McGinnis, R.; Steinthorsdottir, V.; Williams, N.O.; Thorleifsson, G.; Shooter, S.; Hjartardottir, S.; Bumpstead, S.; Stefansdottir, L.; Hildyard, L.; Sigurdsson, J.K.; Kemp, J.P.; Silva, G.B.; Thomsen, L.C.V.; Jskelinen, T.; Kajantie, E.; Chappell, S.; Kalsheker, N.; Moffett, A.; Hiby, S.; Lee, W.K.; Padmanabhan, S.; Simpson, N.A.B.; Dolby, V.A.; Staines-Urias, E.; Engel, S.M.; Haugan, A.; Trogstad, L.; Svyatova, G.; Zakhidova, N.; Najmutdinova, D.; FINNPEC Consortium, .; GOPEC Consortium, .; Dominiczak, A.F.; Gjessing, H.K.; Casas, J.P.; Dudbridge, F.; Walker, J.J.; Pipkin, F.B.; Thorsteinsdottir, U.; Geirsson, R.T.; Lawlor, D.A.; Iversen, A.C.; Magnus, P.; Laivuori, H.; Stefansson, K.; Morgan, L.; , COLLABORATORS; Laivuori, H.; Heinonen, S.; Kajantie, E.; Kere, J.; Kivinen, K.; Pouta, A.; Morgan, L.; Pipkin, F.B.; Kalsheker, N.; Walker, J.J.; Macphail, S.; Kilby, M.; Habiba, M.; Williamson, C.; O'Shaughnessy, K.; O'Brien, S.; Cameron, A.; Poston, L.; Miedzybrodzka, Z.; Redman, C.W.G.; Farrall, M.; Caulfield, M.; Dominiczak, A.F. (2017)

Variants in the fetal genome near _FLT1_ are associated with risk of preeclampsia

Ralph McGinnis[1*†], Valgerdur Steinthorsdottir[2*†], Nicholas O. Williams[1*], Gudmar Thorleifsson[2], Scott Shooter[1], Sigrun Hjartardottir[3], Suzannah Bumpstead[1], Lilja Stefansdottir[2], Lucy Hildyard[1], Jon K. Sigurdsson[2], John P. Kemp[4,5], Gabriela B. Silva[6,7], Liv Cecilie V. Thomsen[6,8], Tiina Jääskeläinen[9], Eero Kajantie[10,11,12], Sally Chappell[13], Noor Kalsheker[13], Ashley Moffett[14], Susan Hiby[14], Wai Kwong Lee[15], Sandosh Padmanabhan[15], Nigel A. B. Simpson[16], Vivien A. Dolby[16], Eleonora Staines-Urias[17,18], Stephanie M. Engel[19], Anita Haugan[20], Lill Trogstad[20], Gulnara Svyatova[21], Nodira Zakhidova[22], Dilbar Najmutdinova[23], The FINNPEC Consortium[24], The GOPEC Consortium[24], Anna F. Dominiczak[15], Håkon K. Gjessing[20,25], Juan P. Casas[26], Frank Dudbridge[17], James J. Walker[16], Fiona Broughton Pipkin[27], Unnur Thorsteinsdottir[2,28], Reynir T. Geirsson[3], Debbie A. Lawlor[4,29], Ann-Charlotte Iversen[6], Per Magnus[20], Hannele Laivuori[9,30,31], Kari Stefansson[2,28], Linda Morgan[13,†]

[1]Wellcome Trust Sanger Institute, Cambridge, UK. [2]deCODE Genetics/Amgen, Reykjavik, Iceland. [3]Department of Obstetrics and Gynecology, Landspitali University Hospital, Reykjavik, Iceland. [4]MRC Integrative Epidemiology Unit at the University of Bristol, UK. [5]University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, Australia. [6]Centre of Molecular Inflammation Research (CEMIR) and Department of Cancer Research and Molecular Medicine, Norwegian University of Science and Technology, Trondheim, Norway. [7]St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway. [8]Department of Obstetrics and Gynecology, Haukeland University Hospital, Bergen, Norway. [9]Medical and Clinical Genetics, University of Helsinki and Helsinki University Hospital, Finland. [10]National Institute for Health and Welfare, Helsinki, Finland. [11]Children's Hospital, University of Helsinki and Helsinki University Hospital, Finland. [12]PEDEGO Research Unit, Oulu University Hospital and University of Oulu, Finland. [13]School of Life Sciences, University of Nottingham, UK. [14]Department of Pathology, University of Cambridge, UK. [15]BHF Glasgow Cardiovascular Research Centre, Institute of Cardiovascular and Medical Sciences, University of Glasgow, UK. [16]Leeds Institute of Biomedical and Clinical Sciences, University of Leeds, UK. [17]Department of Non-communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, UK. [18]Nuffield Department of Obstetrics & Gynaecology, University of Oxford, UK. [19]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, USA. [20]Norwegian Institute of Public Health, Oslo, Norway. [21]Scientific Center of Obstetrics, Gynecology and Perinatology, Almaty, Kazakhstan. [22]Institute of Immunology, Uzbek Academy of Sciences, Tashkent, Uzbekistan. [23]Republic Specialized Scientific Practical Medical Centre of Obstetrics and Gynecology, Tashkent, Uzbekistan. [24]A full list of members and affiliations appears at the end of the paper. [25]Department of Global Public Health and Primary Care, University of Bergen, Norway [26]Farr Institute of Health Informatics, University College London, UK. [27]Medical School, University of Nottingham, UK. [28]Faculty of Medicine, School of Health Sciences, University of Iceland, Reykjavik, Iceland. [29]School of Social and Community Medicine, University of Bristol, UK. [30]Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Finland. [31]Obstetrics and Gynecology, University of Helsinki and Helsinki University Hospital, Finland.

*R.M, V.S. and N.O.W. contributed equally to this work

† Correspondence should be addressed to R.M. (rm2@sanger.ac.uk), V.S. (Valgerdur.Steinthorsdottir@decode.is) or L.M. (linda.morgan@nottingham.ac.uk)

43 **Preeclampsia, which affects approximately 5% of pregnancies, is a leading cause of maternal and**

44 **perinatal death[1]. The causes of preeclampsia remain unclear, but there is evidence for inherited**

45 **susceptibility[2]. Genome-wide association studies (GWAS) have not identified maternal sequence**

46 **variants of genome-wide significance which replicate in independent datasets[3,4]. We report the**

47 **first GWAS of offspring of preeclamptic pregnancies and discovery of the first genome-wide**

48 **significant susceptibility locus (rs4769613; P = 5.4x10^{-11}) in 4380 cases and 310,238 controls. The**

49 **locus is near the gene encoding Fms-like tyrosine kinase 1 (*FLT1*), providing biological support**

50 **since an isoform (sFlt-1) of placental origin is implicated in the pathology of preeclampsia[5]. The**

51 **strongest association is in pregnancies where preeclampsia developed in late gestation and**

52 **offspring birthweights exceeded the 10th centile. An additional nearby variant, rs12050029,**

53 **associates with preeclampsia independent of rs4769613. The newly discovered locus may enhance**

54 **understanding of the pathophysiology of preeclampsia and its subtypes.**

55

56 Our initial GWAS meta-analysis tested 7,476,169 sequence variants in 2,658 offspring of

57 preeclamptic pregnancies and 308,292 controls of European descent from Iceland (deCODE cohort)

58 and the UK (GOPEC and ALSPAC cohorts). We observed a single genome-wide significant association

59 (P=3.2×10^{-8}, rs4769613) located on chromosome 13 near the *FLT1* gene (Fig. 1a). We genotyped

60 rs4769613 and a correlated surrogate in 1722 independent cases and 1946 controls from Norway

61 and Finland along with 26 variants marking GWAS meta-analysis signals elsewhere in the genome

62 whose P values showed suggestive evidence of association (Supplementary Table 1). rs4769613 was

63 significantly associated with preeclampsia in the replication datasets (P=3.6×10^{-4}) and joint analysis

64 of GWAS and replication data yielded robust genome-wide association (P=5.4×10^{-11}) with an allelic

65 odds ratio (OR) of 1.21 for allele C (Table 1). Forest plots show the frequency of allele C is

66 consistently elevated in cases in all GWAS and replication datasets with no evidence of

67 heterogeneity (P_{het}=0.678; Supplementary Fig. 1). None of the other genotyped loci achieved

68 genome-wide significance (P<5x10^{-8}) in joint analysis of GWAS and replication data.

69

70 We then examined genomic features of the *FLT1* locus in detail (Fig. 1b) and found that some of the

71 association signal remained after conditioning out the effect of rs4769613 (Fig 1b, bottom panel).

72 This suggested that other variants near *FLT1* might associate with preeclampsia independent of

73 rs4769613. We therefore genotyped the replication datasets for 21 additional variants at the *FLT1*

74 locus, representing 9 linkage disequilibrium (LD) blocks (Supplementary Table 2; Supplementary Fig.

75 2). rs12050029 and surrogates in the same LD block were significant in combined analysis of GWAS

76 and replication data after *FLT1* region-wide correction for testing all common variants within 1 Mb of

77 rs4769613 (P=3.9×10^{-6}, Table 1). Table 1 shows that rs149427560 also achieved *FLT1* region-wide

78 significance but with an association signal weaker than rs12050029. In summary, our results imply

79 that in addition to rs4769613, other independent variants near *FLT1* may modulate preeclampsia

80 susceptibility.

81 As expected if risk allele rs4769613[C] increases susceptibility by acting through the fetal genome,

82 Supplementary Table 3 shows that allele C frequency in preeclampsia mothers is midway between

83 control frequency and the significantly elevated frequency in preeclampsia offspring. Preeclampsia

84 offspring also preferentially inherited rs4769613[C] from heterozygous parents in the only dataset

85 with DNA available for both parents, again implying that rs4769613[C] increases susceptibility by

86 acting on the fetal genome (Supplementary Table 4). To examine if rs4769613 exerts effects on *both*

87   the fetal and maternal genomes, we applied the EMIM algorithm which simultaneously evaluates
88   maternal cases, offspring and controls to calculate ORs for preeclampsia risk corresponding to one or
89   two risk alleles carried in the fetus (R_1, R_2) or in the mother (S_1, S_2) [6]. Fig. 2a shows that fetal ORs are
90   above 1.0 and that each fetal copy of rs4769613[C] increases these ORs. By contrast, maternal ORs
91   are near 1.0 and are not significant. We conclude that rs4769613 exerts influence primarily through
92   the fetal genome.
93
94   As there is evidence that genetic imprinting may operate in placental development[7] we examined
95   rs4769613 allele transmissions from heterozygous parents, but found no parental gender difference
96   in allele transmission in preeclampsia and hence no evidence for imprinting (χ^2=0.046, P=0.83;
97   Supplementary Table 5).   We also applied EMIM[6] to meta-analyse cohorts with available DNA from
98   one or both parents, but again found no evidence that maternal and paternal alleles at rs4769613
99   confer differential preeclampsia risk (P=0.90).
100
101   Sub-classifications of preeclampsia are based on clinical features, in particular gestation at diagnosis
102   and evidence of fetal growth restriction (FGR)[8]. Early-onset preeclampsia (EO-preeclampsia),
103   affecting 12-15% of all preeclamptic pregnancies and defined as onset before 34 weeks gestation, is
104   associated with higher maternal and perinatal mortality than later onset preeclampsia (LO-
105   preeclampsia). It has been proposed that LO-preeclampsia results predominantly from maternal
106   maladaptation to the physiological stresses of pregnancy, whilst EO-preeclampsia is primarily the
107   result of sub-optimal placental implantation into the uterine wall, leading to inadequate placental
108   perfusion and the release of damaging placental factors into the maternal circulation[8]. In keeping
109   with this, EO-preeclampsia is frequently associated with FGR, resulting in babies who are small for
110   gestational age (SGA) at birth. SGA defined as birthweight <10th centile is widely used as a surrogate
111   for FGR.
112   To assess the impact of rs4769613 on gestation at onset and fetal growth subtypes we noted that
113   rs4769613 risk allele C had higher frequency in LO-preeclampsia than EO-preeclampsia cases (case-
114   control OR 1.23 vs. 1.06) and found the difference was significant in case-case meta-analysis
115   (P=0.017, Fig. 2b). Similarly, allele C had higher frequency in nonSGA-preeclampsia than SGA-
116   preeclampsia cases (case-control OR 1.25 vs 1.10) and the difference was significant in case-case
117   comparison (P=0.019, Fig. 2b). Further division of cases in Fig 2b into the four possible subcategories
118   found that rs4769613 [C] confers greatest risk to LO+nonSGA cases (case-control OR=1.26, P=1.2×10^{-7}
119   ) and least risk to EO+SGA cases (case-control OR=1.03, P=0.72) with case-case comparison of the
120   two subcategories being significant (P=5.8×10^{-3}). In summary, the results indicate that rs4769613
121   exerts its greatest influence in pregnancies where preeclampsia develops in late gestation and
122   birthweights exceed the 10th centile. rs12050029 was also associated with LO-preeclampsia, but the
123   strength of the association did not differ between SGA- and nonSGA-preeclampsia (Supplementary
124   Fig. 3).
125
126   *FLT1* encodes a trans-membrane tyrosine kinase receptor Flt-1 that mediates angiogenesis
127   promoted by binding vascular endothelial growth factor (VEGFA) and placental growth factor (PlGF)[9].
128   The alternatively spliced soluble isoform sFlt-1 antagonizes angiogenesis by also binding VEGFA and
129   PlGF. During pregnancy, *FLT1* is mainly expressed in fetal trophoblasts which release sFlt-1 as the
130   most abundant isoform into the maternal circulation.   The excessive release of sFlt-1 in

131   preeclampsia appears to mediate widespread maternal endothelial dysfunction, manifesting as

132   hypertension, proteinuria, and vascular compromise to major organs. High sFlt-1 and low PlGF

133   concentrations are established markers of EO-preeclampsia[8], but our evidence that *FLT1*

134   polymorphisms are strongly associated with LO-preeclampsia suggests trophoblast function is also

135   important in this preeclampsia subgroup.

136

137   The signals around rs4769613 and rs12050029 are both located in placental enhancer regions (Fig.

138   1b) suggesting a mechanism by which variants could affect *FLT1* expression. We explored possible

139   association between fetal *FLT1* genotype and protein expression by placental immunohistochemistry

140   and intensity scanning in 37 preeclamptic and 44 control pregnancies. There was no detectable

141   association between fetal rs4769613 genotype and trophoblast Flt-1 and sFlt-1 expression in cases

142   (P=0.47) or controls (P=0.26). We also compared maternal serum sFlt-1 from the first or third

143   trimester with fetal rs4769613 genotype in mother-baby pairs from 242 control and 276

144   preeclamptic pregnancies. Control pregnancies exhibited a trend towards increasing maternal serum

145   sFlt-1 levels with each fetal copy of rs4769613[C] in the third trimester, which reached nominal

146   significance (P = 0.04), while in case pregnancies the levels were higher (P<0.001) but with no

147   detectable difference between genotype groups (P=0.47) (Fig. 2c). We did not have suitable

148   placental tissue for mRNA studies, but the Genotype-Tissue expression (GTEx) database

149   (www.gtexportal.org) does not provide evidence for rs4769613 or rs12050029 allele-specific

150   differences in *FLT1* expression in 42 tissues, although data for placental tissue are not recorded. The

151   evidence that fetal rs4769613 genotype affects maternal serum levels of sFlt-1 is therefore modest.

152   Subtle changes in sFlt-1 concentration driven by fetal *FLT1* genotype, as suggested by the data from

153   control pregnancies, may be masked in preeclampsia, where the overall levels are already high. Also,

154   this effect is minimal compared to the increase in serum sFlt-1 seen in preeclamptic pregnancies so

155   it may not reflect the role of the preeclampsia associated variants in the pathophysiology of

156   preeclampsia.

157

158   We explored whether the preeclampsia associated variants affected other diseases or traits by using

159   the deCODE database of common diseases and traits routinely measured at hospitals and clinical

160   laboratories (see Methods). Given that rs4769613 and rs12050029 are not in LD (Supplementary Fig.

161   2), it is noteworthy that the only significant database association for both variants was red blood cell

162   (RBC) count (P=5.0 x 10^{-4} and P=1.5 x 10^{-7} for rs4769613 and rs12050029 respectively), where the

163   preeclampsia risk allele consistently associated with reduced RBC count (Supplementary Table 6).

164   The RBC association with *FLT1* is intriguing since its VEGF ligand has previously been implicated in

165   regulation of erythropoiesis, but the mode and sites of action are complex[10,11]. Our preeclampsia

166   results suggest that in the fetus the two variants lead to an increase in sFlt-1, while the same alleles

167   are associated with reduced RBC count in the general (non-pregnant) population. The effect on both

168   preeclampsia and RBC is consistent with the variants acting through the neighbouring *FLT1*.

169   We note that SNP rs4769613 is located between *FLT1* and *POMP,* which encodes proteasome

170   maturation protein, a ubiquitously expressed protein involved in proteasome assembly and MHC

171   class I antigen presentation[12]. We cannot exclude the possibility that sequence variants at this locus

172   affect expression of *POMP* or more distant genes, but the GTEx database does not provide any

173   evidence to support this contention.

174

175   Evidence presented here implies that altered trophoblastic *FLT1* expression is not merely a
176   secondary consequence of placental pathology in preeclampsia, but is central to its aetiology. A role
177   for fetal sequence variants in susceptibility to preeclampsia is consistent with patterns of inheritance
178   implicating both maternal and paternal factors[2]. The fetal *FLT1* gene has been indirectly implicated
179   previously in pregnancies with fetal trisomy 13, which are associated with increased placental
180   expression of sFlt-1, and an increased incidence of preeclampsia[13]. sFlt-1 is a marker of placental
181   malfunction, a hallmark of EO-preeclampsia[8]; our observation that *FLT1* genotype is associated even
182   more strongly with LO-preeclampsia implies that placental pathology is also a feature of late-onset
183   disease. The variants we describe provide tools for experimental testing of whether, how, when and
184   where they affect *FLT1* expression, and how this relates to the pathophysiology of preeclampsia and
185   its subtypes. The discovery of sequence variants in the genome of the fetus that increase the risk of
186   disease in the mother is an ultimate demonstration of the closeness of the remarkable symbiosis we
187   call pregnancy.
188

189   **Acknowledgements**

218 laboratory technicians, clerical workers, research scientists and receptionists who made this study
219 possible.
220

221 **Author contributions**
222 Manuscript preparation: R.M, V.S, N.O.W, L.C.V.T, A-C.I, L.M. All authors contributed to critical
223 analysis and revision of the manuscript.
224 Study design: R.M, V.S, N.O.W, T. J, S.l, N.K, N.A.B.S, V.A.D, E.S-U, S.M.E, A.H, L.T, G.S, N.Z, D.N,
225 A.F.D, H.K.G, J.P.C, F.D, J.J.W, F.B.P, U.T, D.A.L, A-C.I, P.M, H.L, K.S, L.M.
226 Phenotyping: S.H, G.B.S, L.C.V.T, T.J, E.K, FINNPEC Consortium, GOPEC Consortium, J.J.W, F.B.P,
227 R.T.G, D.A.L, H.L, L.M.
228 Genotyping and quality control: S.S, S.B, L.H, W.K.L, S.P, U.T, H.L.
229 GWAS data analysis: R.M, V.S, N.O.W, G.T, S.S, LS, J.K.S, J.P.K.
230 Immunohistochemistry: G.B.S, L.C.V.T, A-C.I.
231 Biomarker measurement: H.L.
232 Interpretation: R.M, V.S, N.O.W, G.T, J.P.K, L.C.V.T, T.J, E.K, S.C, N.K, N.S, V.D, E.S-U, L.T, H.K.G, F.D,
233 J.J.W, F.B.P, D.A.L, A-C.I, P.M, H.L, K.S, L.M.
234

235 **Competing financial interests statement**

236 Valgerdur Steinthorsdottir, Gudmar Thorleifsson, Lilja Stefansdottir, Jon K. Sigurdsson, Unnur
237 Thorsteinsdottir and Kari Stefansson are employees of the biotechnology firm deCODE Genetics, a
238 subsidiary of Amgen. Debbie A. Lawlor has received industry funding for biomarker research
239 unrelated to this paper from Medtronic, Roche Diagnostics, and Ferring Pharmaceuticals.
240

241 **Consortia**

242 **GOPEC Consortium**: Linda Morgan[13], Fiona Broughton Pipkin[27], Noor Kalsheker[13], James. J. Walker[16],
243 Sheila Macphail[32], Mark Kilby[33], Marwan Habiba[34], Catherine Williamson[35], Kevin O'Shaughnessy[36],
244 Shaughn O'Brien[37], Alan Cameron[38], Lucilla Poston[35], Zofia Miedzybrodzka[39], Christopher W. G.
245 Redman[18], Martin Farrall[40], Mark Caulfield[41], Anna F. Dominiczak[15]

246 **FINNPEC Consortium**: Hannele Laivuori[9,30,31], Seppo Heinonen[31], Eero Kajantie[10,11,12], Juha Kere[42,43],
247 Katja Kivinen[44], Anneli Pouta [10,45]

248 [32]Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. [33]The Centre for
249 Women's & Newborn Health, College of Medical and Dental Sciences, University of Birmingham, UK.
250 [34]University of Leicester, Leicester, UK. [35]Division of Women's Health, Kings College London, UK.
251 [36]Department of Medicine, University of Cambridge, UK. [37]Keele University School of Medicine,
252 Stoke-on-Trent, UK. [38]Institute of Cardiovascular and Medical Sciences, University of Glasgow, UK.
253 [39]College of Life Sciences and Medicine, University of Aberdeen, UK. [18]Nuffield Department of
254 Obstetrics and Gynaecology, University of Oxford, UK. [40]Radcliffe Department of Medicine,
255 University of Oxford, UK. [41]William Harvey Research Institute, Barts and the London School of
256 Medicine and Dentistry, London, UK. [42]Department of Biosciences and Nutrition, Karolinska
257 Institutet, Huddinge, Sweden. [43]Molecular Neurology Research Program, Research Programs Unit,
258 Helsinki, Finland. [44]Division of Cardiovascular Medicine, University of Cambridge, UK. [45]Department

259 of Obstetrics and Gynaecology and MRC Oulu, Oulu University Hospital and University of Oulu,
260 Finland.

261

262 **References**

263 1. Souza, J. P. *et al.* Moving beyond essential interventions for reduction of maternal mortality (the
264   WHO Multicountry Survey on Maternal and Newborn Health): a cross-sectional study. *Lancet*
265   **381**, 1747-1755 (2013).
266 2. Cnattingius, S. Reilly, M. Pawitan, Y. & Lichtenstein, P. Maternal and fetal genetic factors account
267   for most of familial aggregation of preeclampsia: a population-based Swedish cohort study. *Am.*
268   *J. Med. Genet. A.* **130A,** 365-371 (2004).
269 3. Johnson, M. P. *et al.* Genome-wide association scan identifies a risk locus for preeclampsia on
270   2q14, near the inhibin, beta B gene. *PLoS One* **7**, e33666. doi: 10.1371/journal.pone.0033666
271   (2012).
272 4. Zhao, l. *et al.* Genome-wide association study identifies a maternal copy-number deletion in
273   PSG11 enriched among preeclampsia patients. *BMC Pregnancy Childbirth* **12**, 61 (2012).
274 5. Maynard, S. E. *et al*. Excess placental soluble fms-like tyrosine kinase 1 (sFlt1) may contribute to
275   endothelial dysfunction, hypertension, and proteinuria in preeclampsia. *J. Clin. Invest.* **111**, 649-
276   658 (2003).
277 6. Ainsworth, H. F. *et al.* Investigation of maternal effects, maternal-fetal interactions and parent-
278   of-origin effects (imprinting), using mothers and their offspring. *Genet. Epidemiol.* **35,** 19-45
279   (2011).
280 7. Monk, D. Genomic imprinting in the human placenta. *Am. J. Obstet. Gynecol.* **213,** S152-S162
281   (2015).
282 8. Staff A. C. *et al.* Redefining preeclampsia using placenta-derived biomarkers. *Hypertension* **61**,
283   932-942 (2013).
284 9. Cerdeira, A. S. & Karumanchi, S. A. Angiogenic factors in preeclampsia and related
285   disorders. *Cold Spring Harb. Perspect. Med.* pii: a006585. doi: 10.1101/cshperspect.a006585
286   (2012).
287 10. Tam, B. Y. *et al.* VEGF modulates erythropoiesis through regulation of adult hepatic
288   erythropoietin synthesis. *Nat. Med.* **12,** 793-800 (2006).
289 11. Rehn, M. *et al.* Hypoxic induction of vascular endothelial growth factor regulates murine
290   hematopoietic stem cell function in the low-oxygenic niche. *Blood* **118,** 1534-1543 (2011).
291 12. Heink, S. Ludwig, D. Kloetzel, P. M. Krüger, E. IFN-gamma-induced immune adaptation of the
292   proteasome system is an accelerated and transient response. *Proc. Natl. Acad. Sci. USA.* **102,**
293   9241-9246 (2005).
294 13. Bdolah, Y. *et al.* Circulating angiogenic proteins in trisomy 13. *Am. J. Obstet. Gynecol.* **194,** 239-
295   245 (2006).
296

297   **Figure legends**

298

299   **Figure 1 | Manhattan Plots showing GWAS results across all autosomes and detailed view near**

300   ***FLT1* on chromosome 13**.

301   a) Genome-wide Manhattan plot showing strength of association with PE in GWAS meta-analysis

302   plotted as -log10(P value) on the y-axis and corresponding variant position on the x-axis.   A single

303   peak whose apex is the sentinel SNP rs4769613 near *FLT1* on chromosome 13 crosses the blue line

304   denoting genome-wide significance ($P = 5 \times 10^{-8}$).   Variants within 100 Kb of rs4769613 are coloured

305   purple.

306   b) Detailed view near *FLT1* highlighting variants in Table 1. Panels from top to bottom show:

307   unconditional -log10(P value) from GWAS meta-analysis and recombination rate shown as a blue

308   line quantified by right hand y-axis; pattern of regional Linkage Disequilibrium (LD) shown by

309   pairwise values of the LD metric D'; gene names with approximate length and position; the

310   corresponding chromatin state annotations for selected Epigenome Roadmap tissues

311   (http://egg2.wustl.edu/roadmap/web_portal/chr_state_learning.html#core_15state); landscape of

312   inferred chromatin interactions for 31 tissue types using an integrated method for predicting

313   enhancer targets (IM-PET) (http://4dgenome.research.chop.edu/); conditional -log10(P value) for

314   GWAS meta-analysis using logistic regressions with rs4769613 as a covariate.

315

316   **Figure 2 | Key observations about rs4769613 in relation to preeclampsia.**

317   a) Forest plot showing Odds Ratio (OR) and 95% confidence intervals (95% CI) calculated by the
318   EMIM algorithm[6] for preeclampsia risk conferred by one or two copies of rs4769613 risk allele C
319   carried in the fetus (R_1, R_2) or carried in the mother (S_1, S_2). Individual datasets (GOPEC, MoBa,
320   FINNPEC) and meta-analysis across the datasets ([Meta]) show the OR is increased by each fetal
321   copy of risk allele C where for R_1 P=3.7x10^{-4} and for R_2 P=3.7x10^{-9}. By contrast, the OR for maternal
322   copies (S_1, S_2) are not significantly different from 1 implying that, after accounting for fetal copies,
323   maternal copies of allele C confer no additional increased risk of preeclampsia.

324   b) Forest Plot for preeclampsia subtypes defined by early and late onset (EO-PE, LO-PE) and by
325   birthweight that is small-for-gestational age (SGA-PE) or not (nonSGA-PE). Case-case comparisons in
326   blue font show risk allele C is more significantly associated with LO-PE and nonSGA-PE than with EO-
327   PE and SGA-PE. Dividing cases into the four possible subcategories found that allele C confers
328   greatest risk to LO+nonSGA cases and least risk to EO+SGA cases.

329   c) Box-and-whisker plots of first and third trimester maternal serum sFlt-1 concentration in
330   preeclampsia cases and controls, showing the effect of fetal rs4769613 genotype. Boxes span the
331   first to the third quartile of sFlt-1 concentration, with horizontal bars within the box denoting the
332   median; whiskers extend to the 10th and 90th centiles. Maternal sFlt-1 is higher in cases than controls
333   in the third trimester across all fetal genotypes (t-test ln(sFlt-1): t=7.79; 200 d.f.; 2-tailed P< 0.001).
334   In third trimester controls, each copy of the rs4769613[C] allele carried by the fetus is associated
335   with an increase in maternal sFlt-1 (linear regression of ln(sFlt) with SNP genotype (coded 0, 1 and 2)
336   and gestational age as covariates: P=0.04).

337

338
339
340
341

342 **Table 1 Meta-analysis results at three independent variants near *FLT1* giving evidence for association with**
343 **preeclampsia**

Variant	Chr13 Position	Risk/Alt Allele	RAF	Covariate	GWAS N=2,658 / 308,292			Replication N=1,722 / 1,946			GWAS+Replication N=4,380 / 310,238		
					OR	95%CI	*P*	OR	95%CI	*P*	OR	95%CI	*P*
rs4769613	**29138609**	**C/T**	**0.53**	**None**	**1.22**	**1.14-1.31**	**3.2×10⁻⁸**	**1.18**	**1.08-1.30**	**3.6×10⁻⁴**	**1.21**	**1.14-1.28**	**5.4×10⁻¹¹**
rs12050029	29227519	G/A	0.14	None	1.20	1.09-1.33	1.5×10⁻⁴	1.18	1.05-1.32	5.9×10⁻³	1.19	1.11-1.28	3.0×10⁻⁶
rs149427560	29105870	G/GGT	0.06	None	1.30	1.14-1.49	9.3×10⁻⁵	1.16	0.96-1.40	1.1×10⁻¹	1.23	1.09-1.38	4.1×10⁻⁵
rs12050029	29227519	G/A	0.14	rs4769613	1.19	1.09-1.32	2.9×10⁻⁴	1.18	1.05-1.33	4.2×10⁻³	1.19	1.11-1.28	3.9×10⁻⁶
rs149427560	29105870	G/GGT	0.06	rs4769613	1.31	1.15-1.50	6.7×10⁻⁵	1.17	0.97-1.41	9.1×10⁻²	1.23	1.10-1.38	2.4×10⁻⁵

344
345 Results are ordered by strength of association with preeclampsia in GWAS+Replication meta-analysis at three

346 variants near *FLT1* not in linkage disequilibrium. Row for rs4769613 is bold because its GWAS+Replication *P* value is

347 genome-wide significant (p<5×10⁻⁸). rs12050029 and rs149427560 are included in the table because their

348 GWAS+Replication *P* values are below *FLT1* region-wide significance threshold (6.01×10⁻⁵) calculated by the method

349 of Gao (see Main Text and Methods). "N", total cases / controls in meta-analysis; "Chr13 Position", NCBI Build 37

350 position on chromosome 13; "Risk/Alt Allele", allele with higher frequency in cases than controls and alternate allele;

351 "RAF", risk allele frequency in UK GWAS controls; "Covariate", covariate in conditional logistic regression; "OR" and

352 "95%CI", allelic odds ratio and 95% confidence interval; "*P*", *P* values of case-control association. Genotypes of

353 rs149427560 for the FINNPEC cohort were proxied from rs11619261 (pairwise r^2=0.88 in Finland 1000Genomes

354 phase 3).

355
356

Methods

357

Cohorts

358

Three European cohorts of offspring from pregnancies affected by preeclampsia provided cases for

359

GWAS meta-analysis: the GOPEC and ALSPAC cohorts from the UK, and the Icelandic deCODE cohort.

360

Two independent cohorts were used for replication genotyping: the Finnish FINNPEC collection, and

361

the Norwegian Mother and Child Cohort Study (MoBa). Recruitment criteria were not identical in all

362

cohorts, so subsets were selected for this study based on an internationally recognised definition of

363

preeclampsia[14]: new-onset hypertension after the 20th week of gestation, with systolic blood

364

pressure ≥140mmHg or diastolic blood pressure ≥90mmHg on at least two occasions; and new-onset

365

proteinuria of 0.3g/24 hours or more, or ≥1+ on dipstick analysis of urine. All were singleton

366

pregnancies in a white Western European woman; preeclamptic pregnancies in women with a

367

previous history of essential hypertension, type 1 or type 2 diabetes mellitus, ischaemic heart

368

disease, cerebrovascular accident or chronic renal disease were excluded. Phenotypic details are

369

summarised in Supplementary Tables 7 and 8.

370

371

Written informed consent was obtained from participants, or from parents on behalf of minors, and

372

all studies were approved by local Research Ethics Committees.

373

GOPEC (Genetics of Pre-eclampsia)

374

The UK GOPEC collection includes 1157 DNA samples from mother-baby pairs with preeclampsia

375

recruited at diagnosis between 1992 and 2009 for genetic studies of preeclampsia[15]. Control data

376

were derived from the WTCCC2 genome-wide analysis of 2930 samples from the 1958 Birth

377

Cohort and 2737 samples from the National Blood Services, providing control data for 5297

378

individuals after QC[16].

379

380

ALSPAC (Avon Longitudinal Study of Parents and Children)

381

The ALSPAC prospective birth cohort study recruited pregnant women living in the South West of

382

England between 1991 and 1992, and has been described elsewhere[17]. They included 13,678

383

singleton pregnancies resulting in a live birth. After exclusion of women with existing hypertension,

384

diabetes and gestational diabetes, there were 7382 pregnancies for which blood pressure and

385

proteinuria measurements and fetal GWAS data were available. Of these, 146 met the definition of

386

preeclampsia and were included. The control group of 6130 subjects was derived from all other

387

included pregnancies with fetal GWAS data in women not affected by essential hypertension or

388

gestational hypertension.

389

390

deCODE Pre-eclampsia cohort

391

The deCODE preeclampsia cohort is part of an ongoing sample collection including a large part of the

392

Icelandic population. Preeclamptic pregnancies occurring between 1970 and 2009 were identified

393

through scrutiny of hospital records at the Landspitali University Hospital, which provides secondary

394

and tertiary services for the whole of Iceland. Initially, a group of women with hypertensive disease

395

in pregnancy (ICD-9:642.0–9 and ICD-10: O10–16) in the years 1984-1999 were selected for further

396

study based on familial relationships. All maternity records for these women were scrutinised and

397

each affected pregnancy reclassified[18]. This identified 491 singleton preeclamptic pregnancies.

398

Preeclamptic pregnancies from 2000-2009 were identified based on ICD-10 codes (O14-15), yielding

399

1,311 additional singleton preeclamptic pregnancies. Overall information on 1,802 singleton

400

401   preeclamptic pregnancies of 1,662 mothers is available. GWAS data was available from 1507
402   offspring of these pregnancies, identified retrospectively based on the national register. The control
403   group comprises 296,865 individuals from the deCODE sample collection.
404

405   *MoBa (Norwegian Mother and Child Cohort Study)*
406   The Norwegian Mother and Child Cohort Study is a longitudinal study of over 110,000 pregnant
407   women, their children and partners, recruited between 1999 and 2008 from maternity units
408   throughout Norway[19]. 1200 pregnancies affected by preeclampsia were identified from Medical
409   Birth Register of Norway records; the validity of the diagnosis has been assessed by retrieval and
410   examination of antenatal records. 1200 non-hypertensive pregnancies provided the control group.
411   Pregnancies were excluded from case and control groups if a maternal history of essential
412   hypertension, chronic renal disease or diabetes mellitus was recorded in the Medical Birth Registry
413   of Norway.
414

415   *FINNPEC (Finnish Genetic of Pre-eclampsia Consortium)*
416   The FINNPEC collection was assembled in Finland between 2008 and 2011 from two recruitment
417   arms[20]. Samples were collected at the time of diagnosis of preeclampsia from 879 mothers, and
418   during pregnancy from 922 non-pre-eclamptic mothers from antenatal and labour wards. Their
419   children and partners were also enrolled. A further 525 pregnancies affected by preeclampsia were
420   identified by examination of hospital records, and women and offspring were invited to participate
421   by letter. After exclusion of pregnancies which did not meet the entry criteria for this study,
422   offspring of 605 preeclamptic pregnancies were included as cases, and offspring of 800 non-
423   hypertensive pregnancies provided the control group.
424

425   **Genotyping, quality control, genotype imputation and association analysis in GWAS datasets**

426   *GOPEC*
427   1157 offspring of preeclamptic pregnancies were assayed on the Illumina OmniExpress chip;
428   maternal samples where available were similarly genotyped. A total of 730,525 variants were called
429   with the GenCall algorithm. We carried out QC using PLINK
430   (http://pngu.mgh.harvard.edu/~purcell/anal.shtml) and SMARTPCA[21]. A subset of the samples (186)
431   were whole genome amplified (WGA). WGA genotype calls can be prone to calling artefacts. To
432   address this we removed variants with either low call rate (95%) or Mendelian errors in the WGA
433   samples and we then performed a pseudo-case control analysis of WGA vs. non-WGA and removed
434   variants with significant genotypic association ($P<0.001$). We then applied standard QC to the
435   combined WGA and non-WGA dataset on this reduced set of variants (670,435). Briefly, standard
436   QC comprises the following subject level exclusion criteria: individual call-rate < 95%; heterozygosity
437   >3 s.d. from the mean; any of the first 3 HapMap (based on CEU, YRI, CHB, JPT and GIH) principal
438   axes of variation >4 s.d. from the mean and gender mismatch. Related individuals (IBD>0.1) with
439   lowest call-rate were preferentially removed. The variant level exclusion criteria are: call-rate <95%,
440   exact Hardy-Weinberg equilibrium $P<1 \times 10^{-6}$, minor allele frequency (MAF) <1% and non-random
441   missingness of uncalled genotypes ("plink --test-mishap") with Bonferroni corrected $P<0.05$. These
442   filters left 1005 samples (89 WGA) and 574,919 variants.
443

444   We used WTCCC2 population controls from the National Blood Donors Cohort and UK 1958 Birth
445   Cohort[16]. These samples were genotyped on the Illumina 1.2M chip and called using GenCall.
446   Strand ambiguous markers were removed and the standard QC described above was then applied to
447   the two control datasets. The merged control datasets consisted of 5,297 samples and 438,912
448   variants. This control dataset was merged with the case dataset resulting in 429,754 post QC variants
449   that were genotyped in both cases and controls.

450

451   Cases and controls were imputed together with IMPUTE2 (impute_v2.3.0)[22] and SHAPEIT[23] using the
452   pre-phasing workflow against the 1000 Genomes Phase 1 reference panel (Dec. 2013) downloaded
453   from the IMPUTE2 website. Imputation resulted in 10,404,388 bi-allelic variants with MAF > 0.25%
454   that were either directly genotyped or imputed with IMPUTE2 INFO score >0.6.

455

456   Post imputation association analysis was carried out using SNPTEST (v2.4.1)[22] with the "expected"
457   method with no ancestry principal components. We calculated the genomic control on variants with
458   MAF>0.5% as λ_{GC} =1.005.

459

460   *deCODE*
461   Details of GWAS genotyping, QC and imputation of the Icelandic dataset including the preeclampsia
462   cases and controls used in this study have been described[24]. Briefly, samples were assayed with the
463   Illumina HumanHap300, HumanCNV370, HumanHap610, HumanHap1M, HumanHap660, Omni-1,
464   Omni 2.5 or Omni Express bead chips at deCODE genetics. Following QC a final set of 676,913
465   autosomal SNPs were used for long range phasing of all chip-genotyped samples. Making use of the
466   Icelandic genealogy untyped first and second degree relatives of chip-typed individuals were also
467   included in the analysis to increase power[24]. In total 104,220 chip-typed individuals and 294,212 of
468   their untyped relatives were imputed based on a panel of sequence variants identified through
469   whole genome sequencing of 2,636 Icelanders to a mean depth of 20x.

470

471   GWAS analysis of Icelandic preeclampsia offspring included a total of 1,507 cases (380 chip typed)
472   and 296,865 controls (91,326 chip-typed). The controls used in this study were Icelandic individuals
473   from other ongoing GWAS studies at deCODE and their relatives. Logistic regression was used to test
474   for association between sequence variants and disease, treating disease status as the response and
475   genotype counts as covariates. Other characteristics also included in the model as nuisance variables
476   were: sex, county of birth, current age or age at death (first and second order terms included),
477   genotyping status and an indicator function for the overlap of the lifetime of the individual with the
478   timespan of phenotype collection[25]. In order to account for relatedness and stratification within the
479   case and control sample sets we applied the method of genomic control[25]. Based on a set of about
480   300,000 common variants distributed across the genome the inflation in the chi-squared statistic for
481   preeclampsia offspring was estimated to be 1.115.

482

483   *ALSPAC*
484   A total of 9,912 ALSPAC children were genotyped using the Illumina HumanHap550 quad genome-
485   wide SNP genotyping platform (Illumina Inc., San Diego, CA, USA) by Logistics and Genotyping
486   Facilities at the Wellcome Trust Sanger Institute and Laboratory Corporation of America (LabCorp
487   Holdings., Burlington, NC, USA). PLINK software (v1.07) was used to carry out quality control
488   measures. Individuals were excluded from further analysis on the basis of having incorrect gender

489   assignments, minimal or excessive heterozygosity (< 0.320 and > 0.345 for the Sanger data and <

490   0.310 and > 0.330 for the LabCorp data), disproportionate levels of individual missingness (> 3%) and

491   being of non-European ancestry (as detected by a multidimensional scaling analysis seeded with

492   HapMap 2 individuals). EIGENSTRAT analysis revealed no additional obvious population stratification

493   and genome-wide analyses with other phenotypes in the same cohort indicate a low lambda. SNPs

494   with a minor allele frequency of < 1% and call rate of < 95% were removed. Furthermore, only SNPs

495   that passed an exact test of Hardy–Weinberg equilibrium ($P > 5 \times 10^{-7}$) were considered for analysis.

496   Related subjects (> 10% IBD) that passed all other quality control thresholds were retained during

497   subsequent phasing and imputation. 9,115 subjects and 500,527 SNPs passed these quality control

498   filters.

499   We combined 477,482 SNP genotypes in common between the sample of children and mothers.

500   Genotyping and QC of the ALSPAC mothers can be found elsewhere[26]. SNPs with genotype

501   missingness > 1% and those that failed the exact test of HWE were removed. A further 321

502   participants were removed due to potential ID mismatches (IBD < 1). The resultant dataset

503   comprised 17,842 subjects of which 6,305 were mother-offspring pairs. An additional 112 SNPs were

504   removed after a liftover of the merged genotyped data from Hg 18 to Hg19. Haplotype phasing was

505   performed using SHAPEIT (v2.r644)[23] and known autosomal variants were imputed with IMPUTE

506   V2.2.2[22] using the 1000 genomes reference panel (Phase 1, Version 3) consisting of 2186 reference

507   haplotypes (including non-Europeans).

508

509   Logistic regression, as implemented in SNPTEST v2.5-beta4[22], was used to test for association

510   between imputed genotype probabilities and disease status. Based on a set of ~9 million SNPs

511   (MAF>0.5% and IMPUTE2 INFO>0.6), no evidence of genomic inflation was observed (λ_{GC}=1.008).

512

513   **Follow-up genotyping**

514   Follow-up variants in the *FLT1* locus were chosen to test association with the rs4769613 peak and to

515   test possible association in the other 8 *FLT1* LD blocks shown in Supplementary Fig. 2. Some

516   genotyped variants were highly correlated surrogates of rs4769613 or of other follow-up SNPs in

517   case the assay for the primary variant failed, and to ensure that assertion of a true-positive

518   association did not rely on genotyping of a single variant. Follow-up variants in non-*FLT1* regions of

519   the genome were chosen based on GWAS meta-analysis *P* value and were selected to further test

520   suggestive evidence of association exhibited by the top GWAS discovery meta-analysis signals.

521   Replication genotyping was performed at the Wellcome Trust Sanger Institute using Sequenom iPLEX

522   assays, and at the British Heart Foundation Glasgow Cardiovascular Research Centre using TaqMan

523   Open Array genotyping. Variants were excluded from analysis if they had call rates < 95%; subjects

524   with call rates < 80%, and families in the MoBa cohort that exhibited more than 1 Mendelian error

525   were also excluded. For four variants (rs7305125, rs149427560, rs12050029 and rs4769628) follow-

526   up data for the MoBa samples was *in silico* data based on 1046 cases and 961 controls assayed on

527   the Illumina HumanCoreExome-12 v1.1 chip and imputed based on the 1000 Genomes Phase 3

528   reference panel. Of those, 908 cases and 909 controls were also included in the directly genotyped

529   MoBa replication set.

530

531   **Meta-analysis**

532   Prior to meta-analysis GWAS results were adjusted by a genomic control λ_{GC} factor where

533   appropriate as described above for each GWAS cohort. Study level variants with a MAF<0.5% or an

534   imputation quality score <0.6 were excluded from the analysis. This left 7,476,169 autosomal

535   variants for analysis. The GWAS and the GWAS+Replication meta-analyses were conducted using the

536   fixed effect inverse variance weighting method implemented in MetaSoft[27]. No genomic control

537   adjustment was applied to the GWAS meta-analysis results since the inflation factor was negligible

538   (λ_{GC} =1.0075).

539   **Conditional and *FLT1* region-wide analyses**

540   The association between disease status on a variant conditional on rs4769613 was assessed by

541   inverse variance weighted meta-analysis of the per cohort conditional analyses. Individual cohorts

542   were analysed by logistic regression of the disease status against expected genotype dose with the

543   expected doses of conditioning variants included as covariates. This approach was implemented for

544   each cohort as follows: MoBa and FINNPEC replication cohorts were analysed using "plink --

545   condition"; GOPEC and MoBa GWAS were analysed using "snptest -condition_on"; the deCODE

546   association analysis is described above; ALSPAC conditional associations were inferred from the

547   summary association statistics with the use of the 1000 Genomes Phase 3 EUR samples to estimate

548   the LD structure using the joint analysis method[28]. We assessed the region-wide effective number of

549   tests using the method of Gao[29] on the imputed WTCCC2 UK control dataset for the 5405 common

550   variants (MAF>5%) within 1Mb of rs4769613, yielding a total of 832 independent tests and hence a

551   *FLT1* region-wide significance threshold of 0.05/832=6.01×10^{-5}.

552   **Maternal, fetal and parent-of-origin effect analysis**

553   The family genotype data was jointly analysed using the EMIM method[6].The subjects were first

554   partitioned into maximal family groups within each cohort (Supplementary Table 9). We then fitted

555   EMIM models assuming Hardy-Weinberg equilibrium and Exchangeable Parental Genotypes. We

556   considered two parameter sets: maternal and fetal effects (R_1, R_2, S_1 and S_2) and maternal, fetal and

557   parent of origin effect (R_1, R_2, S_1, S_2 and I_m) where I_m is the odds-ratio associated with the maternal

558   transmission of the risk allele. The per cohort results were combined using inverse variance

559   weighted meta-analysis.

560   **Preeclampsia subtype analysis**

561   To analyse the relation between rs4769613 and preeclampsia subtypes we pooled genotype and

562   clinical data of the GOPEC, FINNPEC and MoBa cohorts (Supplementary Table 8). The phenotype

563   associations were calculated using logistic regression with cohort and FINNPEC recruitment region

564   included as indicator variables.

565   **deCODE phenotype database**

566   The deCODE Genetics phenotype database contains medical information on diseases and traits

567   obtained through collaboration with specialists in each field. This includes information on

568   cardiovascular diseases (myocardial infarction, coronary arterial disease, peripheral arterial disease,

569   atrial fibrillation, sick sinus syndrome and stroke), metabolic disorders (obesity, diabetes, and

570   metabolic syndrome), psychiatric disorders (schizophrenia, bipolar disorder, anxiety and depression),

571   addictions (nicotine, alcohol), inflammatory diseases (rheumatoid arthritis, lupus, and asthma),

572   musculoskeletal disorders (osteoarthritis, osteoporosis), eye diseases (glaucoma), kidney diseases

573   (kidney stones, kidney failure) and 29 types of cancer. Anthropometric measures have also been

574 collected through several of these projects. Routinely measured traits from patient workups
575 (sodium, potassium, bicarbonate, calcium, phosphate, creatinine, blood cell counts, haemoglobin,
576 haematocrit, 15 immunoglobulins, iron, vitamins, lipids, liver function tests and more) were obtained
577 from the Landspitali University Hospital, Reykjavík, and the Icelandic Medical Center Laboratory in
578 Mjodd (Laeknasetrid), Reykjavik. The number of independent and uncorrelated secondary traits
579 tested for association amounts to 400.
580

581 **Placental expression of Flt1 and sFlt1**
582 Women with singleton pregnancies delivering by caesarean section were recruited to the Pre-
583 eclampsia Study between 2002 and 2012 at St. Olavs Hospital, Trondheim University Hospital and
584 Haukeland University Hospital, Bergen. Healthy and pre-eclamptic pregnancies were included as
585 described previously[30]. A tangential section (100 mg) from the maternal central side of the placenta
586 was collected directly after delivery, fixed in 10% neutral-buffered formalin and paraffin embedded.
587 Tissue sections of 3 μm were pre-treated in Target Retrieval Solution (#K8004, Dako) and stained by
588 Flt-1 antibody (1:175, # ab32152, Abcam) using EnVision (#K4011, Dako) according to the
589 manufacturer's protocol. This Flt-1 antibody recognises membrane-bound Flt-1 and splice isoforms
590 sFlt-1, sFlt1-14, and isoform 4 (61 kDa). Staining was performed using Autostainer Plus (#S3800,
591 Dako) and images taken at two sites per placenta with an Eclipse E400 microscope and DS-Fi1
592 camera. Staining intensity in syncytiotrophoblast was analysed by NIS-Elements BR 4.0 software
593 (Nikon), excluding immature villi, and blinded for pregnancy outcomes. Staining intensity data were
594 analysed separately in a general linear model incorporating SNP genotype, gestational age and
595 hospital of origin as cofactors.
596

597 **Maternal serum sFlt1 and fetal genotype**
598 We identified mother-baby pairs from the FINNPEC collection for whom offspring DNA and maternal
599 serum samples from the first and/or third trimester of pregnancy were available for analysis.
600 Maternal serum sFlt-1 concentration was measured using electrochemiluminescence immunoassays
601 (ECLIA; Roche Diagnostics GmbH, Mannheim, Germany) on a Cobas e 601 analyzer (Hitachi High
602 Technology Co, Tokyo, Japan). Offspring genotype at rs4769613 was determined by Sequenom
603 MassArray iPLEX genotyping in the FiMM Technology Centre (University of Helsinki, Finland).
604 Investigators were blinded for pregnancy outcome during sample analysis. Serum sFlt-1 data were
605 normalized by logarithmic transformation. Case and control data were compared by unpaired t-
606 testing; genotypic associations with serum sFlt-1 were examined separately in cases and controls in a
607 linear model, with SNP genotype and gestational age as covariates.
608

609 **Data Availability**

610 Meta-analysed GWAS data used in this study, and individual-level GWAS data from the
611 GOPEC cohort, are deposited in the European Genome-phenome Archive
612 (www.ebi.ac.uk/ega) with accession numbers EGAD00010001211 and EGAD00010001212.
613

614 **Methods references**

615 14. Brown, M. A., Lindheimer, M. D., de Swiet, M., Van Assche, A. & Moutquin, J. M. The
616   classification and diagnosis of the hypertensive disorders of pregnancy: statement from the

617   International Society for the Study of Hypertension in Pregnancy (ISSHP). *Hypertens. Pregnancy*
618   **20**, IX–XIV (2001).

619 15. The GOPEC Consortium. Disentangling fetal and maternal susceptibility for pre-eclampsia: a
620   British multicenter candidate-gene study. *Am. J. Hum. Genet.* **77**, 127-131 (2005).

621 16. Evans, D. M. *et al.* Interaction between *ERAP1* and HLA-B27 in ankylosing spondylitis implicates
622   peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nat. Genet.* **43,** 761–
623   767 (2011).

624 17. Boyd, A. *et al.* Cohort Profile: the 'children of the 90s'; the index offspring of The Avon
625   Longitudinal Study of Parents and Children. *Int. J. Epidemiol.* **42,** 111-127 (2013).

626 18. Hjartardottir, S., Leifsson, B. G., Geirsson, R. T. & Steinthorsdottir, V. Paternity change and the
627   recurrence risk in familial hypertensive disorder in pregnancy. *Hypertens. Pregnancy*. **23**, 219-
628   225 (2004).

629 19. Magnus, P. *et al.* Cohort Profile Update: The Norwegian Mother and Child Cohort Study (MoBa).
630   *Int. J. Epidemiol.* **45**, 382-388 (2016).

631 20. Jääskeläinen, T. *et al.* Cohort profile: the Finnish Genetics of Pre-eclampsia Consortium
632   (FINNPEC). *BMJ Open.* Nov 10;6(11):e013148. doi:10.1136/bmjopen-2016-013148 (2016).

633 21. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet.*
634   Dec;2(12):e190. doi:10.1371/journal.pgen.0020190 (2006).

635 22. Marchini, J. & Howie, B. Genotype imputation for genome-wide association studies. *Nat. Rev.*
636   *Genet.* **11**, 499-511 (2010).

637 23. O'Connell, J. *et al.* (2014) A general approach for haplotype phasing across the full spectrum of
638   relatedness. *PLoS Genet*. Apr 17:10(4):e1004234. doi: 10.1371/journal.pgen.1004234 (2014).

639 24. Gudbjartsson, D. F. *et al*. Large-scale whole-genome sequencing of the Icelandic population. *Nat.*
640   *Genet.* **47**, 435-444 (2015).

641 25. Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55,** 997-1004 (1999).

642 26. Hellmich, C. *et al*. Genetics, sleep and memory: a recall-by-genotype study of ZNF804A variants
643   and sleep neurophysiology. *BMC Med. Genet.* Oct 24; 16:96. doi: 10.1186/s12881-015-0244-4
644   (2015).

645 27. Han, B. & Eskin, E. Genetics, sleep and memory: a recall-by-genotype study of ZNF804A variants.
646   *Am. J. Hum. Genet.* **88,** 586-598 (2011).

647 28. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies
648   additional variants influencing complex traits. *Nat. Genet.* **44,** 369-375 (2012).

649 29. Hendricks, A. E. *et al.* Correction for multiple testing in a gene region. *Eur. J Hum. Genet.* **22,**
650   414-418 (2014).

651 30. Austdal, M. *et al*. Metabolic profiles of placenta in preeclampsia using HR-MAS MRS
652   metabolomics. *Placenta* **36**, 1455-1462 (2015).

a
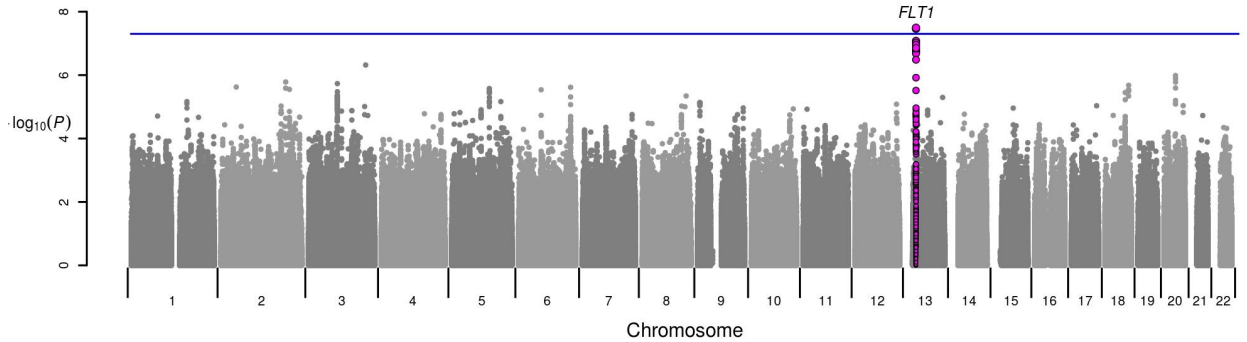
b
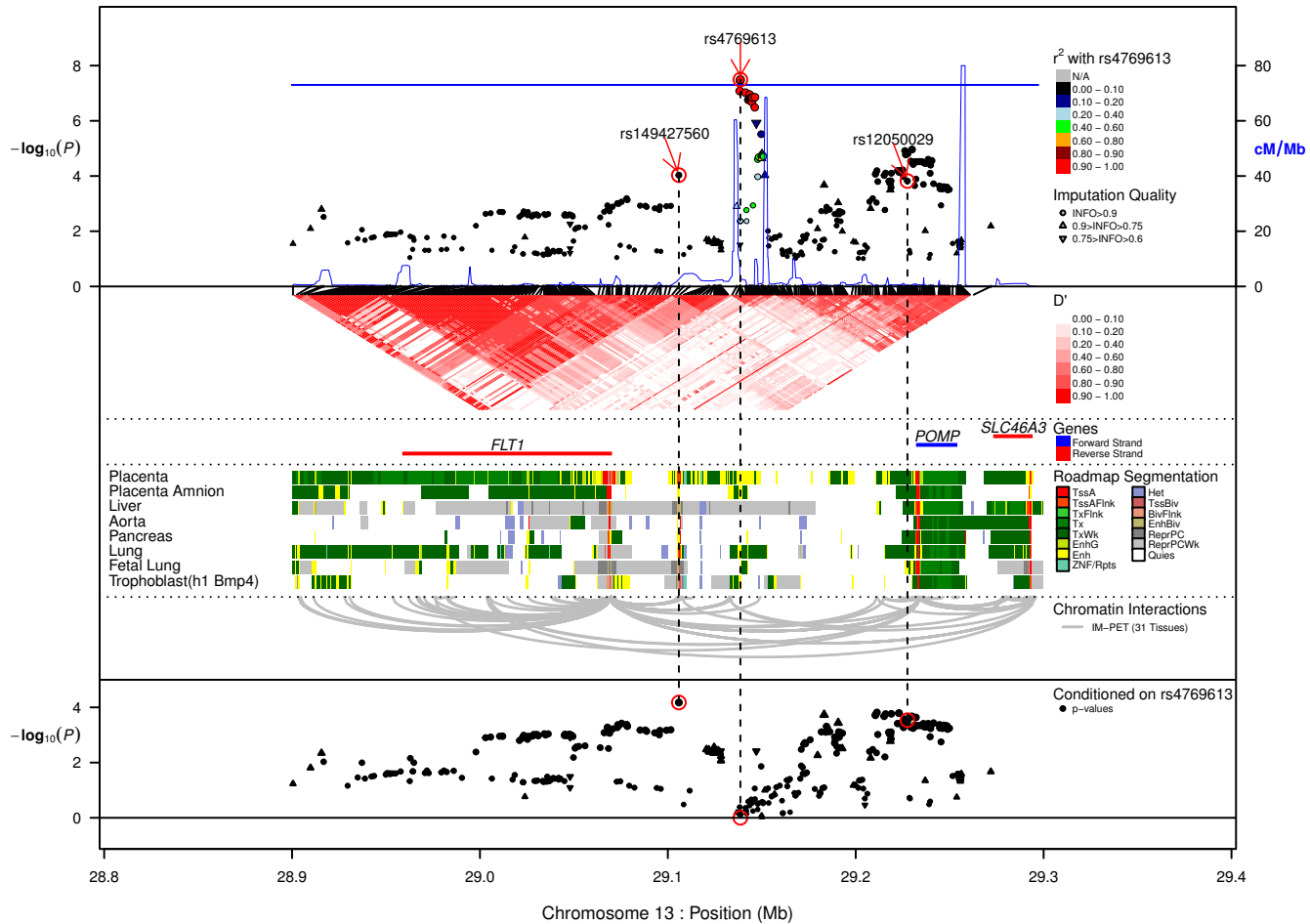
a

rs4769613: Jointly Fitted Maternal (S_1 & S_2) and Fetal(R_1 & R_2) Effects

Cohort	P	
GOPEC	2.3×10^{-3}	R_1
	9.4×10^{-5}	R_2
	5.5×10^{-1}	S_1
	7.4×10^{-1}	S_2
MoBa	5.1×10^{-1}	R_1
	1.2×10^{-2}	R_2
	2.5×10^{-1}	S_1
	2.1×10^{-1}	S_2
FINNPEC	3.4×10^{-3}	R_1
	1.9×10^{-5}	R_2
	7.3×10^{-1}	S_1
	1.7×10^{-1}	S_2
[Meta]	$\mathbf{3.7 \times 10^{-4}}$	R_1
	$\mathbf{3.7 \times 10^{-9}}$	R_2
	$\mathbf{3.5 \times 10^{-1}}$	S_1
	$\mathbf{2.8 \times 10^{-1}}$	S_2

Odds Ratio (95% CI)
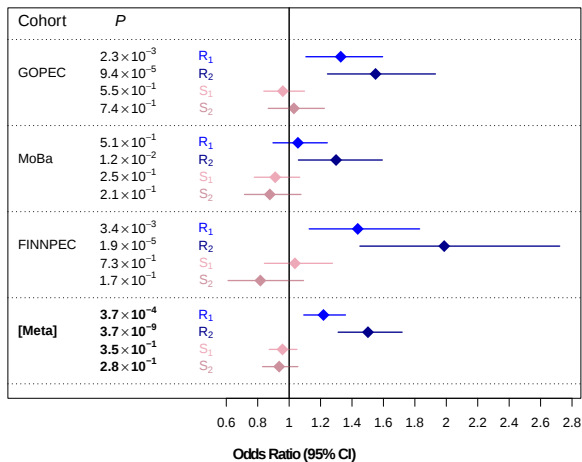
b

rs4769613 : Preeclampsia Subtype Meta−analysis Overview

Analysis	Cases	Controls	P	OR
LO−PE vs Controls	2199	7243	8.1×10^{-8}	1.23
EO−PE vs Controls	455	7243	4.0×10^{-1}	1.06
LO−PE vs EO−PE	**2199**	**455**	$\mathbf{1.7 \times 10^{-2}}$	**1.21**
nonSGA−PE vs Controls	1756	7243	8.9×10^{-8}	1.25
SGA−PE vs Controls	936	7243	6.0×10^{-2}	1.10
nonSGA−PE vs SGA−PE	**1756**	**936**	$\mathbf{1.9 \times 10^{-2}}$	**1.15**
LO+nonSGA PE vs Controls	1611	7243	1.2×10^{-7}	1.26
LO+SGA PE vs Controls	588	7243	4.1×10^{-2}	1.14
EO+nonSGA PE vs Controls	122	7243	2.8×10^{-1}	1.15
EO+SGA PE vs Controls	333	7243	7.2×10^{-1}	1.03
LO+nonSGA vs EO+SGA	**1611**	**333**	$\mathbf{5.8 \times 10^{-3}}$	**1.31**

Odds Ratio (95% CI)

c

Fetal rs4769613 genotype
- TT
- TC
- CC

$P < 0.001$

$P = 0.04$

Maternal serum sFlt-1 (pg/mL)

	Controls	Cases	Controls	Cases
n	223	195	47	156

First trimester   Third trimester

Supplementary Table 1 Meta-analysis of variants selected for post-GWAS follow-up

Variant	Nearest Gene	Chr	Position	Risk/Alt Allele	RAF	GWAS N=2,658 / 308,292			Replication N=1,722 / 1,946			GWAS+Replication N=4,380 / 310,238		
						OR	95%CI	P	OR	95%CI	P	OR	95%CI	P
rs16835173	PCGEM1	2	194213678	T/C	0.02	1.61	1.33-1.96	1.8×10^{-6}	1.26	0.88-1.82	0.20	1.53	1.28-1.81	1.5×10^{-6}
rs78689834	PARD3B	2	205388423	A/C	0.02	1.58	1.31-1.92	3.0×10^{-6}	0.88	0.65-1.19	0.40	1.34	1.14-1.57	4.9×10^{-4}
rs73843515	CADM2	3	85984083	T/C	0.01	2.01	1.48-2.74	8.8×10^{-6}	0.78	0.40-1.49	0.45	1.69	1.28-2.23	2.2×10^{-4}
rs9846396	ZBTB38	3	141140968	C/T	0.56	1.14	1.06-1.22	3.6×10^{-4}	1.00	0.86-1.17	0.98	1.10	1.04-1.17	2.1×10^{-3}
rs112342656	MCC	5	112656059	C/T	0.01	1.90	1.45-2.48	2.7×10^{-6}	0.71	0.46-1.10	0.12	1.45	1.15-1.82	1.5×10^{-3}
rs11740989	JAKMIP2	5	147160125	C/T	0.05	1.37	1.20-1.58	7.1×10^{-6}	1.05	0.84-1.32	0.67	1.28	1.14-1.44	5.0×10^{-5}
rs16898533	SLC25A51P1	6	67269926	G/T	0.04	1.46	1.25-1.71	3.0×10^{-6}	0.87	0.64-1.17	0.35	1.30	1.13-1.50	2.3×10^{-4}
rs9397792	TIAM2	6	155543913	G/T	0.14	1.26	1.14-1.38	2.5×10^{-6}	1.06	0.90-1.25	0.47	1.20	1.11-1.31	9.2×10^{-6}
rs12154986	FAM220A	7	6398651	A/G	0.72	1.15	1.06-1.25	5.4×10^{-4}	1.11	1.01-1.23	0.04	1.14	1.07-1.21	5.8×10^{-5}
rs6577900	FAM135B	8	139255733	C/T	0.32	1.15	1.07-1.24	2.1×10^{-4}	1.02	0.90-1.16	0.70	1.12	1.05-1.19	7.4×10^{-4}
rs2427981	SARDH	9	136581419	A/C	0.58	1.18	1.09-1.27	1.1×10^{-5}	0.95	0.85-1.08	0.45	1.14	1.07-1.21	3.5×10^{-5}
rs3750804	TCF7L2	10	114833850	C/T	0.62	1.19	1.10-1.29	2.9×10^{-5}	1.08	0.95-1.24	0.22	1.16	1.08-1.24	2.6×10^{-5}
rs11197042	ATRNL1	10	116825666	A/G	0.29	1.18	1.09-1.27	2.7×10^{-5}	0.93	0.82-1.05	0.26	1.10	1.03-1.18	2.8×10^{-3}
rs1586382	ZBED5-AS1	11	11074260	T/G	0.73	1.15	1.06-1.24	1.0×10^{-5}	1.02	0.90-1.17	0.72	1.10	1.02-1.17	9.4×10^{-3}
rs11227306	OVOL1	11	65578672	A/C	0.38	1.16	1.08-1.25	4.4×10^{-5}	1.06	0.96-1.17	0.28	1.12	1.06-1.19	7.7×10^{-5}
rs501630	EFEMP2	11	65637273	A/G	0.45	1.15	1.07-1.23	1.1×10^{-4}	1.04	0.92-1.16	0.54	1.12	1.05-1.19	3.0×10^{-4}
rs7305125	ITPR2	12	26960599	C/G	0.31	1.19	1.10-1.29	6.8×10^{-6}	1.06	0.92-1.22	0.88	1.16	1.09-1.24	1.3×10^{-5}
rs118009336	DNAH10	12	124410135	A/C	0.96	1.56	1.28-1.90	8.6×10^{-6}	0.98	0.75-1.26	0.85	1.31	1.12-1.53	6.4×10^{-4}
rs4769613	***FLT1***	**13**	**29138609**	**C/T**	**0.53**	**1.22**	**1.14-1.31**	$\mathbf{3.2\times10^{-8}}$	**1.18**	**1.08-1.30**	$\mathbf{3.6\times10^{-4}}$	**1.21**	**1.14-1.28**	$\mathbf{5.4\times10^{-11}}$
rs7328374	***FLT1***	**13**	**29141327**	**T/C**	**0.52**	**1.21**	**1.13-1.30**	$\mathbf{9.6\,10^{-8}}$	**1.19**	**1.09-1.31**	$\mathbf{2.1\times10^{-4}}$	**1.20**	**1.14-1.27**	$\mathbf{8.6\times10^{-11}}$
rs11623923	SLC35F4	14	58042753	G/A	0.84	1.18	1.07-1.31	1.1×10^{-3}	1.04	0.88-1.22	0.65	1.14	1.05-1.24	2.7×10^{-3}
rs72747221	GCOM1	15	58058582	C/T	0.96	1.74	1.36-2.24	1.1×10^{-5}	0.89	0.65-1.21	0.45	1.33	1.10-1.62	3.4×10^{-3}
rs12962662	TNFRSF11A	18	60104589	T/C	0.16	1.25	1.14-1.38	3.7×10^{-6}	1.03	0.87-1.22	0.72	1.19	1.10-1.29	2.7×10^{-5}
rs6566644	CBLN2	18	70310180	G/T	0.54	1.18	1.10-1.27	3.0×10^{-6}	1.10	0.98-1.23	0.12	1.16	1.09-1.23	1.5×10^{-6}
rs56090944	MMP24	20	33813993	A/G	0.04	1.44	1.24-1.66	1.4×10^{-6}	1.06	0.82-1.38	0.64	1.34	1.18-1.52	9.2×10^{-6}
rs2071969	L3MBTL1	20	42160962	G/A	0.90	1.25	1.11-1.41	2.4×10^{-4}	0.88	0.73-1.05	0.16	1.12	1.02-1.24	2.1×10^{-2}
rs1412977	C20orf85	20	56709290	G/T	0.51	1.18	1.09-1.26	9.5×10^{-6}	1.08	0.99-1.19	0.09	1.14	1.08-1.21	5.6×10^{-6}
rs73306896	ZNF831	20	57750533	T/C	0.12	1.21	1.09-1.34	2.3×10^{-4}	1.11	0.94-1.31	0.21	1.18	1.08-1.29	1.5×10^{-4}

Variants in the table are ordered by chromosome number ("Chr") and NCBI Build 37 position ("Position"). "N", total cases / controls in meta-analysis; "Risk/Alt Allele", allele with higher frequency in GWAS cases than controls and alternate allele; "RAF", risk allele frequency in UK GWAS controls; "OR" and "95% CI", allelic odds ratio and 95% confidence interval. "*P*", *P* values of case-control association. Rows for rs4769613 and its near-perfect surrogate rs7328374 are bold because their GWAS+Replication *P* values are below genome-wide significance of 5×10^{-8} No other follow-up variants achieved genome-wide significance.

Supplementary Table 2   Meta-analysis results at all variants near *FLT1* followed up in the replication datasets

Variant	Chr13 Position	Risk/Alt Allele	RAF	GWAS N=2,658 / 308,292			Replication N=1,722 / 1,946			GWAS+Replication N=4,380 / 310,238		
				OR	95%CI	*P*	OR	95%CI	*P*	OR	95%CI	*P*
rs3794401	28915659	A/G	0.87	1.20	1.07-1.35	1.6×10^{-3}	1.09	0.95-1.26	0.22	1.16	1.06-1.26	1.3×10^{-3}
rs2296284	28963676	A/G	0.29	1.01	0.94-1.09	7.9×10^{-1}	0.99	0.87-1.12	0.83	1.00	0.94-1.07	9.1×10^{-1}
rs9319429	28973703	T/C	0.30	1.01	0.94-1.09	8.0×10^{-1}	1.00	0.88-1.13	0.97	1.01	0.94-1.07	8.4×10^{-1}
rs11619261	29078077	A/G	0.07	1.25	1.10-1.42	6.4×10^{-4}	1.18	0.99-1.39	0.06	1.22	1.10-1.35	1.1×10^{-4}
rs149427560	29105870	G/GGT	0.06	1.30	1.14-1.49	9.3×10^{-5}	1.16	0.96-1.40	0.11	1.23	1.09-1.38	4.1×10^{-5}
rs4769613	**29138609**	**C/T**	**0.53**	**1.22**	**1.14-1.31**	$\mathbf{3.2\times10^{-8}}$	**1.18**	**1.08-1.30**	$\mathbf{3.6\times10^{-4}}$	**1.21**	**1.14-1.28**	$\mathbf{5.4\times10^{-11}}$
rs7328374	**29141327**	**T/C**	**0.52**	**1.21**	**1.13-1.30**	$\mathbf{9.6\times10^{-8}}$	**1.19**	**1.09-1.31**	$\mathbf{2.1\times10^{-4}}$	**1.20**	**1.14-1.27**	$\mathbf{8.6\times10^{-11}}$
rs17555115	29143824	G/A	0.52	1.21	1.13-1.30	1.6×10^{-7}	1.12	0.99-1.27	0.07	1.19	1.12-1.26	5.1×10^{-8}
rs3829387	29149693	C/A	0.21	1.22	1.12-1.32	3.1×10^{-6}	1.09	0.95-1.26	0.22	1.19	1.10-1.27	3.3×10^{-6}
rs9508065	29151651	A/C	0.22	1.19	1.09-1.30	9.4×10^{-5}	1.08	0.97-1.21	0.16	1.15	1.07-1.23	8.1×10^{-5}
rs9508079	29169711	G/A	0.54	1.06	0.99-1.14	1.1×10^{-1}	1.05	0.92-1.20	0.45	1.06	0.99-1.12	8.0×10^{-2}
rs117488563	29183168	A/G	0.06	1.32	1.14-1.53	2.1×10^{-4}	1.20	0.86-1.67	0.28	1.30	1.14-1.48	1.3×10^{-4}
rs2096035	29184460	A/G	0.81	1.17	1.06-1.28	1.3×10^{-3}	1.05	0.90-1.22	0.56	1.13	1.04-1.23	2.4×10^{-3}
rs4769620	29184950	C/T	0.81	1.16	1.06-1.28	1.5×10^{-3}	1.03	0.91-1.17	0.64	1.11	1.03-1.20	4.7×10^{-3}
rs9508092	29186162	T/C	0.80	1.16	1.06-1.27	1.7×10^{-3}	1.07	0.94-1.21	0.34	1.13	1.05-1.22	1.8×10^{-3}
rs35242283	29211289	A/AGAT	0.14	1.22	1.11-1.34	6.4×10^{-5}	1.25	1.07-1.47	6.4×10^{-3}	1.23	1.13-1.33	1.4×10^{-6}
rs71433277	29218967	T/C	0.14	1.22	1.10-1.34	7.9×10^{-5}	1.14	1.02-1.29	0.03	1.19	1.10-1.28	7.6×10^{-6}
rs1185049	29226634	G/A	0.34	1.17	1.09-1.26	1.7×10^{-5}	1.09	0.99-1.20	0.06	1.14	1.08-1.21	5.6×10^{-6}
rs12050029	29227519	G/A	0.14	1.20	1.09-1.33	1.5×10^{-4}	1.18	1.05-1.32	5.9×10^{-3}	1.19	1.11-1.28	3.0×10^{-6}
rs9551517	29230045	A/G	0.34	1.18	1.09-1.26	1.1×10^{-5}	1.09	1.00-1.20	0.06	1.15	1.08-1.21	3.5×10^{-6}
rs4769628	29232064	G/A	0.15	1.21	1.10-1.33	1.3×10^{-4}	1.17	1.04-1.32	7.1×10^{-3}	1.19	1.11-1.28	3.0×10^{-6}

All results are from unconditional logistic regression and are listed by NCBI Build 37 variant position on chromosome 13 ("Chr13 Position"). "N", total cases / controls in meta-analysis; "Risk/Alt Allele", allele with higher frequency in GWAS cases than controls and alternate allele; "RAF", risk allele frequency in UK GWAS controls; "OR" and "95% CI", allelic odds ratio and 95% confidence interval. Rows for rs4769613 and its near-perfect surrogate rs7328374 are in bold to highlight genome-wide significance ($P<5\times10^{-8}$) of their GWAS+Replication *P* values. Genotypes of rs149427560 for the FINNPEC replication cohort were proxied from rs11619261 (pairwise r^2=0.88 in Finland 1000Genomes phase 3); results for several variants do not include FINNPEC data (rs17555115, rs3829387, rs2096035, rs2296284 and rs9319429) because they were not genotyped in FINNPEC. Most of these variants were not genotyped since they are in strong LD and redundant with other variants genotyped in FINNPEC.

Supplementary Table 3 | Frequency of rs4769613 risk allele C in preeclampsia offspring cases, maternal cases, and controls

Cohort	Offspring Cases				Maternal Cases				Controls	
	Freq	se	OR	P	Freq	se	OR	P	Freq	se
GOPEC	0.576	0.011	1.231	2.74×10^{-5}	0.555	0.008	1.131	0.0014	0.524	0.005
deCODE	0.567	0.017	1.229	5.49×10^{-4}	0.557	0.010	1.095	0.0242	0.531	0.001
MoBa	0.544	0.011	1.100	7.71×10^{-2}	0.519	0.010	0.996	0.9610	0.520	0.008
FINNPEC	0.538	0.015	1.334	2.69×10^{-4}	0.489	0.013	1.096	0.2080	0.466	0.012

Cohort sample numbers (Offspring cases/Maternal cases/Controls) are: GOPEC(1004/1875/5083), deCODE(411/1205/135190), MoBa(1125/1169/1927), FINNPEC(527/729/870); "Freq" and "se" are mean allele frequency and standard error for risk allele C; "OR" and "P" are allelic odds ratio and case-control P value calculated by Pearson χ^2.

Supplementary Table 4 | Transmission Disequilibrium Test of rs4769613 alleles transmitted from heterozygous (C/T) parents to offspring in preeclampsia and control trios of the MoBa cohort

Phenotype	Allele Transmissions			Test for preferential inheritance of allele C	
	C	T	%C	TDT χ^2	P
Preeclampsia case trios	417	361	53.6%	4.031	0.045
Control trios	351	341	50.7%	0.145	0.704

All MoBa trios with genotyped DNA from both parents and child are included.

"C" and "T" show total counts of risk allele C and alternate allele T transmitted from each heterozygous (C/T) parent to an offspring. "%C" is percentage of risk allele C among total transmitted alleles and shows preferential inheritance of allele C (53.6%) by preeclampsia offspring but nearly equal inheritance (50.7%) by offspring in control families; TDT $\chi^2 = (C\text{-}T)^2/(C\text{+}T)$ derived in Spielman *et al* (American Journal of Human Genetics 52:506-516, 1993) shows statistical significance (p<0.045) of preferential allele C inheritance by preeclampsia cases but non-significance (p=0.704) in control trios.

Supplementary Table 5 | Test for sex difference between heterozygous (C/T) mothers and fathers in transmission of rs4769613 alleles to preeclampsia offspring of MoBa cohort

Parental Sex	Allele Transmissions			Test Sex Difference	
	C	T	%C	χ^2	P
Mother	164	134	55.0%	0.0462	0.83
Father	169	143	54.2%		

"C" and "T" show total counts of risk allele C and alternate allele T transmitted from each heterozygous (C/T) parent to an offspring. "%C" is percentage of allele C among total transmitted alleles. Preeclampsia trios are included only if maternal and paternal allele transmissions are unambiguous (i.e. preeclampsia trios were excluded if both parents and offspring were heterozygous). "χ^2" is 2x2 Pearson χ^2 Test on counts of allele C and T transmitted from heterozygous mothers versus fathers in unambiguous preeclampsia trios.

Supplementary Table 6 | Association results for Red Blood Cell (RBC) count, haemoglobin and haematocrit for preeclampsia associated variants

SNP	Risk allele*	Other allele	Risk allele frequency	Trait[†]	N samples[‡]	*P* value	Effect[§]	95%CI
rs4769613	C	T	0.53	RBC count	270,314	5.0×10^{-4}	-0.011	-0.005 to -0.017
				Haemoglobin	272,616	7.1×10^{-4}	-0.010	-0.004 to -0.016
				Haematocrit	268,150	2.5×10^{-3}	-0.009	-0.003 to -0.015
rs12050029	G	A	0.14	RBC count	270,314	1.5×10^{-7}	-0.024	-0.033 to -0.015
				Haemoglobin	272,616	1.1×10^{-4}	-0.016	-0.024 to -0.008
				Haematocrit	268,150	4.0×10^{-6}	-0.019	-0.027 to -0.011

* Allele associated with increased risk of preeclampsia

[†] Results are shown for three correlated traits

[‡] Number of individuals included in the analysis

[§] Effect estimate in units of standard deviation is reported for the designated Risk allele. Association was tested using generalized linear regression. Measurements were adjusted for age, sex and measurement site, and average was taken over the available measurements after adjustment and inverse normal transformation.

Supplementary Table 7 | Characteristics of pregnancies from GWAS and Replication cohorts

Cohort and country of origin	Group	Maternal age		Primi-parous	Highest SBP (mm Hg)		Highest DBP (mm Hg)		Gestation at delivery (weeks)		Offspring birthweight (grams)	
		Mean	SD		Mean	SD	Mean	SD	Median	IQR	Median	IQR
GOPEC$^{\lozenge}$ UK	Cases n=1157	29.6	5.6	78%	167	18	111	9	37	34-38	2580	1848-3193
ALSPAC UK	Cases n=146	28.7	5.3	68%	159	13	108	9	39	37-40	3240	2710-3680
ALSPAC UK	Controls n=6130	28.6	4.8	41%	127	10	78	7	40	39-41	3460	3160-3780
deCODE$^{\square}$ Iceland	Cases n=1507	27.8	6.0	66%	155‡	16	107‡	9	39	37-40	3310§	2752-3744
MoBa Norway	Cases n=1200	29.1	4.8	66%	Not available		Not available		39	38-41	3325	2812-3740
MoBa Norway	Controls n=1200	30.1	4.4	42%	Not available		Not available		40	39-41	3690	3341-4010
FINNPEC Finland	Cases n=605	29.9	5.6	75%	166	17	109	9	38	36-39	2913	2390-3380
FINNPEC Finland	Controls n=800	29.7	5.1	56%	128	14	85	10	40	39-41	3590	3260-3930

Phenotypic data were available for >99% of pregnancies unless otherwise indicated.

SBP, systolic blood pressure; DBP, diastolic blood pressure; IQR, interquartile range

‡ 491 pregnancies

§ 1294 pregnancies

$^{\lozenge}$ Control data for GOPEC GWAS were derived from 5297 unselected subjects from the UK population included in the 1958 Birth Cohort and donors to the National Blood Service

$^{\square}$ Control data for deCODE GWAS were derived from 296,865 unselected subjects from the Icelandic population

Supplementary Table 8 | Early and late onset cases of pre-eclampsia

Cohort	EO-preeclampsia	LO-preeclampsia
GOPEC	274	625
deCODE	176	360
FINNPEC	123	479
MoBa	58	1095

Number of cases of early onset and late onset preeclampsia used in data analysis. Only cases where gestational age at diagnosis of pre-eclampsia was unequivocally recorded were included in sub-group analysis.

"EO-preeclampsia" is early onset pre-eclampsia, diagnosed at <34 weeks gestation. "LO-preeclampsia" is late onset pre-eclampsia, diagnosed at ≥34 weeks gestation.

Supplementary Table 9 | Per cohort pedigree counts used for EMIM analysis

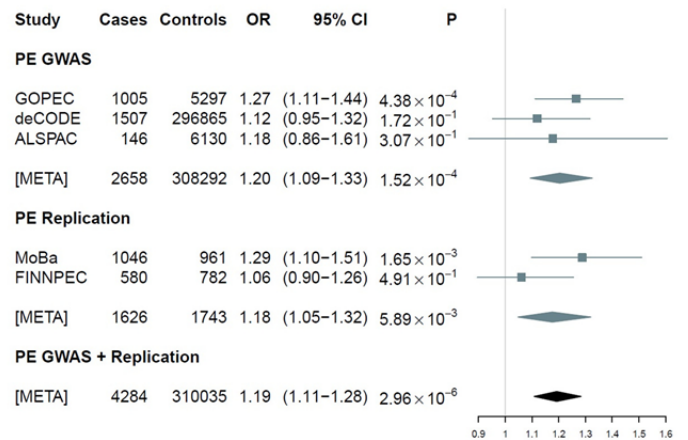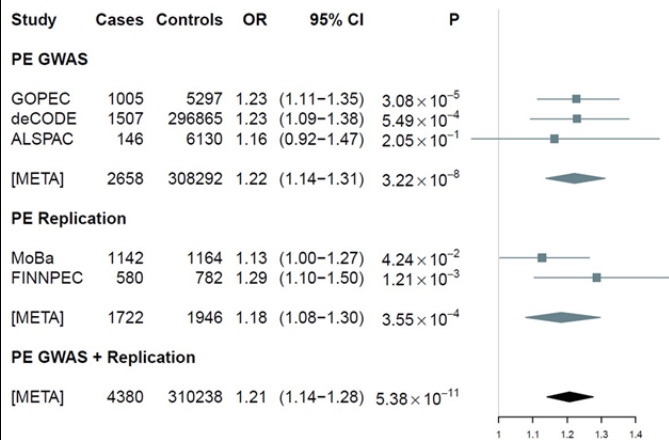| Group	MoBa	GOPEC	FINNPEC	Total
Case Trios	798	0	0	798
Case Mother-Offspring Duos	315	953	403	1671
Case Father-Offspring Duos	9	0	0	9
Case Mothers*	50	922	326	1298
Case Fathers*	33	0	0	33
Case Offspring*	6	51	126	183
Case Parents[†]	9	0	0	9
Control Mother-Offspring Duos	429	0	487	916
Control Father-Offspring Duos	5	0	0	5
Controls[‡]	1494	5083	383	6960

*Subjects without other family members

[†]Two parents without their case offspring
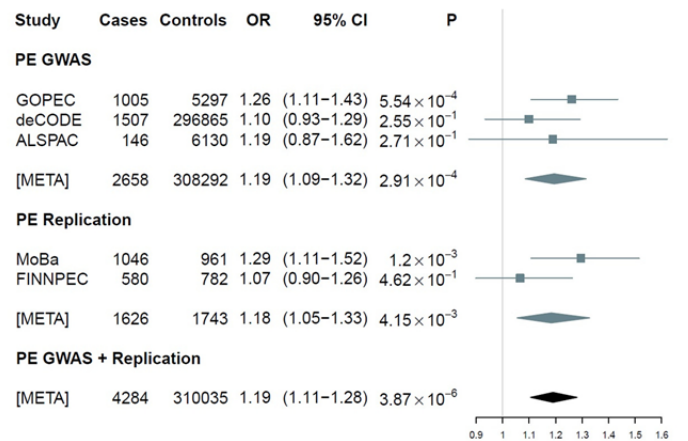
[‡]Includes parents in control trios

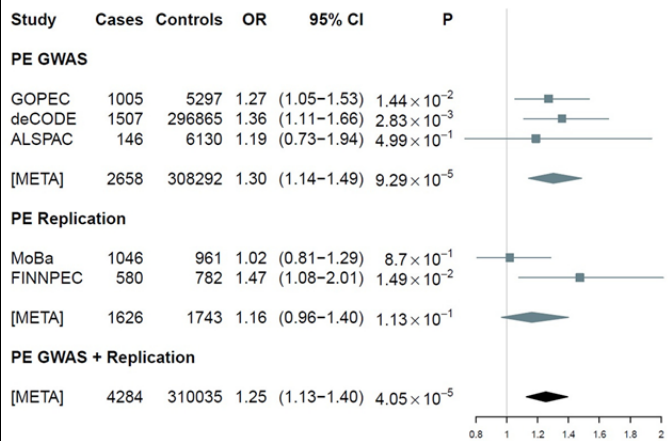**rs4769613 [hg19: chr13−29138609; risk: C(0.525); other: T; Phet: 0.678]**

Study	Cases	Controls	OR	95% CI	P
PE GWAS					
GOPEC	1005	5297	1.23	(1.11−1.35)	3.08×10^{-5}
deCODE	1507	296865	1.23	(1.09−1.38)	5.49×10^{-4}
ALSPAC	146	6130	1.16	(0.92−1.47)	2.05×10^{-1}
[META]	2658	308292	1.22	(1.14−1.31)	3.22×10^{-8}
PE Replication					
MoBa	1142	1164	1.13	(1.00−1.27)	4.24×10^{-2}
FINNPEC	580	782	1.29	(1.10−1.50)	1.21×10^{-3}
[META]	1722	1946	1.18	(1.08−1.30)	3.55×10^{-4}
PE GWAS + Replication					
[META]	4380	310238	1.21	(1.14−1.28)	5.38×10^{-11}

rs12050029 [hg19: chr13−29227519; risk: G(0.143); other: A; Phet: 0.387]

Study	Cases	Controls	OR	95% CI	P
PE GWAS					
GOPEC	1005	5297	1.27	(1.11−1.44)	4.38×10^{-4}
deCODE	1507	296865	1.12	(0.95−1.32)	1.72×10^{-1}
ALSPAC	146	6130	1.18	(0.86−1.61)	3.07×10^{-1}
[META]	2658	308292	1.20	(1.09−1.33)	1.52×10^{-4}
PE Replication					
MoBa	1046	961	1.29	(1.10−1.51)	1.65×10^{-3}
FINNPEC	580	782	1.06	(0.90−1.26)	4.91×10^{-1}
[META]	1626	1743	1.18	(1.05−1.32)	5.89×10^{-3}
PE GWAS + Replication					
[META]	4284	310035	1.19	(1.11−1.28)	2.96×10^{-6}

rs149427560 [hg19: chr13−29105870; risk: G(0.064); other: GGT; Phet: 0.311]

Study	Cases	Controls	OR	95% CI	P
PE GWAS					
GOPEC	1005	5297	1.27	(1.05−1.53)	1.44×10^{-2}
deCODE	1507	296865	1.36	(1.11−1.66)	2.83×10^{-3}
ALSPAC	146	6130	1.19	(0.73−1.94)	4.99×10^{-1}
[META]	2658	308292	1.30	(1.14−1.49)	9.29×10^{-5}
PE Replication					
MoBa	1046	961	1.02	(0.81−1.29)	8.7×10^{-1}
FINNPEC	580	782	1.47	(1.08−2.01)	1.49×10^{-2}
[META]	1626	1743	1.16	(0.96−1.40)	1.13×10^{-1}
PE GWAS + Replication					
[META]	4284	310035	1.25	(1.13−1.40)	4.05×10^{-5}

rs12050029 [hg19: chr13−29227519; risk: G(0.143); other: A; Phet: 0.354]
(Conditional on rs4769613)

Study	Cases	Controls	OR	95% CI	P
PE GWAS					
GOPEC	1005	5297	1.26	(1.11−1.43)	5.54×10^{-4}
deCODE	1507	296865	1.10	(0.93−1.29)	2.55×10^{-1}
ALSPAC	146	6130	1.19	(0.87−1.62)	2.71×10^{-1}
[META]	2658	308292	1.19	(1.09−1.32)	2.91×10^{-4}
PE Replication					
MoBa	1046	961	1.29	(1.11−1.52)	1.2×10^{-3}
FINNPEC	580	782	1.07	(0.90−1.26)	4.62×10^{-1}
[META]	1626	1743	1.18	(1.05−1.33)	4.15×10^{-3}
PE GWAS + Replication					
[META]	4284	310035	1.19	(1.11−1.28)	3.87×10^{-6}

rs149427560 [hg19: chr13−29105870; risk: G(0.064); other: GGT; Phet: 0.287]
(Conditional on rs4769613)

Study	Cases	Controls	OR	95% CI	P
PE GWAS					
GOPEC	1005	5297	1.27	(1.05−1.53)	1.52×10^{-2}
deCODE	1507	296865	1.39	(1.13−1.70)	1.81×10^{-3}
ALSPAC	146	6130	1.20	(0.74−1.96)	4.58×10^{-1}
[META]	2658	308292	1.31	(1.15−1.50)	6.66×10^{-5}
PE Replication					
MoBa	1046	961	1.03	(0.82−1.29)	8.22×10^{-1}
FINNPEC	580	782	1.51	(1.10−2.06)	1.07×10^{-2}
[META]	1626	1743	1.17	(0.97−1.41)	9.1×10^{-2}
PE GWAS + Replication					
[META]	4284	310035	1.26	(1.13−1.41)	2.39×10^{-5}

rs11619261 [hg19: chr13−29078077; risk: A(0.065); other: G; Phet: 0.367]
(Conditional on rs4769613)

Study	Cases	Controls	OR	95% CI	P
PE GWAS					
GOPEC	1005	5297	1.21	(1.01−1.45)	4.13×10^{-2}
deCODE	1507	296865	1.35	(1.10−1.65)	3.6×10^{-3}
ALSPAC	146	6130	1.17	(0.73−1.87)	5.24×10^{-1}
[META]	2658	308292	1.26	(1.11−1.44)	4.18×10^{-4}
PE Replication					
MoBa	1142	1164	1.07	(0.88−1.31)	4.91×10^{-1}
FINNPEC	580	782	1.51	(1.10−2.06)	1.07×10^{-2}
[META]	1722	1946	1.18	(1.00−1.40)	5.12×10^{-2}
PE GWAS + Replication					
[META]	4380	310238	1.23	(1.11−1.37)	6.75×10^{-5}

Supplementary Figure 1

Forest plots for GWAS and replication cohorts at independent variants near *FLT1* giving evidence for association with preeclampsia.

Forests plots are presented in order of strength of association with preeclampsia (PE) (see Table 1). The top line of each plot gives variant rs number, position on chromosome 13 (human genome build 19), risk allele (i.e. allele with higher frequency in cases than controls in the GWAS meta-analysis)and its population frequency in parenthesis, other allele, and Phet giving the *P* value for heterogeneity of odd ratios (OR) in the five cohorts. Subsequent lines provide a breakdown of results for each GWAS and replication cohort including number of cases and controls, the allelic case-control OR and its 95% confidence interval (95% CI) and case-control association *P* value. [META] indicates corresponding meta-analysis results for GWAS cohorts, replication cohorts, or GWAS and replication cohorts combined; each meta-analysis allelic OR is represented by a "diamond" whose width corresponds to the 95% CI. The first three Forest plots give unconditional logistic regression results and the last three plots give logistic regression results that condition out the effect of rs4769613. The SNP rs11619261 is included here because it is used as a proxy for rs149427560 in the FINNPEC cohort (see Table 1).
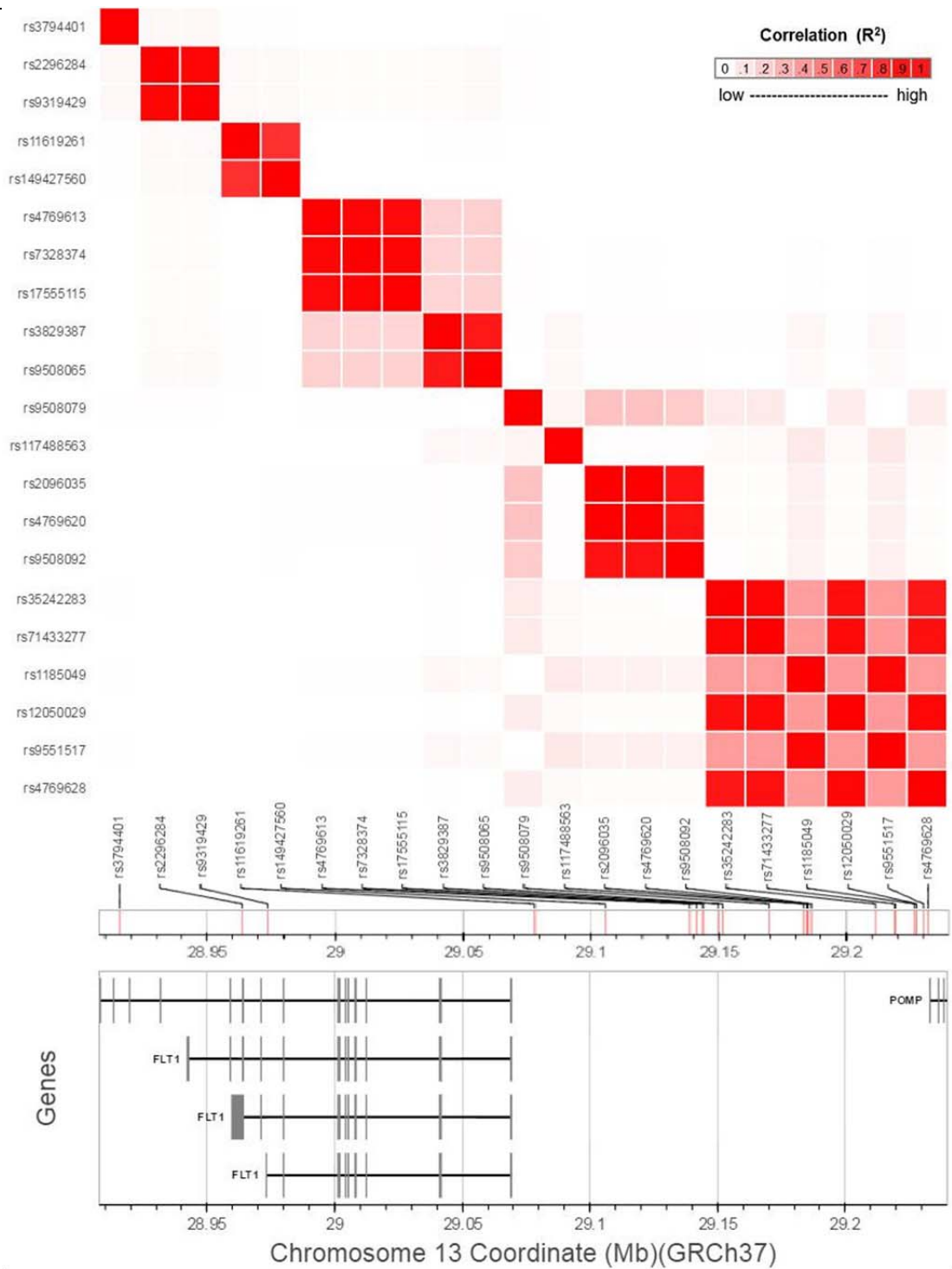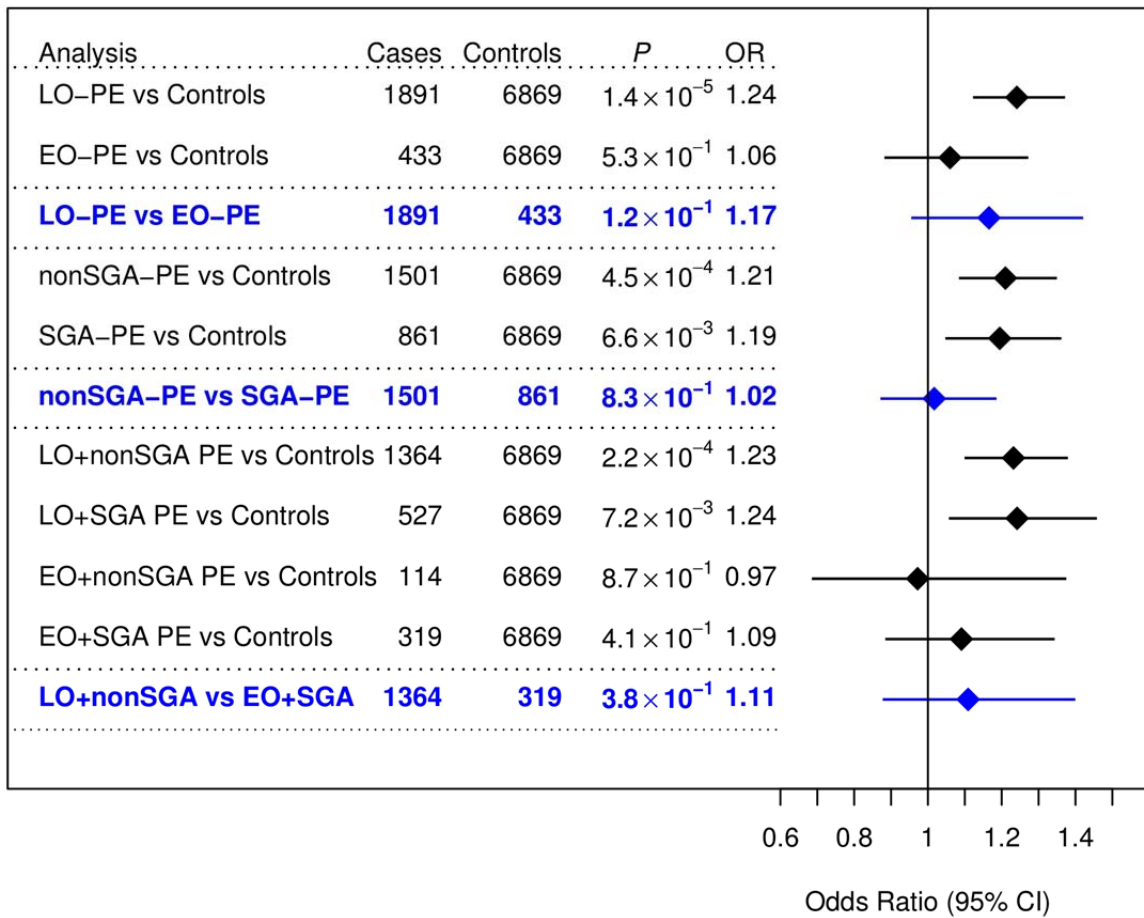
Supplementary Figure 2

Linkage disequilibrium (LD) matrix showing pairwise LD r^2 values among all variants near *FLT1* selected for follow-up in Replication cohorts.

LD r^2 values were generated by LDlink software (http://analysistools.nci.nih.gov/LDlink) from 1000Genomes Phase 3 genotypes for all European populations. Variants are ordered by chromosomal position and shown in

relation to *FLT1* and *POMP*. Intensity of red shading (see colour legend) shows approximate r^2 value for each variant pair and implies that the 21 follow-up variants define 9 LD "blocks" with high pairwise r^2 within each block but low pairwise r^2 for members of different blocks.

rs12050029 : PE Subtype Meta−analysis Overview

Analysis	Cases	Controls	P	OR
LO−PE vs Controls	1891	6869	1.4×10^{-5}	1.24
EO−PE vs Controls	433	6869	5.3×10^{-1}	1.06
LO−PE vs EO−PE	**1891**	**433**	$\mathbf{1.2\times10^{-1}}$	**1.17**
nonSGA−PE vs Controls	1501	6869	4.5×10^{-4}	1.21
SGA−PE vs Controls	861	6869	6.6×10^{-3}	1.19
nonSGA−PE vs SGA−PE	**1501**	**861**	$\mathbf{8.3\times10^{-1}}$	**1.02**
LO+nonSGA PE vs Controls	1364	6869	2.2×10^{-4}	1.23
LO+SGA PE vs Controls	527	6869	7.2×10^{-3}	1.24
EO+nonSGA PE vs Controls	114	6869	8.7×10^{-1}	0.97
EO+SGA PE vs Controls	319	6869	4.1×10^{-1}	1.09
LO+nonSGA vs EO+SGA	**1364**	**319**	$\mathbf{3.8\times10^{-1}}$	**1.11**

Odds Ratio (95% CI)

Supplementary Figure 3

rs12050029: Preeclampsia subtype meta-analysis overview.

Forest plot for preeclampsia (PE) subtypes defined by early and late onset (EO-PE, LO-PE) and by birthweight that is small-for-gestational age (SGA-PE) or not (nonSGA-PE). Case-control comparison shows risk allele G is significantly associated with LO-PE (P=1.4x10^{-5}), nonSGA-PE (P=4.5x10^{-4}) and SGA-PE (P=6.6x10^{-3}). The strength of association does not differ between SGA-PE and nonSGA-PE.