | |
| **Preterm PE** | **Black (n=7)** | | **Non Black (n=9)** | | |
| **Genotype** | **N** | **%** | **N** | **%** | |
| A/A | 6 | 85.7 | 6 | 66.7 | |
| A/C | 0 | 0 | 3 | 33.3 | p=0.213 |
| C/C | 1 | 14.3 | 0 | 0 | |
| **Term PE** | **Black (n=28)** | | **Non Black (n=17)** | | |
| **Genotype** | **N** | **%** | **N** | **%** | |
| A/A | 22 | 78.6 | 16 | 94.1 | |
| A/C | 6 | 21.4 | 0 | 0 | p=0.043 |
| C/C | 0 | 0 | 1 | 5.9 | |

Top: Genotype distribution in the control outcome according to ethnicity showed no significant difference of the three genotypes according to ethnicity.

Middle: Genotype distribution in the preterm PE according to ethnicity showed no significant difference of the three genotypes according to ethnicity

Bottom: Genotype distribution in the term PE outcome according to ethnicity showed significant difference of the three genotypes according to ethnicity

**Table 7. Genotype and other risk factors – adjusted odds ratios.**

| | All PE | | Term PE | |
|---|---|---|---|---|
| **Parameter** | **OR [95% CI]** | **p** | **OR [(95% CI]** | **p** |
| A/A genotype | 2.57 [1.18-5.61] | 0.018 | 3.68 [1.34-9.72] | 0.009 |
| Black Ethnicity | 1.97 [1.03-3.77] | 0.039 | 2.40 [1.16-4.97] | 0.018 |
| BMI ($kg/h^2$) | 1.12 [1.06-1.19] | <0.001 | 1.12 [1.05-1.19] | <0.001 |
| Previous PE | 3.79 [1.44-10.00] | 0.007 | 4.67 [1.63-13.37] | 0.004 |
| Maternal age (years) | 0.98 [0.93-1.03] | 0.410 | 0.96 [0.90-1.02] | 0.194 |
| Overall accuracy (%) | 77.9 | | 83.7 | |
| Overall OR | 7.0 | | 14.0 | |
| Detection Rate (%) | 31.3 | | 30.6 | |
| False Positive Rate (%) | 6 | | 3 | |
| Positive LR | 5.21 | | 10.2 | |
| Negative LR | 0.73 | | 0.72 | |
| Overall LR | 7.13 | | 14.17 | |

Top: Assessments of the adjusted Odds ratios (OR) based on the calculation of the individual confounder OR given the other confounders. The followings confounders showed significant differences: A/A genotype, previous PE, black ethnicity, BMI and maternal age.

Multiple regression was conducted for all these adjusted values combined. Detection rate and false positive rates were calculated from the observed measures. Positive, Negative and overall LR were calculated as described in statistical methods.

# Figure legends

**Figure 1 - Primer details and PCR products.**

A- The primers used for each exon PCR (Ex1-Ex4) are detailed for their forward (F) and reverse (1R) order along with the respective melting temperature (Tm), annealing temperature (Ta) and the amplified DNA fragment size.

B - PCR amplification of each *LGALS13* exons (Ex1-Ex4) presented after separation on Gel electrophoresis. M - 100bp DNA marker

**Figure 2 – Sequencing of the -98 Site.**

A – A multiple sequence alignment of the PCR products of exon 1 from several specimens, including *LGALS13* promoter region and flanking intron.

B – Enlarged vision of 20 nucleotides including the -98 position (peaks after the black vertical line). On top – the subject marked in section A that carries the A/A genotype. The two rows below show additional subjects with C/C genotype (middle) and A/C genotype (bottom).

**Figure 3 – Promoter luciferase assays.**

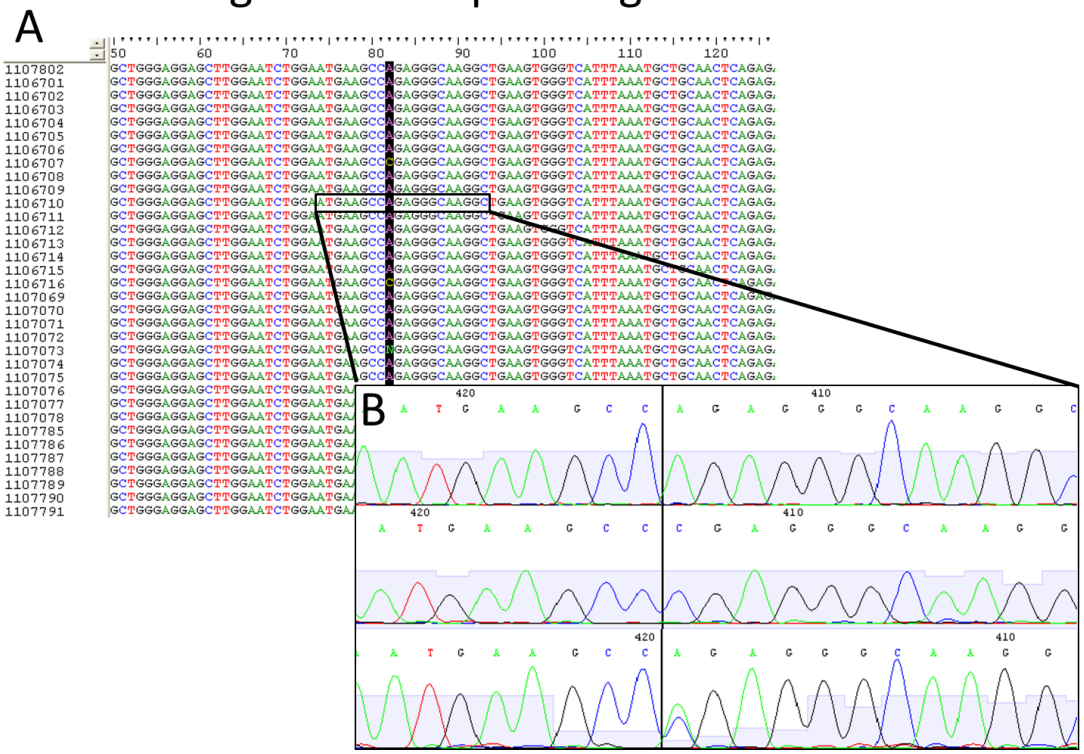
A – Partial promoter sequence of *LGALS13*. Nucleotide positions from the translation initiator codon (underlined) are shown in the left. The first exon is shown with bold black, the predicted TFAP2A binding site is depicted by bold blue underlined letters, and the “C” in -98 position is highlighted with red.

B – Positional Probability Matrix (PPM) and Positional Weight Matrix (PWM) of the canonical TFAP2A binding site (MA0003.1) as downloaded from the Jaspar database (http://jaspar.genereg.net/). PPM values provide the normalized occurrence of each nucleotide at each position in the canonical TFAP2A binding site, while PWM – log-likelihood ratio derived from PPM – values provide binding affinity scores for each nucleotide at each position. The most frequent nucleotides in each position in the PPM are depicted with values colored according to the coloring of nucleotides in the PWM. As shown with red square, the „C” in the fourth position of the TFAP2A binding site has a higher occurrence, and thus affinity, than “A”. C – Luciferase activity of “-98C” and “-98A” *LGALS13* promoter clones. Promoter “-98A" variant had 13% lower expression in non-differentiated BeWo cells (p=0.04) and 26% lower

expression in Forskolin-induced BeWo cells after 48 hours of differentiation ($p<0.001$) than the “-98C" variant. While the expression of both promoter variants increased during the 48 hours of differentiation, the "-98C" variant had 4.55-fold increase while the "-98A" variant had only 3.85-fold increase in expression (-15%, $p<0.001$). All experiments were run in triplicate.

949

# Figure 2 – Sequencing of -98 site




950
951

**Figure 3: Promoter luciferase assays**