Endoglin pathway genetic variation in preeclampsia: A validation study in Norwegian and Latina cohorts

Mandy J. Schmella, James M. Roberts, Yvette P. Conley, Dianxu Ren, Gro L. Storvold, Sue A. Ingles, Melissa L. Wilson, Anne Catherine Staff, Carl A. Hubel

A R T I C L E   I N F O

Keywords:
Preeclampsia
Genetic association
Candidate gene
Endoglin
Pregnancy

O B J E C T I V E: The purpose of this study was to validate our previous genetic association findings related to the endoglin (ENG) pathway from an American Caucasian preeclampsia cohort in independent preeclampsia cohorts. We also sought to explore the ENG pathway for new genetic associations in these independent cohorts.

S T U D Y   D E S I G N: We used a tagging single nucleotide (tSNP) approach to assess genetic variability across five ENG pathway genes (ENG, TGFβ1, TGFβR1, ALK1, and TGFβR2) in a Caucasian cohort from Norway (n = 77 preeclampsia cases & n = 63 normotensive controls) and a White Hispanic cohort from Southern California (n = 69 preeclampsia cases & n = 106 normotensive controls).

M A I N   O U T C O M E   M E A S U R E S: Univariate analyses (Chi Square, Fisher’s Exact) and multivariate logistic regression were conducted to evaluate the association between tSNP genotype distributions and pregnancy outcome in each cohort. Logistic regression models were adjusted for maternal age at delivery, infant sex, parity, smoking during pregnancy, and prepregnancy BMI.

R E S U L T S: Although we were unable to replicate our previous SNP-specific findings (ENG rs11792480, TGFβR2 rs6550005; p’s > 0.05), we found that genetic variation in TGFβR1(ALK5) (rs6478974) and TGFβR2 (rs11129420, rs6802220, rs11551078, rs3773640, rs3773663) was significantly associated with preeclampsia in the Norwegian cohort and genetic variation in ALK1 (rs706819) and TGFβR2 (rs9843942) was significantly associated with preeclampsia in the Latina cohort.

C O N C L U S I O N: Overall, our results provide further support for the involvement and investigation of the endoglin pathway in preeclampsia.

Characterized as one of the “great obstetrical syndromes,” [1] preeclampsia is a heterogeneous, multi-system hypertensive disorder that affects 3–5% of pregnancies [2–4] and significantly contributes to maternal and infant morbidity and mortality worldwide [5]. Preeclampsia is diagnosed when a woman develops new-onset gestational hypertension and proteinuria, or multi-systemic signs in the absence of proteinuria, after 20 weeks’ gestation. However, the underlying pathophysiologic mechanisms that lead to the development of preeclampsia and subsequent maternal endothelial dysfunction have not been fully elucidated [5]. An imbalance of angiogenesis-related factors,
invasive phenotype [11,12]. Throughout pregnancy, ENG mRNA expression is elevated in the placenta and the maternal blood (cellular and cell-free components) in women who develop preeclampsia [13]. Preeclampsia was defined as SBP ≥ 160 or DBP ≥ 110 on ≥ two occasions that were at least six hours apart (in accordance with the criteria outlined by the American College of Obstetricians and Gynecologists in 2002) [25]. HELLP syndrome was defined by the presence of hemolysis, elevated liver enzymes, and low platelets in the participants with preeclampsia. Partial HELLP was diagnosed when only two of the three signs were present. Normotensive (control) pregnancies were defined as those pregnancies in which a woman’s blood pressure did not reach 140/90 mmHg and she did not develop proteinuria. For this validation study, all control participants delivered appropriate for gestational age infants at term. Based on self-reported country of birth and names, all Norwegian participants included in this current study were categorized as White. The Oslo Pregnancy Biobank was approved by the Regional Committee of Medical Research Ethics in Eastern Norway, and use of DNA samples for the current study was approved by the University of Pittsburgh Institutional Review Board (IRB).

Participants of the “Pilot Study of Novel Candidate Genes for Preeclampsia” case-control research study (University of Southern California) were retrospectively recruited from delivery logs at the Los Angeles County and University of Southern California Women’s and Children’s Hospital or during their postpartum stay at the Women’s and Children’s Hospital between the years of 1999 and 2008. Women with lupus, chronic renal disease, multiple gestation, or sickle cell disease/trait were excluded from participation. Preeclampsia was also defined in alignment with the blood pressure and proteinuria criteria outlined by the American College of Obstetricians and Gynecologists in 2002 (please see criteria above) [25]. Severe preeclampsia was defined as SBP ≥ 160 mmHg or DBP ≥ 110 mmHg on ≥ two occasions at least 6 hours apart and proteinuria ≥ 500 mg/dL in a 24-hr specimen or ≥ 3 on a urine dipstick. HELLP syndrome was defined by the presence of hemolysis (abnormal peripheral smear or LDH ≥ 600), elevated liver enzymes (ALT and/or AST ≥ 70), and low platelets (≤ 100,000). Partial HELLP was diagnosed when only two of the three signs were present. For this replication study, all participants with HELLP or partial HELLP also met the preeclampsia criteria. Uncomplicated pregnancies were defined as those pregnancies without significant hypertension. Based on self-report, all 175 Latina participants included in the current study were categorized as Hispanic Whites. The “Pilot Study of Novel Candidate Genes for Preeclampsia” research study was approved by the University of Southern California IRB, and use of the DNA samples in the current study was approved by the University of Pittsburgh IRB.

### 2. Materials and methods

#### 2.1. Study populations

De-identified clinical data and DNA samples were obtained from the Oslo Pregnancy Biobank and the “Pilot Study of Novel Candidate Genes for Preeclampsia” research study conducted at the University of Southern California, for Norwegian and Latina cohort participants, respectively. Participants of the Oslo Pregnancy Biobank were women with singleton pregnancies who delivered via elective cesarean section between the years 2001 and 2008. Women with a history of chronic hypertension, renal disease, diabetes, or other chronic diseases were excluded. Preeclampsia was defined as new-onset hypertension and proteinuria after gestational week 20, in accordance with the American College of Obstetricians and Gynecologists 2002 criteria that were in place when the study was conducted [25]. New-onset hypertension was defined as systolic blood pressure (SBP) ≥ 140 mmHg or diastolic blood pressure (DBP) ≥ 90 mmHg after 20 weeks’ gestation, measured on ≥ two occasions at least six hours apart in a previously normotensive woman. Proteinuria was defined as ≥ 0.3 g in a 24-hr specimen or a urine dipstick ≥ 1+ on ≥ two midstream urine samples collected 6-h apart in the absence of a urinary tract infection. Severe preeclampsia was defined as SBP ≥ 160 or DBP ≥ 110 on ≥ two occasions that were at least six hours apart (in accordance with the criteria outlined by the American College of Obstetricians and Gynecologists in 2002) [25]. HELLP syndrome was defined by the presence of hemolysis, elevated liver enzymes, and low platelets in the participants with preeclampsia. Partial HELLP was diagnosed when only two of the three signs were present. Normotensive (control) pregnancies were defined as those pregnancies in which a woman’s blood pressure did not reach 140/90 mmHg and she did not develop proteinuria. For this validation study, all control participants delivered appropriate for gestational age infants at term. Based on self-reported country of birth and names, all Norwegian participants included in this current study were categorized as White. The Oslo Pregnancy Biobank was approved by the Regional Committee of Medical Research Ethics in Eastern Norway, and use of DNA samples for the current study was approved by the University of Pittsburgh Institutional Review Board (IRB).

Participants of the “Pilot Study of Novel Candidate Genes for Preeclampsia” case-control research study (University of Southern California) were retrospectively recruited from delivery logs at the Los Angeles County and University of Southern California Women’s and Children’s Hospital or during their postpartum stay at the Women’s and Children’s Hospital between the years of 1999 and 2008. Women with lupus, chronic renal disease, multiple gestation, or sickle cell disease/trait were excluded from participation. Preeclampsia was also defined in alignment with the blood pressure and proteinuria criteria outlined by the American College of Obstetricians and Gynecologists in 2002 (please see criteria above) [25]. Severe preeclampsia was defined as SBP ≥ 160 mmHg or DBP ≥ 110 mmHg on ≥ two occasions at least 6 hours apart and proteinuria ≥ 500 mg/dL in a 24-hr specimen or ≥ 3 on a urine dipstick. HELLP syndrome was defined by the presence of hemolysis (abnormal peripheral smear or LDH ≥ 600), elevated liver enzymes (ALT and/or AST ≥ 70), and low platelets (≤ 100,000). Partial HELLP was diagnosed when only two of the three signs were present. For this replication study, all participants with HELLP or partial HELLP also met the preeclampsia criteria. Uncomplicated pregnancies were defined as those pregnancies without significant hypertension. Based on self-report, all 175 Latina participants included in the current study were categorized as Hispanic Whites. The “Pilot Study of Novel Candidate Genes for Preeclampsia” research study was approved by the University of Southern California IRB, and use of the DNA samples in the current study was approved by the University of Pittsburgh IRB.

#### 2.2. Genotyping methods

DNA aliquots obtained from the Norwegian cohort participants were extracted from EDTA peripheral blood samples using the MagNA Pure LC DNA Isolation Kit (Roche Diagnostics GmbH, Mannheim, Germany). DNA aliquots obtained from the Latina cohort participants were extracted from either peripheral blood samples or buccal swabs using the QIAamp DNA mini kit (Qiagen, Valencia, CA, USA), saliva samples using the Oragenekit (DNA Genotek, Ontario, Canada), or mouthwash samples using a phenyl-chloroform protocol previously described [26]. The genomic DNA aliquots provided by both the Norwegian and Latina cohorts were whole genome amplified using the GE™ Healthcare illustra™ GenomiPhi V2 DNA Amplification kit (GE Healthcare Life Sciences, Little Chalfont, United Kingdom). Methods for polymorphism selection, genotyping (iPLEX® Gold-SNP Genotyping assay (Sequenom® Inc, San Diego, CA; now Agena Bioscience, San Diego, CA), and data reliability checks are fully described by Bell et al. (2013) [24] (Table 1). The UCSC Genome Browser was used to identify tSNP positions within the genome (Human Assembly Dec. 2013 [24] (Table 1). The UCSC Genome Browser, dbSNP) and conducted a literature search that combined gene name AND rs number as keywords (e.g., TGFβR2 AND rs1129420) to learn about the potential function of the tSNPs that were significantly associated with pregnancy outcome in the multivariate models.
Table 1
Tagging and potentially functional single nucleotide polymorphisms assayed.

<table>
<thead>
<tr>
<th>Candidate Gene</th>
<th>SNP rs numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endoglin (ENG)</td>
<td>rs10987746, rs10819309, rs10765055, rs11792480, rs10121110</td>
</tr>
<tr>
<td>Transforming growth factor beta 1 (TGFβ1)</td>
<td>rs8179181, rs4803455, rs11466314, rs1800469, rs1800468, rs4803457</td>
</tr>
<tr>
<td>Transforming growth factor beta 1 receptor (TGFβR1)</td>
<td>rs6478974, rs10739778, rs420549</td>
</tr>
<tr>
<td>Activin receptor like kinase 1 (ALK1)</td>
<td>rs3795178, rs11169955, rs706819</td>
</tr>
<tr>
<td>Transforming growth factor beta receptor 2 (TGFβR2)</td>
<td>rs3087465, rs6550005, rs11129420, rs6802220, rs17025785, rs4522809, rs4955212, rs5020833, rs6809777, rs12487185, rs11924422, rs13083813, rs13075948, rs1155708, rs13086588, rs2082224, rs1079885, rs1036097, rs6995435, rs6792117, rs479794, rs3773640, rs3773664, rs3773645, rs3773652, rs2043136, rs1346907, rs876688, rs877572, rs9843942, rs3773663, rs744751</td>
</tr>
</tbody>
</table>

Abbreviations: SNP, single nucleotide polymorphism. Notes: Bolded SNPs indicate the high priority SNPs analyzed in Tier 1 of this study. Underlined SNPs indicate the SNPs that were not successfully genotyped with iPLEX®.

2.3. Statistical analysis

Demographic and clinical characteristics were compared between cases and controls using parametric and non-parametric tests (e.g., Student’s t-test, Fisher’s exact test). For the main analyses, we utilized a tiered analytic approach. For the first analytic tier, we used Chi-square or Fisher’s exact tests (univariate analysis) and multivariate regression models to evaluate three tSNPs that were found to be significantly associated with pregnancy outcome in our previous study (rs11792480, rs10121110, TGFβR2: rs6550005) and preeclampsia status were not replicated in both the unadjusted and adjusted analyses (p’s > 0.05) (Refer to Supplemental Materials for tSNP/SNP genotype distributions). These associations were also non-significant in our exploratory subgroup analyses of early and late-onset preeclampsia.

3. Results

3.1. Associations between ENG (rs11792480, rs10121110) and TGFβR2 (rs6550005) tSNPs and preeclampsia not replicated in Norwegian and Latina cohorts

The demographic and clinical characteristics for the Norwegian and Latina cohorts are presented in Tables 2 and 3, respectively. In our previous genetic association study, we found that variation in two endoglin tSNPs (rs11792480, rs10121110) was associated with susceptibility to protection from preeclampsia in an American Caucasian cohort, while variation in a TGFβR2 tSNP (rs6550005) was associated with preeclampsia in both an American Caucasian and an African American cohort. We found additional tSNPs that were associated with preeclampsia in the African American cohort that were not found to be associated with preeclampsia in the American Caucasian cohort, but here we focused on tSNPs that were associated with preeclampsia in the American Caucasian cohort given that our replication cohorts were not of African ancestry.

In both the Norwegian and Latina cohorts, associations between each of the three tSNPs of interest (ENG: rs11792480, rs10121110; TGFβR2: rs6550005) and preeclampsia status were not replicated in both the unadjusted and adjusted analyses (p’s > 0.05) (Refer to Supplemental Materials for tSNP/SNP genotype distributions). These associations were also non-significant in our exploratory subgroup analyses of early and late-onset preeclampsia.

3.2. Exploratory analysis provides additional support for the ENG pathway’s involvement in preeclampsia

Six intronic tSNPs in two ENG pathway candidate genes were found to be significantly associated with preeclampsia status in the Norwegian cohort, TGFβR1 (rs6478974) and TFFR2 (rs11129420, rs6802220, rs1155708, rs3773640, rs3773664) (Table 4) (Refer to Supplemental Materials for tSNP/SNP genotype distributions).

Our exploratory subgroup findings suggest that certain tSNPs in two ENG pathway candidate genes were also found to be significantly associated with preeclampsia in the Latina cohort, ALK1 (rs706819, 3’ UTR variant) and TGFBR2 (rs984394, intronic) (Table 5). Because tSNP genotype distributions may differ by preeclampsia subgroup, we also compared preterm and term preeclampsia distributions to control distributions.

4. Discussion

Preeclampsia is a heterogeneous disorder of pregnancy whose etiology may stem from a variety of underlying mechanisms. Women who develop early-onset and late-onset preeclampsia differ on average on several aspects, including their risk of future cardiovascular disease [28], but a dysfunctional placenta is seen as a major common pathophysiological pathway for triggering the maternal syndrome [29]. Some of the proposed mechanisms include defects in placental implantation, oxidative stress, a susceptible maternal constitution, and anti-angiogenic/modulatory factors that are thought to have a role in the development of preeclampsia, and associations have been found at the gene, mRNA, and protein level. We have previously shown that genetic variants in the ENG pathway are associated with an increased susceptibility to preeclampsia in both American Caucasian and African American women [24]. Others have shown that mRNA expression of ENG is increased in the placenta, maternal blood (cellular and non-cellular components), and maternal microvascular endothelial cells [13–20] while levels of sENG protein in the maternal circulation are elevated weeks to months before overt disease in women with preeclampsia [21,22]. Although the evidence to support the alterations in ENG mRNA...
Table 2
Demographic and clinical characteristics for Norwegian cohort.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Cases (n = 77)</th>
<th>Controls (n = 63)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal age at delivery, years [mean ± SD]</td>
<td>31.8 ± 4.7</td>
<td>32.0 ± 4.6</td>
<td>.75</td>
</tr>
<tr>
<td>Nulliparous [n (%)]</td>
<td>48 (62.3%)</td>
<td>35 (55.6%)</td>
<td>.42</td>
</tr>
<tr>
<td>Gestational age at delivery (weeks) [median (min-max)]</td>
<td>33.9 (24.9–39.4)</td>
<td>38.7 (37.1–41.9)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Pre-pregnancy BMI (kg/m²) [median (min-max)]</td>
<td>23.7 (18.9–41.1)</td>
<td>22.3 (17.4–30.7)</td>
<td>.004</td>
</tr>
<tr>
<td>Smoked during pregnancy [n (%)]</td>
<td>8 (10.4%)</td>
<td>6 (9.5%)</td>
<td>.87</td>
</tr>
</tbody>
</table>

Abbreviations: BMI (kg/m²), body mass index (kilograms/square meters). Notes: Pre-pregnancy BMI includes BMI pre-pregnancy or within the first trimester of pregnancy.

Table 3
Demographic and clinical characteristics for Latina cohort.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Cases (n = 69)</th>
<th>Controls (n = 106)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal age at delivery, years [median (min-max)]</td>
<td>26.0 (17.0, 45.0)</td>
<td>25.0 (15.0, 42.0)</td>
<td>.3</td>
</tr>
<tr>
<td>Nulliparous [n (%)]</td>
<td>31 (44.9%)</td>
<td>41.0</td>
<td>.2</td>
</tr>
<tr>
<td>Gestational age at delivery (weeks) [median (min-max)]</td>
<td>37.0 (25.0, 40.0)</td>
<td>40.0 (37.0, 42.0)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Pre-pregnancy BMI (kg/m²) [median (min-max)]</td>
<td>26.8 (19.6, 39.6)</td>
<td>24.2 (17.0, 41.8)</td>
<td>.02</td>
</tr>
<tr>
<td>Smoked during pregnancy [n (%)]</td>
<td>7 (10.9%)</td>
<td>4 (4.0%)</td>
<td>.08</td>
</tr>
</tbody>
</table>

Abbreviations: BMI (kg/m²), body mass index (kilograms/meters squared). Notes: * n = 11 controls and n = 2 cases were missing pre-pregnancy BMI values. ** n = 6 controls and n = 5 cases were missing values on smoking during pregnancy.

In this study we investigated the association between maternal genetic variation in the ENG pathway and susceptibility to/protection from preeclampsia in independent, well-characterized preeclampsia cohorts. We first sought to validate our genetic association findings from an American Caucasian preeclampsia cohort (ENG tSNPs rs11792480, rs10121110; TGFBIR2 tSNP rs6550005) in Norwegian and Latina cohorts. We further explored the other gene candidates from the ENG pathway for new genetic associations with preeclampsia in these cohorts. We were unable to validate/replicate the SNP-specific previous findings in these independent cohorts. However, we identified other associations between ENG pathway candidate genes and preeclampsia in the Norwegian (TGFBIR1[ALK5], TGFBIR2) and Latina (ALK1, TGFBIR2) cohorts in our adjusted models. We have previously demonstrated that variation in TGFBIR1[ALK5] was associated with preeclampsia in an African American cohort [24]. Moreover, we have demonstrated that genetic variation in TGFBIR2 is associated with susceptibility to/protection from preeclampsia in all of the populations we have studied. Although the TGFBIR2 tSNPs that were associated with a particular cohort may differ, these findings, along with the other associations in the Norwegian and Latina cohorts, provide additional evidence to support the ENG pathway’s role/involvement in the development of preeclampsia.

4.1. Failure to validate/replicate associations between genetic variation in ENG and TGFBIR2

Several potential explanations could account for our inability to validate our previous results in independent cohorts. First, these

Table 4
Multivariate logistic regression results in the Norwegian cohort: ENG Pathway Candidate Gene tSNPs Associated (p < .05) with Pregnancy Outcome.

<table>
<thead>
<tr>
<th>Gene/tSNP</th>
<th>Genotype Distributions</th>
<th>HWE</th>
<th>Genotype Comparisons</th>
<th>aOR [95% CI] **</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGFBR1[ALK5]:rs6478974</td>
<td>AA: 16 (20.8) TT: 16 (20.8) TA: 45 (58.4)</td>
<td>NS</td>
<td>TA vs AA</td>
<td>2.38 [1.007–5.60]</td>
<td>0.05</td>
</tr>
<tr>
<td>TGFBR2[rs11129420] MAF: T = 0.48</td>
<td>AA: 12 (16.4) TT: 16 (21.9) TA: 45 (61.6)</td>
<td>0.02***</td>
<td>TT vs AA</td>
<td>5.01 [1.39–18.01]</td>
<td>0.01</td>
</tr>
<tr>
<td>TGFBR2[rs6802220] MAF: A = 0.43</td>
<td>GG: 24 (32.9) AA: 6 (8.2) AG: 43 (58.9)</td>
<td>0.04***</td>
<td>AA vs. GG</td>
<td>0.17 [0.05–0.64]</td>
<td>0.009</td>
</tr>
<tr>
<td>TGFBR2[rs1155708] MAF: A = 0.40</td>
<td>GG: 32 (43.8) AA: 5 (6.8) GA: 36 (49.3)</td>
<td>NS</td>
<td>AA vs. GG</td>
<td>0.27 [0.09–0.85]</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Abbreviations: tSNP, tagging single nucleotide polymorphism; MAF, minor allele frequency; HWE, Hardy-Weinberg Equilibrium; aOR, adjusted odds ratio; CI, confidence interval; NS, non-significant (p > .05).

Note: * HWE assessed with Chi-square or Exact tests.
** The reference group for Pregnancy Outcome is the Case group. The multivariate logistic regression models were adjusted for maternal age, parity, pre-pregnancy BMI, smoking, and infant sex.
*** HWE was also assessed separately in cases and controls. The control group was in HWE while the case group was out of HWE.
Table 5
Multivariate logistic regression results in the Latina cohort: ENG pathway Candidate gene tSNPs associated (p < .05) with pregnancy outcome.

<table>
<thead>
<tr>
<th>Gene/tSNP</th>
<th>Genotype</th>
<th>HWE</th>
<th>Genotype</th>
<th>aOR [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALK1/ρ</td>
<td>GG: 24</td>
<td>35</td>
<td>GG vs GA</td>
<td>3.59 [1.51-8.53]</td>
</tr>
<tr>
<td></td>
<td>GG: 19</td>
<td>18</td>
<td></td>
<td>0.02</td>
</tr>
<tr>
<td>s706819</td>
<td>AA: 21</td>
<td>14</td>
<td>AA + GA</td>
<td>2.62 [1.20-5.72]</td>
</tr>
<tr>
<td></td>
<td>Aa: 33</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAF:</td>
<td>G: 0.48</td>
<td>50</td>
<td>GA vs GA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A: 53</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>s6478974</td>
<td>AA: 29</td>
<td>50</td>
<td>GG vs GA</td>
<td>0.37 [0.14-1.0]</td>
</tr>
<tr>
<td></td>
<td>Aa: 29</td>
<td>50</td>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td>TGFβR2/ρ</td>
<td>AA: 29</td>
<td>50</td>
<td>GG vs GA</td>
<td>0.27 [0.09-0.76]</td>
</tr>
<tr>
<td></td>
<td>Aa: 29</td>
<td>50</td>
<td></td>
<td>0.02</td>
</tr>
</tbody>
</table>

Abbreviations: tSNP, tagging single nucleotide polymorphism; MAF, minor allele frequency; HWE, Hardy-Weinberg Equilibrium; aOR, adjusted odds ratio; CI, confidence interval; NS, non-significant (p > .05).

Notes:
* HWE assessed with Chi-square or Exact tests.
** The reference group for Pregnancy Outcome is the Case group. The multivariate logistic regression models were adjusted for maternal age, parity, pre-pregnancy BMI, smoking, and infant sex.

cohort may not have been sufficiently powered to detect statistically significant associations when they truly existed. Second, tSNPs that were selected based on Caucasian ancestry (CEU population-European descendants living in Utah) may not be as informative in the Norwegian and Latina cohorts if the haplologs tagged by the tSNPs differ by ancestry. Finally, differences in the definition of preeclampsia phenotype may underlie our failure to validate our previous results. In the original cohort, a diagnosis of preeclampsia was based on hypertension, proteinuria, and hyperuricemia. Hyperuricemia was included in the original cohort’s research definition as a way to identify a more homogeneous and potentially more severe form of preeclampsia. In both the Norwegian and Latina cohorts, hyperuricemia was not included as a criterion for the preeclampsia phenotype. Given the heterogeneous nature of preeclampsia, these cohorts may represent different preeclampsia subtypes with differing etiologies.

4.2. Additional evidence to support the ENG pathway’s involvement in preeclampsia

Genetic variation in TGFβR1 (ALK5) and TGFβR2 was associated with susceptibility to/protection from preeclampsia in the Norwegian cohort, while genetic variation in ALK1 and TGFβR2 was associated with susceptibility to/protection from preeclampsia in the Latina cohort. Seven of the eight significantly associated tSNPs (Table 5; rs6478974, rs1129420, rs680220, rs1155708, rs3773640, rs3773663, rs9843942) are intronic, while the other SNP was located in the 3’ UTR of the gene (rs706819). Of the eight tSNPs, we were only able to locate a small amount of literature on TGFβR1 tSNP rs6478974, which has been associated with endometrial and gastric cancer in other populations [31,32].

TGFβR1 (ALK5) and ALK1 are type 1 receptors of the TGFβ signaling cascade, which form a heteromeric complex with TGFβR2, the type 2 receptor of the TGFβ signaling cascade. In most cells, TGFβ signals via ALK5, but TGFβ can also signal via ALK1 in endothelial cells. The TGFβ1 ligand binds to these complexes, and propagation of the cell signal is involved in the control of a variety of processes including cellular proliferation, migration, and differentiation [10].

In vitro research has demonstrated that TGFβ1 is a negative regulator of trophoblast invasion toward the invasive phenotype [33-36], and that endoglin is needed to facilitate this inhibitory effect [11,12]. Shallow trophoblast invasion and failed remodeling of the spiral arteries, affects mostly the early-onset subtype of preeclampsia [37-40]. Furthermore, implantation and placental development is dependent on the successful interaction between fetal/placental and maternal/deciudal factors [34]. As such, it is plausible that maternal or fetal genetic variation in any of the members of the ENG/TGFβ pathway could impact TGFβ signaling and the regulation of trophoblast invasion. Moreover, it is possible that maternal genetic variation in this pathway could also impact systemic endothelial function in the mother, as this pathway is involved in the maintenance of vascular homeostasis, including vasomotor tone [10,41,42].

4.3. Study limitations

In addition to small sample size and differences in ancestry that were previously discussed, there were several other limitations associated with this study. First, the cohorts’ sample sizes were further reduced when we carried out our exploratory preeclampsia subgroup analyses, which limited our ability to adequately evaluate preeclampsia subgroups. Second, we did not adjust for multiple comparisons and therefore we cannot rule out the potential for type 1 error. As such, follow-up studies with larger sample sizes, including larger preeclampsia subgroups, will be needed to validate these results. Third, we were unable to successfully genotype three SNPs (TGFβ1 rs1800468; TGFβR2 rs1078985 and rs995435).

5. Summary

Overall, our results provide further support for the involvement and investigation of the endoglin pathway in preeclampsia, yet additional research with larger sample sizes, including larger samples of different preeclampsia subgroups, is needed. Due to the relatively low frequency of preeclampsia, there is a need for investigators across the globe to collaborate and share data/biological samples with each other so that we can continue to put preeclampsia’s puzzle pieces together.

Acknowledgements

We would like to thank the Global Pregnancy Collaboration, which provided us with the Norwegian cohort samples for this project. We would also like to acknowledge the University of Pittsburgh HSCRF Genomics Research Core for their iPLEX services. We would also like to thank Ms. Jacqueline Lee (undergraduate nursing student at the University of Pittsburgh) for her help with the construction of the Supplemental data file.
Funding

This work was supported by the National Institutes of Health (F32NR014622, T32NR009759) and the Preeclampsia Foundation.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.preghy.2017.10.005.

References