

1 **Article [Working Title]:** A multi-ancestry genome-wide association study identifies
2 novel candidate loci in the *RARB* gene associated with hypertensive disorders of
3 pregnancy

4 **Short Title:** Multi-ancestry GWAS of hypertension in pregnancy

5 **Jasmine A. Mack**^{1,2}, Adam Burkholder¹, Farida S. Akhtari¹, John S. House¹, Ulla
6 Sovio², Gordon C.S. Smith², Charles P. Schmitt¹, David C.Fargo¹, Janet E. Hall¹, Alison
7 A. Motsinger-Reif¹

8 ¹ National Institute of Environmental Health Sciences, Research Triangle Park, NC
9 27709, USA

10 ² University of Cambridge, Department of Obstetrics and Gynaecology, Cambridge, UK,
11 CB2 0SW

12

13 **Corresponding author:** Alison Motsinger-Reif

14 alison.motsinger-reif@nih.gov

15

16

17 **Word count:** 7917

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32 **Abstract**

33 **Background**

34 Genetic factors related to pregnancy-related traits are understudied, especially among
35 ancestrally diverse cohorts. This study assessed maternal contributions to hypertensive
36 disorders of pregnancy (HDP) in multi-ancestry cohorts.

37 **Methods**

38 We performed a genome-wide association study of HDP using data from the
39 Personalized Environment and Genes Study (PEGS) cohort (USA) with validation in the
40 UK Biobank (UKBB). We performed gene-level and gene-set analyses and tested the
41 association of polygenic scores (PGS) for systolic blood pressure (SBP), preeclampsia
42 (PE), and gestational hypertension (GH).

43 **Results**

44 We identified two novel maternal genome-wide significant associations with HDP. The
45 lead independent variants were rs114954125 on chromosome 2 (near *LRP1B*; OR
46 (95% CI): 3.03 (2.05, 4.49); $P=3.19 \times 10^{-8}$) and rs61176331 on chromosome 3 (near
47 *RARB*; OR (95% CI): 3.09 (2.11, 4.53); $P=7.97 \times 10^{-9}$). We validated rs61176331 in the
48 UKBB ($P=3.73 \times 10^{-2}$). When aggregating SNPs by genes, *RARB* ($P=1.36 \times 10^{-3}$) and
49 *RN7SL283P* ($P=2.56 \times 10^{-2}$) were associated with HDP. Inflammatory and
50 immunological biological pathways were most strongly related to HDP-associated
51 genes. While all blood pressure and HDP-related PGS were significantly associated
52 with HDP in PEGS, the SBP PGS was a stronger predictor of HDP (area under the
53 curve (AUC): 0.57; $R^2=0.7\%$) compared to the PE PGS (AUC: 0.53; $R^2=0.2\%$).

54 **Conclusion**

55 Our study is the first to identify and validate maternal genetic variants near *RARB*
56 associated with HDP. The findings demonstrate the power of multi-ancestry studies for
57 genetic discovery and highlight the relationship between immune response and HDP
58 and the utility of PGS for risk prediction.

59

60 ClinicalTrials.gov Identifier for PEGS: NCT00341237

61

62

63

64

65

66 **Non-standard Abbreviations and Acronyms**

AFR	African ancestry
AMR	Admixed American ancestry
EAF	Effect allele frequency
EAS	East Asian ancestry
EUR	European ancestry
GH	Gestational hypertension
GWAS	Genome-wide association study
HDP	Hypertensive disorders of pregnancy
KEGG	Kyoto Encyclopedia of Genes and Genomes
MAC	Minor allele count
MIM	Mendelian Inheritance in Man
PCA	Principal components analysis
PE	Preeclampsia
PEGS	Personalized Environment and Genes Study
PGS	Polygenic score
SAS	South Asian ancestry
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant
UKBB	UK Biobank

67

68

69

70

71

72

73

74

75

76

77

78

79

80

81

82 **Introduction**

83 Hypertensive disorders of pregnancy (HDP) are a heterogeneous set of adverse
84 pregnancy outcomes that occur in the perinatal period comprising gestational
85 hypertension (GH), preeclampsia/eclampsia (PE), chronic, pre-existing hypertension
86 with or without superimposed PE, and postpartum PE. HDP occur in approximately 5-
87 10% of pregnancies worldwide¹ although prevalence varies by disease subtype and
88 region^{2,3}. HDP incidence increased by 11% from 1990 to 2019⁴. Globally, an estimated
89 50,000-75,000 deaths occur annually due to PE⁵. HDP are the primary cause of
90 maternal morbidity and mortality² in high-resource countries and a secondary cause
91 globally⁶.

92

93 There is racial and ethnic disparity regarding HDP prevalence and incidence as well as
94 the contributing risk factors. Black, Hispanic, and Indigenous individuals experience
95 disproportionately high rates of HDP compared to White individuals^{7,8}. Risk factors for
96 HDP include pre-pregnancy body mass index (BMI) > 25 kg/m², nulliparity, prior GH,
97 maternal age of more than 35 years, family history of HDP, and pre-existing
98 hypertension⁹.

99

100 The pathophysiology of HDP is not fully understood, but dysfunction in placental growth
101 factors and regulators of angiogenesis are suggested mechanisms, with maternal and
102 fetal contributions¹⁰. This mechanism is hypothesized to be associated with diminished
103 placental perfusion causing systemic vascular endothelial dysfunction through the

104 release of antiangiogenic factors from the placenta^{11,12}. Placental hypoperfusion is
105 thought to be caused by failure of transformation of uterine spiral arteries into low-
106 resistance conduits, which is a normal feature of invasion by extravillous
107 cytotrophoblast^{13,14}. Failure of spiral artery transformation has been linked to fetal
108 growth restriction, miscarriage, and PE¹⁴. Moreover, placental factors can interact with
109 maternal genetics, epigenetics, and environmental factors to influence the response to
110 placental dysfunction¹⁵. Maternal immune system changes are involved in the
111 development and pathogenesis of PE^{16–22}, but the mechanisms are incompletely
112 understood.

113
114 Early genetic studies of HDP have been predominantly focused on individuals of
115 European ancestry^{23,24}. In research in large cohorts with genetic data, participants from
116 underrepresented racial/ethnic groups are consistently excluded from published
117 analyses due to small sample sizes and the use of sparse methods to account for
118 population structure^{25–27}.

119
120 Prior genome-wide association studies (GWAS) of PE have revealed strong signals
121 near the *FLT1* gene (MIM:165070) based on a meta-analysis of fetal genotypes from
122 five cohorts, of which all participants were of European ancestry: Genetics of Pre-
123 Eclampsia (GOPEC), Avon Longitudinal Study of Parents and Children (ALSPAC),
124 deCODE, Norwegian Mother and Child (MoBa), and Finnish Genetics of Pre-eclampsia
125 Consortium (FINNPEC)²⁴. The largest maternal genome-wide multi-ancestry meta-
126 analysis of PE and GH was conducted in cohorts comprising FinnGen, Estonian

127 biobank, Genes and Health, Michigan Genomics Initiative, Mass General Brigham
128 Biobank, Biobank Japan, BioMe Biobank, InterPregGen Consortium, Trøndelag Health
129 Study (HUNT), UKBB, Penn Medicine Biobank, and Nulliparous Pregnancy Outcomes
130 Study: Monitoring Mothers-to-Be (nuMoM2b)²⁸. Among 31,091 HDP cases assessed in
131 the meta-analysis, 83% were of European ancestry. Multi-ancestry meta-analyses have
132 been conducted in diverse cohorts such as the Pan UK Biobank (UKBB), but the control
133 groups of the GWAS also included female UKBB participants who have never been
134 pregnant²⁹.

135
136 To increase diversity and support the representation of non-European populations in
137 genomics research, GWAS that include minority racial and ethnic groups can be
138 conducted in currently available data from diverse cohorts. The results can be stratified
139 by ancestry, and meta-analyses can jointly model effects in a multi-ancestry analysis.
140 Building upon previous multi-ancestry genetic studies, we conducted multi-ancestry
141 GWAS in whole-genome sequencing (WGS) data from a discovery cohort from the
142 North Carolina-based Personalized Environment and Genes Study (PEGS) using a
143 case-control study design. The joint modeling exhibited improved power for genetic
144 discovery, enabling the first discovery of a maternal genetic association near the retinoic
145 acid receptor (*RARB*) gene (MIM: 180220) with HDP. We validated this association in
146 the UKBB and a meta-analysis in PEGS and the UKBB. In downstream analyses, we
147 performed gene-level analysis and gene-set analyses to elucidate the biological
148 mechanisms involved in HDP. Further, in the PEGS dataset, we assessed the predictive

149 ability of existing polygenic scores (PGS) for PE and GH, with HDP in predicting risk of
150 these conditions.

151

152 **Methods**

153 The data used in this study are available to researchers through application to dbGAP
154 the UK Biobank (<https://www.ukbiobank.ac.uk/register-apply/>). Reference panel data
155 from the 1000 Genomes Project and polygenic scores from the PGS Catalog are
156 publicly available. The Major Resources Table provides details. The summary statistics
157 that support the study findings are available from the corresponding author upon
158 reasonable request.

159

160 **Study Populations**

161 PEGS Cohort

162 For the discovery cohort, we selected participants from PEGS, a long-term, prospective
163 cohort based in North Carolina³⁰. PEGS collects extensive clinical, genetic,
164 environmental, and self-reported health information to investigate disease etiology and
165 risk factors across phenotypes in a diverse group of 19,685 individuals.

166

167 Participants are administered the Health and Exposure Survey, External Exposome
168 Survey, and Internal Exposome Survey that request information on sociodemographic
169 factors, individual and family medical history, lifestyle factors, medication use, diet, and
170 physical environmental exposures from childhood to adulthood. While PEGS exclusion
171 criteria include women who self-report being pregnant, a small number of participants

172 were pregnant when they completed the Internal Exposome Study, as this was
173 administered later in the data collection process.

174
175 For a subset of 4,737 PEGS participants who provided blood and urine samples, WGS
176 data were processed, with the optional return of results for select health conditions. The
177 Broad Institute sequenced all samples on the NovaSeq 6000 platform with a target
178 genome-wide read depth of 30x (24 samples per flow cell).

179
180 For downstream analyses, we removed related individuals in the PEGS genomic cohort
181 by analyzing the genotypes of all individuals using KING 2.2.5³¹ to generate pairwise
182 kinship estimation. We excluded second-degree or closer relatives using a kinship
183 threshold of 0.0884. We estimated genome-wide local ancestry on the autosomes and
184 X chromosome using RFMix version 2, a discriminative modeling approach³². We
185 classified genomic positions by five genetic similarity groups: African (AFR), Admixed
186 American (AMR), East Asian (EAS), European (EUR), and South Asian (SAS). We
187 classified individuals into one of the genetic similarity groups if any genome-wide local
188 ancestry proportion was greater than 0.50. Otherwise, we classified participants as
189 “None”.

190
191 The PEGS cohort is a registered clinical trial, and all participants provided written
192 informed consent (ClinicalTrials.gov Identifier: NCT00341237)³⁰. We obtained the data
193 used for the analyses from PEGS Data Freeze version 2.0 created in August 2021. The

194 **Supplemental Methods** section provides details on genome sequencing, quality
195 control, and local ancestry inference.

196

197 UK Biobank

198 For the validation cohort, we selected participants from the UK Biobank (UKBB). UKBB
199 is a large, deeply phenotyped genomic cohort comprising 488,366 participants aged 40-
200 69 years at recruitment from 2006 to 2010. About 6% of participants (>30,000) identified
201 as having a non-White ethnic background^{25,33}. Genetic data are linked to electronic
202 health records and survey data, and UKBB has information on diagnosis history, brain
203 imaging, medication use, lifestyle factors, deprivation indices, environmental exposure
204 data, and sociodemographic characteristics.

205

206 Ethical approval for UKBB was obtained from the North West Centre for Research
207 Committee (11/NW/0382). Informed consent was obtained for all participants. The
208 following work described was approved by the UKBB team under Application No.
209 57849. We obtained predominant ancestry assignments from UKBB Returned Dataset
210 2442³⁴. The **Supplemental Methods** provides details on genotyping and genetic
211 similarity groups in UKBB.

212

213 *Case-Control Definition*

214 **Figure S1** outlines the steps taken to define cases and controls in the PEGS cohort.
215 Because the data available for PEGS and UKBB are not standardized, we defined
216 cases and controls separately for each cohort and harmonized the case-control sets as

217 much as possible. The **Supplemental Methods** section provides details on case-
218 control definition.

219

220 PEGS

221 We defined cases and controls for 4,933 PEGS participants who self-reported at least
222 one pregnancy. After filtering, there were 202 GH cases, which is the primary outcome,
223 92 PE cases, 78 GH only cases, 223 GH/PE cases, and 1,569 controls. **Table S1** lists
224 the phenotypes and survey questions used to define cases and controls.

225

226 UKBB

227 We defined cases and controls in the UKBB population using the International
228 Classification of Diseases (ICD) codes (ICD10) for pregnancy-induced hypertension
229 with and without significant proteinuria (PE and hemolysis, elevated liver enzymes and
230 low platelets (HELLP) syndrome), eclampsia, and unspecified maternal hypertension
231 (**Table S1**). After filtering, there were 1,319 HDP cases and 2,050 controls.

232

233 ***Statistical Analysis***

234 For each cohort, we calculated frequencies and proportions for categorical variables
235 and mean/standard deviation for normally distributed continuous variables, including
236 covariates and sociodemographic characteristics. For continuous variables without
237 normal distribution, we report median and interquartile ranges. We estimated *P* values
238 by Fisher's exact test to assess associations among covariates and performed t-tests

239 for continuous variables. For variables without normal distribution, such as BMI, we
240 used the Mann-Whitney U test.

241

242 Genome-wide Association Study for Common Variants

243 We conducted GWAS (MAF \geq 0.03) using multivariable logistic regression to model the
244 odds of HDP, assuming an additive allelic genetic model. We adjusted for the top 10
245 genetic principal components, self-reported race/ethnicity, age and BMI at enrollment,
246 and gravidity. We used PLINK2.0³⁵ for all GWAS with Firth-corrected logistic regression
247 modeling to account for case-control imbalance³⁶. Because information on gravidity was
248 available for few UKBB participants, we excluded this covariate in the UKBB analyses.

249 We included genotyping batch as a covariate in the UKBB analyses given that two types
250 of arrays were used in genotyping with 106 batches²⁵. We used a threshold of $P <$
251 5×10^{-8} for genome-wide significance and $P < 1 \times 10^{-6}$ for suggestive significance.

252

253 In addition to the main outcome of HDP, we performed GWAS for three other
254 phenotypes in the PEGS discovery cohort (PE, GH without PE, and GH and/or PE) for
255 all individuals and separately for the AFR and EUR similarity groups. We adjusted for
256 within-group variation using the top 10 principal components, age and BMI at
257 enrollment, and gravidity. The AMR, EAS, and SAS similarity groups had insufficient
258 numbers of participants to perform stratified analyses (**Table 1**).

259

260 To validate our findings, we filtered GWAS summary statistics from the PEGS multi-
261 ancestry analysis to include only variants with suggestive association ($P < 1 \times 10^{-6}$).

262 We assessed these variants in the UKBB using harmonized cases and controls. We
263 used GCTA-COJO for conditional and joint analyses to identify independently
264 associated loci³⁷. We used METAL to perform inverse variance-weighted meta-analyses
265 for variants with suggestive association in PEGS and UKBB³⁸.

266

267 Gene-based Association Analysis and Gene-set Analysis

268 To increase statistical power for discovery, we used MAGMA to perform gene-set
269 association analysis, adjusting for the top 10 genetic principal components, self-
270 reported race/ethnicity, age and BMI at enrollment, and gravidity. We used Ensembl
271 annotation on GRCh38 to map the SNVs from the GWAS to 31,427 genes³⁹. We utilized
272 the results from the gene-level analysis to perform gene-set analysis in MAGMA^{40,41}.
273 We used the Kyoto Encyclopedia of Genes and Genomes (KEGG) to define gene sets
274 by biological pathways and conditions^{42,43}. To control the false discovery rate (FDR), we
275 used the Benjamini-Hochberg method to adjust P values from the statistical analyses for
276 multiple comparisons. We define statistical significance by an adjusted P value of 0.05
277 and suggestive FDR-corrected significance by an adjusted P value of 0.10.

278

279 ***Application of Polygenic Scores***

280 To assess the predictive value of SBP, PE, and GH for HDP in PEGS, we applied an
281 SBP PGS derived from the UKBB female sub-cohort in Kauko et al.⁴⁴ and separate
282 scores for SBP, PE, and GH from meta-analyses in Honigberg et al.²⁸. We separately
283 computed the four PGS in the PEGS pregnancy ($n = 1,771$) and female sub-cohorts (n

284 = 3,102) to model HDP and essential hypertension, respectively. We standardized each
285 PGS. The **Supplemental Methods section** provides details.

286
287 We performed logistic regression with Firth correction to assess the association
288 between HDP risk and the standardized PGS, adjusting for the top ten genetic principal
289 components, self-reported race/ethnicity, age and BMI at enrollment, and gravidity. For
290 further validation, we tested the association of each PGS with essential hypertension in
291 the PEGS female sub-cohort, adjusting for the same covariates except gravidity.

292 Additionally, we performed stratified analyses for the AFR and EUR genetic similarity
293 groups. For each PGS, we report odds ratios per standard deviation with a 95% CI. We
294 calculated the AUC for each model using the pROC package in R to assess
295 discriminative ability. All PGS analyses were performed in R 4.3.1 with a significance
296 threshold of 0.05.

297
298 The Major Resources Table in the **Supplemental Materials** provides details on data
299 and code availability.

300

301 **Results**

302 In the PEGS cohort, participants with HDP (11.4% of the study sample) were younger,
303 had a higher BMI at enrollment, and had more pregnancies compared to participants in
304 the control group ($P < 0.001$; **Table 1**). Among HDP cases, 4.5% experienced recurrent
305 diagnoses of PE, and 18.8% experienced GH during two or more pregnancies.

306 Race/ethnicity distribution, education level, and income were similar for the case and

307 control groups (**Table 1**). In the UKBB cohort, race/ethnicity, predominant genetic
308 similarity group, and education level were similar for cases and controls (**Table S2**).
309 Similar to PEGS, UKBB participants with HDP were younger ($P=0.024$) and had a
310 higher BMI at recruitment ($P<0.001$) compared to controls.

311

312 Novel Maternal Loci Identified on *RARB*

313 In the PEGS GWAS, 11 variants in two loci met genome-wide significance (**Figure 1A**;
314 **Table S3**) for HDP. An association was found for a significant novel candidate region in
315 the maternal genome (202 cases and 1,569 controls) consisting of intergenic variants
316 near the low-density lipoprotein (LDL) receptor-related protein 1B (*LRP1B*;
317 MIM:608766) gene on chromosome 2 (**Figure 1B**). A stronger signal was observed for
318 a region identified as intronic to the *RARB* gene (MIM:180220) on chromosome 3
319 (**Figure 1C**). The lead variant was rs61176331, with an OR of 3.09 for HDP (EAF:
320 0.048; 95% confidence interval (95% CI): (2.11, 4.53); $P=7.97 \times 10^{-9}$). The second most
321 significant variant was rs114954125, with an OR of 3.03 (EAF: 0.051; 95% CI: (2.05,
322 4.49); $P=3.19 \times 10^{-8}$) (**Table 2**). A genomic inflation factor of 1.02 indicated sufficient
323 adjustment for population stratification (**Figure S2A**). Conditional and joint analyses
324 confirmed that rs61176331 ($P_{\text{joint}}=9.56 \times 10^{-9}$) and rs114954125 ($P_{\text{oint}}=3.79 \times 10^{-8}$) are
325 independent loci (**Table 2**). In results stratified by genetic similarity group, these loci had
326 larger effect sizes although the P values were attenuated. rs61176331 had an OR of
327 6.98 and 2.37 for the AFR and EUR genetic similarity groups, respectively (**Table 2**;
328 **Figure S2B**). rs114954125 had a similar effect size as the multi-ancestry GWAS for the
329 EUR genetic similarity group, but this variant did not reach the quality control threshold

330 in the AFR genetic similarity group. We validated rs61176331 in the UKBB analysis
331 (OR: 1.29; 95% CI: (1.02, 1.65); $P=3.73 \times 10^{-2}$) (**Figure S2B**). In the meta-analysis
332 including PEGS and UKBB, strong associations remained for the candidate region near
333 *RARB* (**Table 2; Table S3**).

334
335 In a sensitivity analysis with three HDP subtypes as the outcome (PE, GH only, and GH
336 and/or PE; **Figure S3**), there was distinct attenuation of signals in the multi-ancestry
337 HDP GWAS with PE as the outcome (21 cases) (**Figure S3A**). In this GWAS, the
338 candidate signal on *RARB* was attenuated below the genome-wide significance
339 threshold while rs114954125 on chromosome 2 remained significant. (**Figure S3E**).

340

341 Significant Gene-wide Associations for *RARB*

342 **Figure 2** is a gene-based Manhattan plot. Two gene signals met the FDR-corrected
343 significance threshold—*RARB* ($P=1.36 \times 10^{-3}$) and *RN7SL283P* ($P=2.56 \times 10^{-3}$). Among
344 the top 15 most significant genes, there were strong peaks on chromosome 6 (*HLA-*
345 *DRB5*, *BTNL2*, *RNU1-61P*) and chromosome 17 (*LRRC59*, *EME1*, *ACSF2*; **Table S4**)
346 although these associations did not reach gene-wide significance.

347

348 Immune-related Pathways Linked to HDP

349 Among the ten most significant KEGG gene sets associated with HDP were the pentose
350 phosphate pathway and the hedgehog signaling pathway (**Figure 3**). Gene-set analyses
351 revealed additional immune-related diseases and processes related to HDP-associated
352 genes (**Table S5**), but these sets, which included autoimmune thyroid disease, graft-

353 versus-host disease, adipocytokine signaling pathway, and allograft rejection
354 (uncorrected $P < 0.05$), did not meet the FDR-corrected significance threshold.

355

356 Polygenic Score for Systolic Blood Pressure is Associated with HDP

357 We tested the association of HDP with PGS for SBP, PE, GH with HDP, and essential
358 hypertension in the PEGS female sub-cohort. **Table S6** shows the number of variants
359 missing in the PEGS cohort for each PGS. In the overall model, all four PGS were
360 significantly associated with HDP ($P < 0.05$). The AUC (0.57; 95% CI: (0.52-0.61); **Figure**
361 **4; Figure S5**) and pseudo- R^2 (0.7%; **Table S7**) were the highest for the SBP PGS [1]
362 based on the UKBB. One standard deviation increase in the SBP PGS [1] was
363 associated with a 2.09-fold increase in HDP odds (95% CI: 1.40, 3.13). GH-PGS had
364 the highest OR for HDP (**Figure 4**). **Figure S6** shows HDP prevalence by percentile for
365 each PGS. When the results were stratified by genetic similarity group, the four PGS
366 were not significantly associated with HDP in the AFR genetic similarity group, but there
367 were similar effect sizes for the overall models and for the EUR genetic similarity group
368 (**Figure 4**).

369

370 For essential hypertension, the two SBP PGS were strongly associated with HDP in
371 both the overall and stratified analyses (**Figure S4**). The pseudo- R^2 for SBP PGS [2]
372 was the highest at 5.6% (**Table S7**). The receiver operator characteristic curves indicate
373 that SBP PGS [2] is the most predictive of essential hypertension in PEGS (**Figure S7**),
374 evidenced by the clear linear trend in the percentile plot (**Figure S8**).

375

376 **Discussion**

377 Our GWAS is the first to identify and validate genetic variants that are significantly
378 associated with HDP in the maternal genome in *RARB* on chromosome 3. By
379 leveraging a multi-ancestry study design, we increased statistical power for genetic
380 discovery, which enabled us to confirm our findings in the clinically phenotyped UKBB
381 cohort.

382

383 The two genome-wide significant loci in PEGS were also significantly associated with
384 HDP in the gene-level analysis, where *RARB* and *RN7SL283P* were the top two genes.
385 Retinoic acid receptors are essential for healthy placental and fetal development. The
386 receptor binds retinoic acid, which is the biologically active form of vitamin A that
387 facilitates cellular signaling in embryonic morphogenesis, cell proliferation,
388 differentiation, and apoptosis^{45–47}. Maternal retinol is transferred across the placenta to
389 embryos⁴⁷. All-trans-retinoic acid plays a significant role in maintaining immune balance
390 in some tissues and is linked to the etiology of PE, where greater dysregulation in
391 retinoic acid levels in the decidua (maternal endometrium) is observed in PE cases
392 compared to those with a normal pregnancy⁴⁸. A link has been suggested between
393 decidualization and the anti-angiogenic protein soluble fms-like tyrosine kinase-1
394 (sFLT1), with retinoic acid directly reducing sFLT1 production in the decidua, which
395 affects early vascular development⁴⁸. Furthermore, in GWAS in a Peruvian cohort
396 focused on maternal and offspring genetic effects on PE, rs4241542, a variant that is 80
397 kb from the lead variant rs61176331, was associated with proteinuria in offspring
398 ($P=8.48 \times 10^{-6}$)⁴⁹, further supporting the biological validity of this region in *RARB* for

399 HDP. Future research is needed to investigate the interaction at the maternal-fetal
400 interface, including the influence of gene-gene interaction in placental pathophysiology.

401

402 While the variant near *LRP1B* was not validated in UKBB, rs114954125 is near
403 *RN7SL283P*, which is a pseudogene. *LRP1B* belongs to the LDL receptor family and is
404 a tumor suppressor gene⁵⁰. In a previous candidate SNP study, the T-allele variant in
405 *LRP1B*, rs35821928, was associated with increased maternal height and decreased
406 maternal weight⁵¹.

407

408 While the MAGMA gene-set analysis did not yield pathways that met the stringent FDR-
409 corrected significance threshold, the top ten most significant gene sets provide
410 intriguing insights into the probable molecular mechanisms of HDP. The pentose
411 phosphate pathway is among these notable gene sets. This metabolic pathway is
412 involved in cellular redox homeostasis, which is related to oxidative stress as a factor in
413 HDP^{11,12}. The identification of immune-response pathways and conditions such as
414 autoimmune thyroid disease, graft-versus-host disease, and allograft rejection points to
415 an immunological dysregulation component of HDP. This is consistent with a growing
416 body of research focused on the immunological basis of HDP^{52,19,22,48}.

417

418 The gene-level analysis revealed a strong peak of genes at chromosome 6 that did not
419 meet the gene-wide significance threshold. *HLA-DRB5* and *BTNL2* belong to the human
420 major histocompatibility (HLA) complex, where they are involved in the generation of an
421 immune response. Additionally, the adipocytokine signaling pathway connects to

422 signaling molecules originating from adipose tissue, which has been linked to various
423 aspects of cardiovascular health. This could indicate adipose tissue malfunction plays a
424 role in HDP and other cardiometabolic disorders of pregnancy such as gestational
425 diabetes⁵³. These suggestive findings related to immune response, inflammation, and
426 metabolic dysfunction call for further analysis and interpretation in large pregnancy
427 cohorts.

428
429 The differing patterns of suggestive associations in the GWAS of the composite HDP
430 outcome, GH, and PE in PEGS may indicate differences in etiology (**Figure S3**). The
431 heterogeneous nature of HDP has led to the understanding that there are several
432 distinct PE subtypes. A two-stage placental model related to the development of PE^{19,54}
433 demonstrated that an early-onset, more severe subtype may be associated with
434 underlying genetic or environmental factors that result in abnormal placentation. In
435 contrast, a later-onset, less severe subtype may be a consequence of factors such as
436 obesity, diabetes, cardiovascular disorders, or multi-fetal pregnancy¹. Further, risk
437 factors for GH and PE vary⁵⁵, supporting evidence for differing etiology across the
438 collection of hypertensive disorders. However, a holistic assessment of HDP that
439 combined PE with other maternal hypertensive disorders as a phenotype for GWAS
440 yielded novel, overlapping risk loci identified in models of PE only and PE or other
441 HDP⁵⁶.

442
443 Considering polygenic prediction, using a blood pressure PGS to predict HDP has been
444 found to be more effective than using a PE PGS for women who are non-hypertensive

445 before pregnancy⁵⁷. In this study, we assessed the predictive utility of four PGS in the
446 context of HDP and essential hypertension. While these scores were developed in the
447 White British sub-cohort of UKBB^{28,44}, our results demonstrate that an SBP PGS is a
448 strong predictor of HDP. In the stratified results, none of the scores exhibited predictive
449 value for HDP for the AFR genetic similarity group, and only the SBP scores were
450 significantly associated with essential hypertension. These findings highlight the need
451 for customizing PGS development to be inclusive of varying genetic similarity
452 backgrounds to improve the accuracy of HDP risk assessment, as the study's
453 effectiveness in predicting HDP using a PE PGS or GH PGS does not apply to all
454 populations. Furthermore, while the SBP PGS showed promise for HDP prediction and
455 prevention, it is worth investigating how these genetic scores could be combined with
456 other clinical and environmental factors to develop a more holistic approach to HDP
457 prediction and prevention.

458
459 This study has several limitations. First, a composite outcome like HDP can obscure
460 disparate etiologies, for example, between PE and GH. Because of the limited sample
461 size, we grouped the heterogeneous conditions that comprise HDP to increase power to
462 detect genetic variant associations. Second, the results are based on data from a
463 particular time point for each participant. The lack of longitudinal data does not allow us
464 to account for the continuum of risk due to factors such as maternal age and gestational
465 age at the onset of HDP. Further, the study design does not consider the recurrence of
466 complications in subsequent pregnancies. Third, the outcomes are dichotomized due to
467 the nature of the questionnaires completed, which may result in the oversimplification of

468 adverse pregnancy outcomes. Fourth, misclassification bias may be present, since
469 participants who enrolled in PEGS before a pregnancy complication occurred were
470 included in the control group. Fourth, there is a lack of harmonization of phenotypes
471 between study cohorts. While PEGS data are limited, the UKBB uses a wide variety of
472 diagnosis codes for phenotyping, and thus not all pregnancy-related adverse outcomes
473 were excluded from the control group.

474

475 **Perspectives (max 250 words - Conclusion)**

476 Variants near *RARB* were significantly associated with HDP in a multi-ancestry cohort.
477 We validated this association in a clinically phenotyped cohort. In gene-set analyses,
478 probable mechanisms associated with HDP included inflammatory and immunological
479 pathways. These findings are supported by recent research focused on the relationship
480 between retinoic acid dysregulation and PE. In-depth functional genomics investigations
481 can elucidate the molecular pathways by which genetic variations near *RARB* contribute
482 to the etiology of HDP, such as the influence of these variants on *RARB* expression,
483 epigenetic alterations, and downstream signaling pathways within the placenta and
484 maternal vascular system. The study demonstrates the strong predictive value of SBP
485 PGS for HDP, but further work should investigate approaches to improve the accuracy
486 and inclusivity of risk assessment for HDP.

487

488 HDP is a complex and challenging area of obstetrics, with a significant, disproportionate
489 impact on maternal and fetal health across populations. While we focus on the genetic
490 contribution to HDP, examining the role of social determinants of health is also

491 important, in part because environmental factors can influence genetic risk and the
492 development of HDP. Addressing social, lifestyle, and genetic factors is imperative to
493 reduce the burden of HDP and improve maternal and fetal outcomes. Further research
494 in large, diverse pregnancy cohorts is needed to fully understand the underlying
495 mechanisms and develop effective preventative and therapeutic strategies.

496

497 **Novelty and Relevance (max 100 words)**

498 *What is new?*

- 499 • Genetic studies of hypertensive disorders of pregnancy (HDP) have
500 predominantly focused on individuals of European ancestry.
- 501 • Our multi-ancestry genome-wide association study is the first to identify and
502 validate genetic variants that are significantly associated with HDP in the
503 maternal genome in the *RARB* gene.

504 *What is relevant?*

- 505 • We identified two novel genome-wide significant associations with HDP in the
506 maternal genome, namely rs114954125 on chromosome 2 and rs61176331 on
507 chromosome 3. The variant rs61176331 was validated in the UK Biobank. Gene-
508 set analyses revealed a suggestive association between HDP and
509 inflammatory/immunological pathways such as the pentose phosphate pathway
510 and autoimmune thyroid disease.

511

512 *Clinical/Pathophysiological Implications?*

- 513 • Retinoic acid, which binds to *RARB*, inhibits the expression of sFLT1 in decidua.
514 The effects of genetic variants near *RARB* on the pathophysiology of HDP could
515 provide insights into the underlying molecular mechanisms.
- 516 • Existing genetic scores for SBP, PE, and GH were predictive of HDP and
517 essential hypertension in the overall PEGS cohort but were not predictive of HDP
518 for those of predominantly African ancestry. Future directions include the
519 development of genetic risk prediction modeling inclusive of large, ancestrally
520 diverse populations to improve utility.

521

522 **Acknowledgements**

523 We would like to acknowledge the participants in the PEGS and UK Biobank cohorts for
524 making this work possible. We would also like to thank Hannah Collins Cakar for
525 assistance with manuscript preparation.

526

527 **Sources of Funding**

528 This study was supported by the National Institutes of Health Intramural Research
529 Program and the NIH-Oxford-Cambridge Scholars program.

530

531 **Disclosures**

532 None

533

534

535

536 **Supplemental Material**

537 Supplemental Methods

538 Tables S1–S7

539 Figures S1-S8

540 References #58-#90

541

542 **References**

- 543 1. Hutcheon JA, Lisonkova S, Joseph KS. Epidemiology of pre-eclampsia and the
544 other hypertensive disorders of pregnancy. *Best Pract. Res. Clin. Obstet.*
545 *Gynaecol.* 2011;25:391–403.
- 546 2. Khan KS, Wojdyla D, Say L, Gülmezoglu AM, Van Look PF. WHO analysis of
547 causes of maternal death: a systematic review. *Lancet Lond. Engl.*
548 2006;367:1066–1074.
- 549 3. Sole KB, Staff AC, Räisänen S, Laine K. Substantial decrease in preeclampsia
550 prevalence and risk over two decades: A population-based study of 1,153,227
551 deliveries in Norway. *Pregnancy Hypertens.* 2022;28:21–27.
- 552 4. Wang W, Xie X, Yuan T, Wang Y, Zhao F, Zhou Z, Zhang H. Epidemiological trends
553 of maternal hypertensive disorders of pregnancy at the global, regional, and
554 national levels: a population-based study. *BMC Pregnancy Childbirth.*
555 2021;21:364.
- 556 5. Ghulmiyyah L, Sibai B. Maternal mortality from preeclampsia/eclampsia. *Semin.*
557 *Perinatol.* 2012;36:56–59.
- 558 6. GBD 2015 Maternal Mortality Collaborators. Global, regional, and national levels of
559 maternal mortality, 1990-2015: a systematic analysis for the Global Burden of
560 Disease Study 2015. *Lancet.* 2016;388:1775–1812.
- 561 7. Fasanya HO, Hsiao CJ, Armstrong-Sylvester KR, Beal SG. A critical review on the
562 use of race in understanding racial disparities in preeclampsia. *J. Appl. Lab. Med.*
563 2021;6:247–256.
- 564 8. Johnson JD, Louis JM. Does race or ethnicity play a role in the origin,
565 pathophysiology, and outcomes of preeclampsia? An expert review of the
566 literature. *Am. J. Obstet. Gynecol.* 2022;226:S876–S885.

- 567 9. Bartsch E, Medcalf KE, Park AL, Ray JG. Clinical risk factors for pre-eclampsia
568 determined in early pregnancy: systematic review and meta-analysis of large
569 cohort studies. *BMJ*. 2016;353:i1753.
- 570 10. Braunthal S, Brateanu A. Hypertension in pregnancy: Pathophysiology and
571 treatment. *SAGE Open Med*. 2019;7:2050312119843700.
- 572 11. Gilbert JS, Ryan MJ, LaMarca BB, Sedeek M, Murphy SR, Granger JP.
573 Pathophysiology of hypertension during preeclampsia: linking placental ischemia
574 with endothelial dysfunction. *Am. J. Physiol.-Heart Circ. Physiol*. 2008;294:H541–
575 H550.
- 576 12. LaMarca B. Endothelial dysfunction. An important mediator in the pathophysiology
577 of hypertension during preeclampsia. *Minerva Ginecol*. 2012;64:309–320.
- 578 13. Braunthal S, Brateanu A. Hypertension in pregnancy: pathophysiology and
579 treatment. *SAGE Open Med*. 2019;7:2050312119843700.
- 580 14. Lyall F, Bulmer JN, Kelly H, Duffie E, Robson SC. Human trophoblast invasion
581 and spiral artery transformation. *Am. J. Pathol*. 1999;154:1105–1114.
- 582 15. Roberts JM, Rich-Edwards JW, McElrath TF, Garmire L, Myatt L, Global
583 Pregnancy Collaboration. Subtypes of preeclampsia: recognition and determining
584 clinical usefulness. *Hypertension*. 2021;77:1430–1441.
- 585 16. Lu H-Q, Hu R. The role of immunity in the pathogenesis and development of pre-
586 eclampsia. *Scand. J. Immunol*. 2019;90:e12756.
- 587 17. Staff AC, Fjeldstad HE, Fosheim IK, Moe K, Turowski G, Johnsen GM, Alnaes-
588 Katjavivi P, Sugulle M. Failure of physiological transformation and spiral artery
589 atherosclerosis: their roles in preeclampsia. *Am. J. Obstet. Gynecol*. 2022;226:S895–
590 S906.
- 591 18. Harmon AC, Cornelius DC, Amaral LM, Faulkner JL, Cunningham MW Jr, Wallace
592 K, LaMarca B. The role of inflammation in the pathology of preeclampsia. *Clin.*
593 *Sci*. 2016;130:409–419.
- 594 19. LaMarca B. The role of immune activation in contributing to vascular dysfunction
595 and the pathophysiology of hypertension during preeclampsia. *Minerva Ginecol*.
596 2010;62:105–120.
- 597 20. Aneman I, Pienaar D, Suvakov S, Simic TP, Garovic VD, McClements L.
598 Mechanisms of key innate immune cells in early- and late-onset preeclampsia.
599 *Front. Immunol*. 2020;11:1864.
- 600 21. Smith SD, Dunk CE, Aplin JD, Harris LK, Jones RL. Evidence for immune cell
601 involvement in decidual spiral arteriole remodeling in early human pregnancy. *Am.*
602 *J. Pathol*. 2009;174:1959–1971.

- 603 22. LaMarca B, Cornelius D, Wallace K. Elucidating immune mechanisms causing
604 hypertension during pregnancy. *Physiology*. 2013;28:225–233.
- 605 23. Gray KJ, Kovacheva VP, Mirzakhani H, Bjornes AC, Almoguera B, DeWan AT,
606 Triche EW, Saftlas AF, Hoh J, Bodian DL, et al. Gene-centric analysis of
607 preeclampsia identifies maternal association at PLEKHG1. *Hypertension*.
608 2018;72:408–416.
- 609 24. McGinnis R, Steinthorsdottir V, Williams NO, Thorleifsson G, Shooter S,
610 Hjartardottir S, Bumpstead S, Stefansdottir L, Hildyard L, Sigurdsson JK, et al.
611 Variants in the fetal genome near FLT1 are associated with risk of preeclampsia.
612 *Nat. Genet*. 2017;49:1255–1260.
- 613 25. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A,
614 Vukcevic D, Delaneau O, O’Connell J, et al. The UK Biobank resource with deep
615 phenotyping and genomic data. *Nature*. 2018;562:203–209.
- 616 26. Mills MC, Rahal C. A scientometric review of genome-wide association studies.
617 *Commun. Biol*. 2019;2:1–11.
- 618 27. Peterson RE, Kuchenbaecker K, Walters RK, Chen C-Y, Popejoy AB, Periyasamy
619 S, Lam M, Iyegbe C, Strawbridge RJ, Brick L, et al. Genome-wide association
620 studies in ancestrally diverse populations: opportunities, methods, pitfalls, and
621 recommendations. *Cell*. 2019;179:589–603.
- 622 28. Honigberg MC, Truong B, Khan RR, Xiao B, Bhatta L, Vy HMT, Guerrero RF,
623 Schuermans A, Selvaraj MS, Patel AP, et al. Polygenic prediction of preeclampsia
624 and gestational hypertension. *Nat. Med*. 2023;29:1540–1549.
- 625 29. Pan UKBB | Pan UKBB [Internet]. [cited 2021 Mar 30]; Available from:
626 <https://pan.ukbb.broadinstitute.org/>
- 627 30. National Institute of Environmental Health Sciences (NIEHS). Personalized
628 Environment and Genes Study [Internet]. clinicaltrials.gov; 2022 [cited 2022 Sep
629 1]. Available from: <https://clinicaltrials.gov/ct2/show/NCT00341237>
- 630 31. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen W-M. Robust
631 relationship inference in genome-wide association studies. *Bioinformatics*.
632 2010;26:2867–2873.
- 633 32. Maples BK, Gravel S, Kenny EE, Bustamante CD. RFMix: a discriminative
634 modeling approach for rapid and robust local-ancestry inference. *Am. J. Hum.*
635 *Genet*. 2013;93:278–288.
- 636 33. Diaz-Papkovich A, Anderson-Trocmé L, Ben-Eghan C, Gravel S. UMAP reveals
637 cryptic population structure and phenotype heterogeneity in large genomic
638 cohorts. *PLOS Genet*. 2019;15:e1008432.

- 639 34. : Return 2442 [Internet]. [cited 2023 Jul 19]; Available from:
640 <https://biobank.ndph.ox.ac.uk/ukb/dset.cgi?id=2442>
- 641 35. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-
642 generation PLINK: rising to the challenge of larger and richer datasets.
643 *GigaScience* [Internet]. 2015 [cited 2021 Aug 30];4. Available from:
644 <https://doi.org/10.1186/s13742-015-0047-8>
- 645 36. FIRTH D. Bias reduction of maximum likelihood estimates. *Biometrika*.
646 1993;80:27–38.
- 647 37. Yang J, Ferreira T, Morris AP, Medland SE, Genetic Investigation of
648 ANthropometric Traits (GIANT) Consortium, DIAbetes Genetics Replication And
649 Meta-analysis (DIAGRAM) Consortium, Madden PAF, Heath AC, Martin NG,
650 Montgomery GW, et al. Conditional and joint multiple-SNP analysis of GWAS
651 summary statistics identifies additional variants influencing complex traits. *Nat.*
652 *Genet.* 2012;44:369–375, S1-3.
- 653 38. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of
654 genomewide association scans. *Bioinformatics*. 2010;26:2190–2191.
- 655 39. Aken BL, Ayling S, Barrell D, Clarke L, Curwen V, Fairley S, Fernandez Banet J,
656 Billis K, García Girón C, Hourlier T, et al. The Ensembl gene annotation system.
657 *Database J. Biol. Databases Curation*. 2016;2016:baw093.
- 658 40. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set
659 analysis of GWAS data. *PLoS Comput. Biol.* 2015;11:e1004219.
- 660 41. Zhang S. A simplified protocol for performing MAGMA/H-MAGMA gene set
661 analysis utilizing high-performance computing environments. *STAR Protoc.*
662 2022;3:101083.
- 663 42. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes.
664 *Nucleic Acids Res.* 2000;28:27–30.
- 665 43. Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG:
666 integrating viruses and cellular organisms. *Nucleic Acids Res.* 2021;49:D545–
667 D551.
- 668 44. Kauko A, Aittokallio J, Vaura F, Ji H, Ebinger JE, Niiranen T, Cheng S. Sex
669 differences in genetic risk for hypertension. *Hypertension*. 2021;78:1153–1155.
- 670 45. Huebner H, Hartner A, Rascher W, Strick RR, Kehl S, Heindl F, Wachter DL,
671 Beckmann MW, Fahlbusch FB, Ruebner M. Expression and regulation of retinoic
672 acid receptor responders in the human placenta. *Reprod. Sci.* 2018;25:1357–
673 1370.

- 674 46. RARB retinoic acid receptor beta [Homo sapiens (human)] - Gene - NCBI
675 [Internet]. [cited 2022 Sep 24]; Available from:
676 <https://www.ncbi.nlm.nih.gov/gene/5915#summary>
- 677 47. Rhinn M, Dollé P. Retinoic acid signalling during development. *Development*.
678 2012;139:843–858.
- 679 48. Rajakumar A, Kane MA, Yu J, Taylor RN, Sidell N. Aberrant retinoic acid
680 production in the decidua: Implications for pre-eclampsia. *J. Obstet. Gynaecol.*
681 *Res.* 2020;46:1007–1016.
- 682 49. Nieves-Colón MA, Badillo Rivera KM, Sandoval K, Villanueva Dávalos V,
683 Enriquez Lencinas LE, Mendoza-Revilla J, Adhikari K, González-Buenfil R, Chen
684 JW, Zhang ET, et al. Clotting factor genes are associated with preeclampsia in
685 high-altitude pregnant women in the Peruvian Andes. *Am. J. Hum. Genet.*
686 2022;109:1117–1139.
- 687 50. Liu X, White S, Peng B, Johnson AD, Brody JA, Li AH, Huang Z, Carroll A, Wei P,
688 Gibbs R, et al. WGS: an annotation pipeline for human genome sequencing
689 studies. *J. Med. Genet.* 2016;53:111–112.
- 690 51. Løset M, Johnson MP, Melton PE, Ang W, Huang R-C, Mori TA, Beilin LJ, Pennell
691 C, Roten LT, Iversen A-C, et al. Preeclampsia and cardiovascular disease share
692 genetic risk factors on chromosome 2q22. *Pregnancy Hypertens. Int. J. Womens*
693 *Cardiovasc. Health.* 2014;4:178–185.
- 694 52. LaMarca B. Endothelial dysfunction; an important mediator in the Pathophysiology
695 of Hypertension during Preeclampsia. *Minerva Ginecol.* 2012;64:309–320.
- 696 53. Trivett C, Lees ZJ, Freeman DJ. Adipose tissue function in healthy pregnancy,
697 gestational diabetes mellitus and pre-eclampsia. *Eur. J. Clin. Nutr.* 2021;75:1745–
698 1756.
- 699 54. Staff AC. The two-stage placental model of preeclampsia: An update. *J. Reprod.*
700 *Immunol.* 2019;134–135:1–10.
- 701 55. Li X, Tan H, Huang X, Zhou S, Hu S, Wang X, Xu X, Liu Q, Wen SW. Similarities
702 and differences between the risk factors for gestational hypertension and
703 preeclampsia: A population based cohort study in south China. *Pregnancy*
704 *Hypertens.* 2016;6:66–71.
- 705 56. Tyrmi JS, Kaartokallio T, Lokki AI, Jääskeläinen T, Kortelainen E, Ruotsalainen S,
706 Karjalainen J, Ripatti S, Kivioja A, Laisk T, et al. Genetic risk factors associated
707 with preeclampsia and hypertensive disorders of pregnancy. *JAMA Cardiol.*
708 2023;8:674–683.
- 709 57. Nurkkala J, Kauko A, FinnGen, Laivuori H, Saarela T, Tyrmi JS, Vaura F, Cheng
710 S, Bello NA, Aittokallio J, et al. Associations of polygenic risk scores for

- 711 preeclampsia and blood pressure with hypertensive disorders of pregnancy. *J.*
712 *Hypertens.* 2023;41:380.
- 713 58. gatk4-genome-processing-pipeline [Internet]. 2022 [cited 2022 Sep 6];Available
714 from: <https://github.com/gatk-workflows/gatk4-genome-processing-pipeline>
- 715 59. Auwera GAV der, O'Connor BD. Genomics in the cloud: Using Docker, GATK, and
716 WDL in Terra. O'Reilly Media, Inc.; 2020.
- 717 60. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-
718 MEM [Internet]. 2013 [cited 2022 Sep 6];Available from:
719 <http://arxiv.org/abs/1303.3997>
- 720 61. Picard Tools - By Broad Institute [Internet]. [cited 2022 Sep 6];Available from:
721 <https://broadinstitute.github.io/picard/>
- 722 62. applybqsr — Clara Parabricks v3.7 documentation [Internet]. [cited 2022 Sep
723 6];Available from:
724 https://docs.nvidia.com/clara/parabricks/3.7.0/Documentation/ToolDocs/man_applybqsr.html
725
- 726 63. DeepVariant [Internet]. 2022 [cited 2022 Sep 6];Available from:
727 <https://github.com/google/deepvariant>
- 728 64. Liu X, White S, Peng B, Johnson AD, Brody JA, Li AH, Huang Z, Carroll A, Wei P,
729 Gibbs R, et al. WGS: an annotation pipeline for human genome sequencing
730 studies. *J. Med. Genet.* 2016;53:111–112.
- 731 65. Aken BL, Ayling S, Barrell D, Clarke L, Curwen V, Fairley S, Fernandez Banet J,
732 Billis K, García Girón C, Hourlier T, et al. The Ensembl gene annotation system.
733 *Database J. Biol. Databases Curation.* 2016;2016:baw093.
- 734 66. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen W-M. Robust
735 relationship inference in genome-wide association studies. *Bioinformatics.*
736 2010;26:2867–2873.
- 737 67. Population Structure and Relatedness Inference using the GENESIS Package
738 [Internet]. [cited 2021 Feb 28];Available from:
739 <https://www.bioconductor.org/packages/devel/bioc/vignettes/GENESIS/inst/doc/pca.html>
740
- 741 68. Conomos MP, Miller MB, Thornton TA. Robust inference of population structure for
742 ancestry prediction and correction of stratification in the presence of relatedness.
743 *Genet. Epidemiol.* 2015;39:276–293.
- 744 69. Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS. A high-performance
745 computing toolset for relatedness and principal component analysis of SNP data.
746 *Bioinformatics.* 2012;28:3326–3328.

- 747 70. Delaneau O, Zagury J-F, Robinson MR, Marchini JL, Dermitzakis ET. Accurate,
748 scalable and integrative haplotype estimation. *Nat. Commun.* 2019;10:5436.
- 749 71. Maples BK, Gravel S, Kenny EE, Bustamante CD. RFMix: a discriminative modeling
750 approach for rapid and robust local-ancestry inference. *Am. J. Hum. Genet.*
751 2013;93:278–288.
- 752 72. Byrska-Bishop M, Evani US, Zhao X, Basile AO, Abel HJ, Regier AA, Corvelo A,
753 Clarke WE, Musunuri R, Nagulapalli K, et al. High-coverage whole-genome
754 sequencing of the expanded 1000 Genomes Project cohort including 602 trios.
755 *Cell.* 2022;185:3426-3440.e19.
- 756 73. IGSR | samples [Internet]. [cited 2022 May 9];Available from:
757 <https://www.internationalgenome.org/data-portal/sample>
- 758 74. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic
759 D, Delaneau O, O’Connell J, et al. The UK Biobank resource with deep
760 phenotyping and genomic data. *Nature.* 2018;562:203–209.
- 761 75. : Category 100319 [Internet]. [cited 2022 Sep 9];Available from:
762 <https://biobank.ndph.ox.ac.uk/ukb/label.cgi?id=100319>
- 763 76. : Resource 1671 [Internet]. [cited 2022 Sep 9];Available from:
764 <https://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=1671>
- 765 77. : Return 2442 [Internet]. [cited 2023 Jul 19];Available from:
766 <https://biobank.ndph.ox.ac.uk/ukb/dset.cgi?id=2442>
- 767 78. Quality Control (QC) | Pan UKBB [Internet]. [cited 2021 Apr 9];Available from:
768 <https://pan-dev.ukbb.broadinstitute.org/docs/qc>
- 769 79. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation
770 PLINK: rising to the challenge of larger and richer datasets. *GigaScience*
771 [Internet]. 2015 [cited 2021 Aug 30];4. Available from:
772 <https://doi.org/10.1186/s13742-015-0047-8>
- 773 80. Yang J, Ferreira T, Morris AP, Medland SE, Genetic Investigation of ANthropometric
774 Traits (GIANT) Consortium, DIAbetes Genetics Replication And Meta-analysis
775 (DIAGRAM) Consortium, Madden PAF, Heath AC, Martin NG, Montgomery GW,
776 et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics
777 identifies additional variants influencing complex traits. *Nat. Genet.* 2012;44:369–
778 375, S1-3.
- 779 81. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of
780 genomewide association scans. *Bioinformatics.* 2010;26:2190–2191.

- 781 82. Durinck S, Moreau Y, Kasprzyk A, Davis S, De Moor B, Brazma A, Huber W.
782 BioMart and Bioconductor: a powerful link between biological databases and
783 microarray data analysis. *Bioinforma. Oxf. Engl.* 2005;21:3439–3440.
- 784 83. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration
785 of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.*
786 2009;4:1184–1191.
- 787 84. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set
788 analysis of GWAS data. *PLoS Comput. Biol.* 2015;11:e1004219.
- 789 85. Zhang S. A simplified protocol for performing MAGMA/H-MAGMA gene set analysis
790 utilizing high-performance computing environments. *STAR Protoc.*
791 2022;3:101083.
- 792 86. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The
793 Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.*
794 2015;1:417–425.
- 795 87. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic*
796 *Acids Res.* 2000;28:27–30.
- 797 88. Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG:
798 integrating viruses and cellular organisms. *Nucleic Acids Res.* 2021;49:D545–
799 D551.
- 800 89. Kauko A, Aittokallio J, Vaura F, Ji H, Ebinger JE, Niiranen T, Cheng S. Sex
801 differences in genetic risk for hypertension. *Hypertension.* 2021;78:1153–1155.
- 802 90. Honigberg MC, Truong B, Khan RR, Xiao B, Bhatta L, Vy HMT, Guerrero RF,
803 Schuermans A, Selvaraj MS, Patel AP, et al. Polygenic prediction of preeclampsia
804 and gestational hypertension. *Nat. Med.* 2023;29:1540–1549.

805

806

807

808

809

810

811

812

813

814

815 **Tables**

PEGS	Study Sample n = 1,771	HDP Cases n = 202	Controls n = 1,569	33 P value
Hypertensive Disorders of Pregnancy, n (%)				
Gestational hypertension	202 (11.4)	202 (100.0)	0 (0.0)	<0.001*
Number of pregnancies with gestational hypertension				
1	97 (5.5)	97 (48.0)	0 (0.0)	
2	24 (1.4)	24 (11.9)	0 (0.0)	
3 or more	14 (0.9)	14 (6.9)	0 (0.0)	<0.001*
Missing	1636 (92.4)	67 (33.2)	1569 (100.0)	
Gestational hypertension without preeclampsia	78 (4.4)	78 (38.6)	0 (0.0)	<0.001*
Preeclampsia	71 (4.0)	71 (35.2)	0 (0.0)	
Number of pregnancies with preeclampsia complication				
1	58 (3.3)	58 (28.7)	0 (0.0)	
2	4 (0.2)	4 (2.0)	0 (0.0)	
3 or more	5 (0.3)	5 (2.5)	0 (0.0)	<0.001*
Missing	1704 (96.2)	135 (66.8)	1569 (100.0)	
Gestational hypertension and/or preeclampsia	202 (11.4)	202 (100.0)	0 (0.0)	<0.001*
Covariates				
Age at enrollment, mean (SD)	47.2 (13.6)	44.8 (12.7)	47.5 (13.7)	0.0071†
BMI at enrollment, median (IQR)	27.4 (23.4, 32.6)	28.7 (25.0, 34.9)	27.3 (23.2, 32.3)	<0.001‡
Gravidity, median (IQR)	3.0 (2.0, 4.0)	4.0 (3.0, 5.0)	3.0 (2.0, 4.0)	<0.001‡
Race/ethnicity, n(%)				
American Indian/Alaskan Native	16 (0.9)	1 (0.5)	15 (1.0)	
Asian	39 (2.2)	3 (1.5)	36 (2.3)	
Black	358 (20.2)	45 (22.3)	313 (20.0)	
Hispanic/Latino	47 (2.7)	3 (1.5)	44 (2.8)	0.1878*
White	1256 (70.9)	138 (68.3)	1118 (71.3)	
Other (Multiple or Native Hawaiian/Pacific Islander)	22 (1.2)	5 (2.5)	17 (1.1)	
Unknown	33 (1.9)	7 (3.5)	26 (1.7)	

Predominant genetic similarity group				
African	382 (21.6)	49 (24.3)	333 (21.2)	0.4288*
Admixed American	8 (0.5)	0 (0.0)	8 (0.5)	
East Asian	25 (1.4)	1 (0.5)	24 (1.5)	
European	1328 (75.0)	150 (74.3)	1178 (75.1)	
South Asian	15 (0.8)	2 (1.0)	13 (0.8)	
No group	13 (0.7)	0 (0.0)	13 (0.8)	
Socioeconomic Characteristics, n (%)				
Educational Level at Enrollment				
High school education or less	183 (10.3)	14 (6.9)	169 (10.8)	0.2900*
Some college or technical/vocational school	549 (31.0)	73 (36.1)	476 (30.3)	
Bachelor's degree	530 (29.9)	57 (28.2)	473 (30.2)	
Graduate or professional degree	495 (28.0)	57 (28.2)	438 (27.9)	
Missing	14 (0.8)	1 (0.5)	13 (0.8)	
Income at Enrolment (USD)				
Less than \$50,000	659 (37.2)	69 (34.2)	590 (37.6)	0.1810*
\$50,000 - \$79,999	441 (24.9)	62 (30.7)	379 (24.2)	
\$80,000 or more	603 (34.0)	66 (32.7)	537 (34.2)	
Missing	68 (3.8)	5 (2.5)	63 (4.0)	

Table 1. Demographic characteristics of participants in the PEGS discovery cohort.

* The *P* values were computed by Fisher's exact test.

† The *P* value was computed by t-test.

‡ The *P* values were computed by performing the Mann-Whitney U test.

816

817

818

819

820

821

822

823

Multi-ancestry GWAS									
(202 cases/1,569 controls)									
Joint analysis in GWAS									
Nearest Gene	Lead Variant	Chr:Position	Effect Allele/		EAF	OR	95% CI	P	P
			Other Allele						
<i>LRP1B</i>	rs114954125	2:139978665	G/A		0.051	3.03	2.05-4.49	3.19E-08	3.79E-08
<i>RARB</i>	rs61176331	3:24930847	A/C		0.048	3.09	2.11-4.53	7.97E-09	9.56E-09
Genetic Similarity Group for African Descent									
(49 cases/333 controls)									
Nearest Gene	Lead Variant	Chr:Position	Effect Allele/		EAF	OR	95% CI	P	
			Other Allele						
<i>RARB</i>	rs61176331	3:24930847	A/C		0.042	6.98	2.97-16.42	8.32E-06	
Genetic Similarity Group for European Descent									
(150 cases/1,178 controls)									
Nearest Gene	Lead Variant	Chr:Position	Effect Allele/		EAF	OR	95% CI	P	
			Other Allele						
<i>LRP1B</i>	rs114954125	2:139978665	G/A		0.067	2.95	1.97-4.42	1.47E-07	
<i>RARB</i>	rs61176331	3:24930847	A/C		0.051	2.37	1.51-3.74	1.92E-04	
Validation									
(1,319 cases/2,050 controls)									

Nearest Gene	Lead Variant	Chr:Position	Effect Allele/		OR	95% CI	P
			Other Allele	EAF			
<i>LRP1B</i>	rs114954125	2:139978665	G/A	0.061	1.02	0.83-1.27	8.34E-01
<i>RARB</i>	rs61176331	3:24930847	A/C	0.045	1.29	1.02-1.65	3.73E-02

GWAS + Validation
(1,521 cases/3,619 controls)

Nearest Gene	Lead Variant	Chr:Position	Effect Allele/		OR	95% CI	P
			Other Allele	EAF			
<i>LRP1B</i>	rs114954125	2:139978665	G/A	0.059	1.31	1.09-1.58	4.64E-03
<i>RARB</i>	rs61176331	3:24930847	A/C	0.046	1.66	1.35-2.04	1.27E-06

Table 2. Summary statistics for rs114954125 (*LRP1B*) and rs61176331 (*RARB*). Chr = chromosome, EAF = effect allele frequency

825 **Figures with Figure Legends**

826

827

828

829

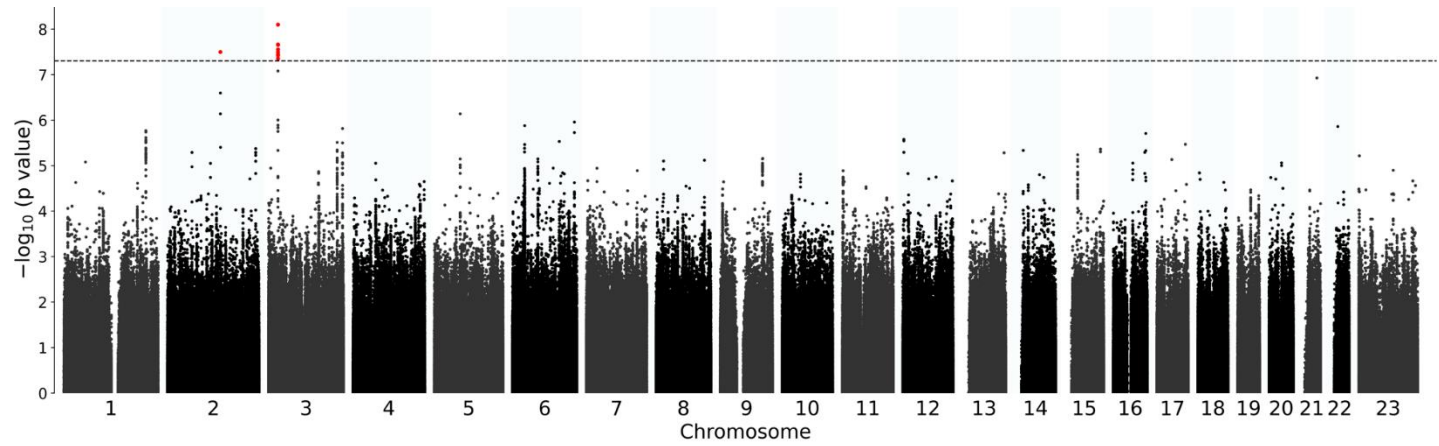
830

831

832

833

834



835

836

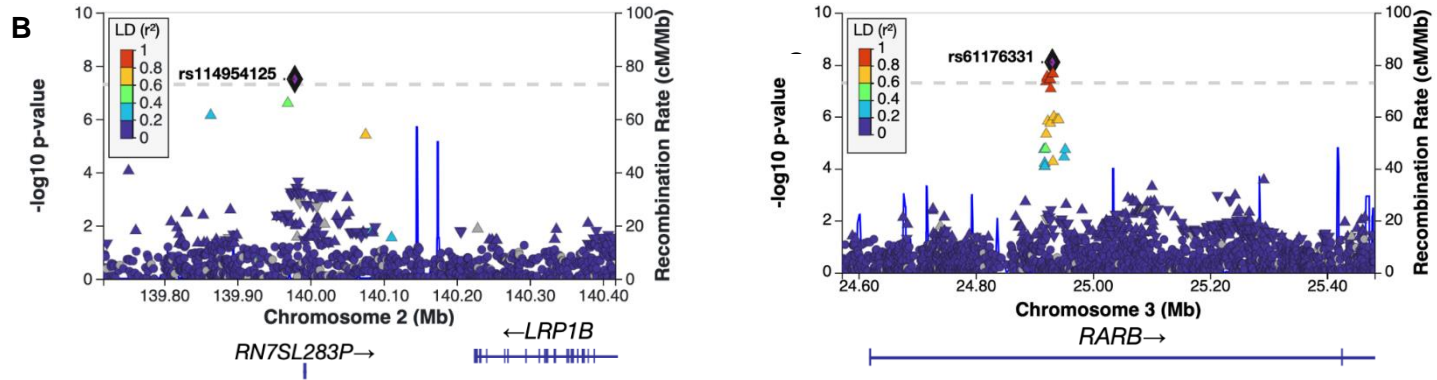
837

838

839

840

841



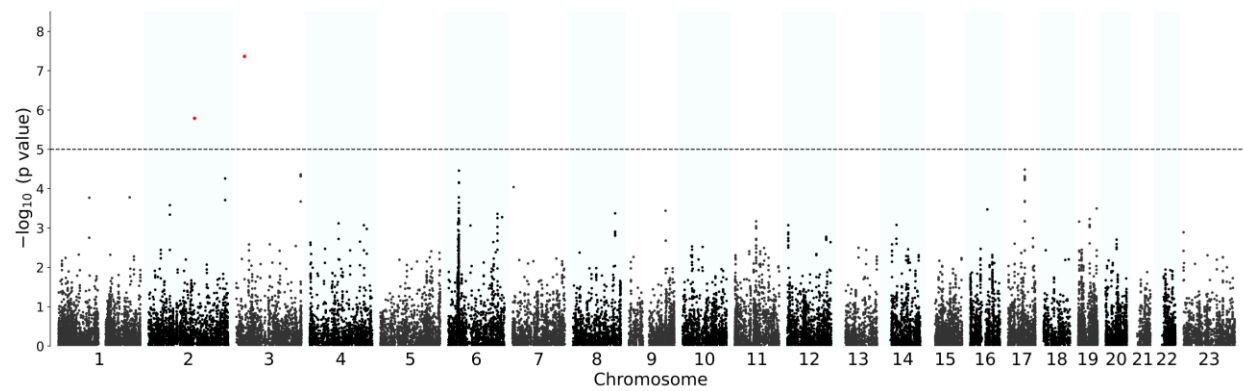
842 **Figure 1.** Manhattan plot of hypertensive disorders of pregnancy GWAS results and LocusZoom plots of genome-wide significant loci
843 in the PEGS cohort (202 cases and 1,569 controls). The genome coordinates for **B** and **C** reflect GRCh38 assembly.

844

845

846

847



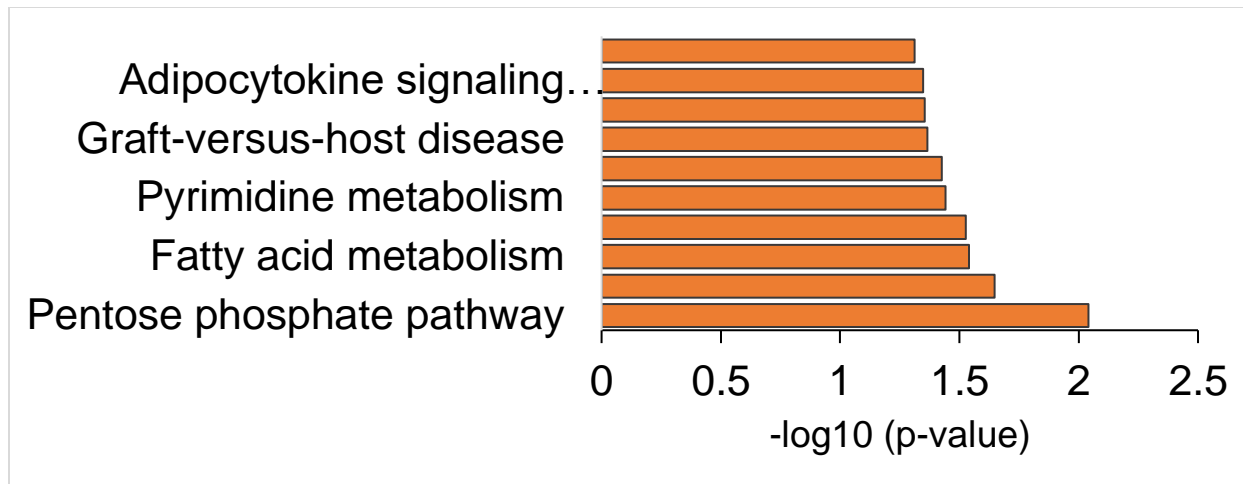
848

849

850 **Figure 2.** Manhattan plot of gene-based association analysis of hypertensive disorders of pregnancy. Labeled genes represent the
851 nearest gene within 50 kb upstream or downstream of the gene.

852

853



854

855

856 **Figure 3.** Gene-set analysis using KEGG with MAGMA. The ten most significant KEGG sets are shown on the y-axis. Corresponding
857 unadjusted $-\log_{10}(p\text{-values})$ are shown on the x-axis. P values were estimated by linear regression model-based testing using
858 MAGMA.

859

860

861

862

863

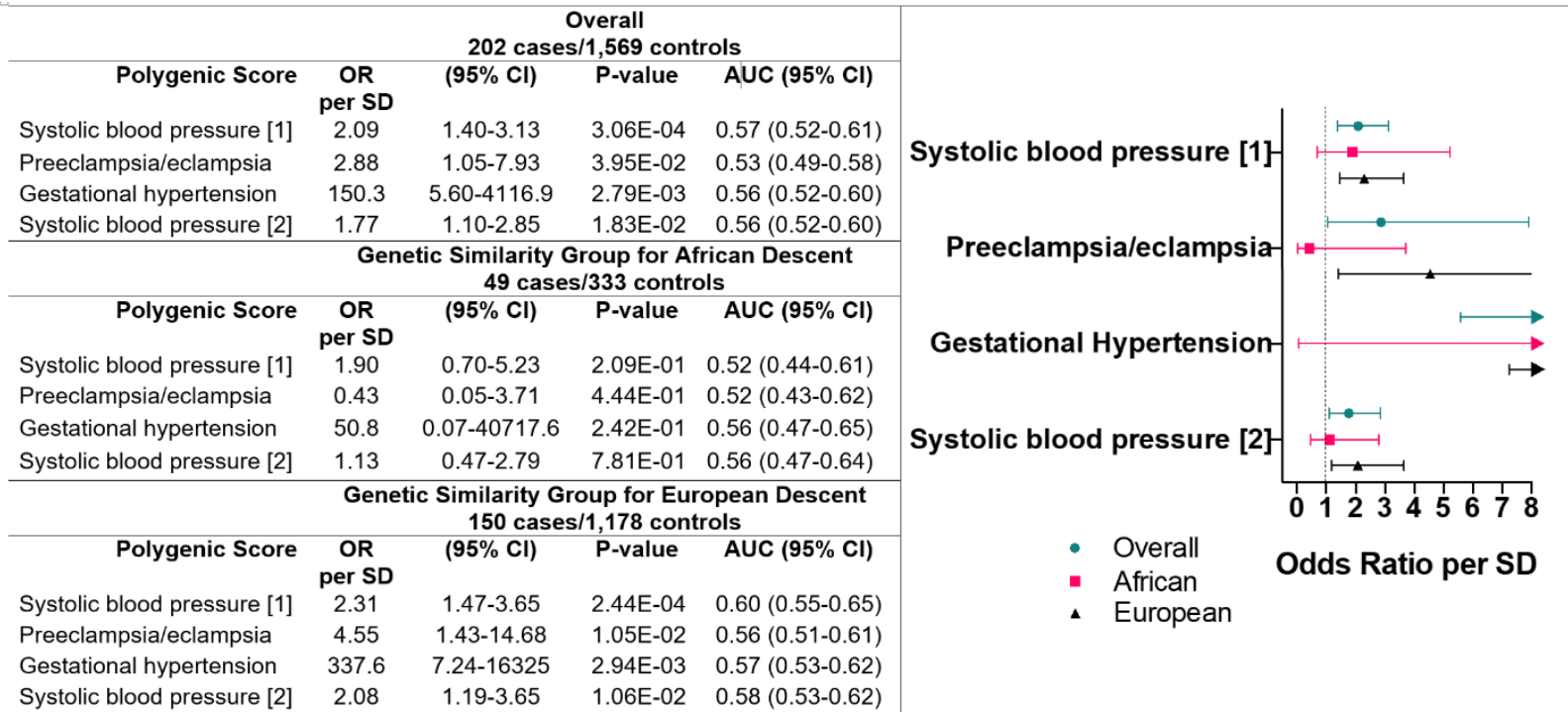
864

865

866

867

868



869

870 **Figure 4.** Polygenic score summary statistics and forest plot for odds ratio (ORs) per standard deviation (SD) for hypertensive
871 disorders of pregnancy in multi-ancestry models and African and European genetic similarity groups. AUC: area under the curve

872