

RESEARCH ARTICLE

# Identifying genes and traits associated with pre-eclampsia using summary statistics

Wenyan Zhu<sup>1\*</sup>, Qi Wang<sup>2\*</sup>, Min Chen<sup>2\*</sup>

**1** Department of Biostatistics, Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing, China, **2** Department of gynecology and obstetrics, Nanjing Women and Children's Healthcare Hospital, Women's Hospital of Nanjing Medical University, Nanjing, China

 These authors contributed to this work equally.

\*[2420912956@qq.com](mailto:2420912956@qq.com)



## Abstract

The occurrence and development of pre-eclampsia (PE) is closely related to genetics. However, multi-omics analysis does not provide sufficient evidence to define significant genes. Therefore, we aimed to identify significant genes and pathways using summary statistics from genome-wide association studies (GWAS). Based on the summary statistics, we used linkage disequilibrium score regression (LDSC) to discover genetic correlation between PE and complex traits. Leveraging summary statistics of tissue-specific expression quantitative trait loci (eQTL), we used FUSION to define significant genes, Bayesian colocalization analysis to identify pleiotropic genes, and Multi-marker Analysis of GenoMic Annotation (MAGMA) to determine the associated pathways. Specifically, considering the potential relationship between PE and tissues, we included 11 tissues, such as kidney cortex. Our integrative analysis revealed that the observed heritability of PE was 0.0179 (standard error [SE] = 0.0021,  $P$ -value < 0.001). Also, based on the Bonferroni correction, we defined 238 traits genetically correlated to PE, such as the other cardiovascular diseases ( $r = -0.55$ ) and furosemide ( $r = 0.79$ ). Integrating eQTL summary statistics across eleven tissues, we identified 30 significant genes, such as *EIF2S1* in the uterus (TWAS.  $Z = 4.44$ , TWAS.  $P = 8.95 \times 10^{-6}$ ), and *PAWRP2* in ovary (TWAS.  $Z = 4.34$ , TWAS.  $P = 1.45 \times 10^{-5}$ ). Based on colocalization, we identified 26 pleiotropic genes. We found that three genes, including *RPS26*, *SULT1A2*, *OBSCN-AS1*, and *SUOX*, were simultaneously defined by FUSION and colocalization. Moreover, we found that the significant enrichment was in the FOXG1\_TARGET\_GENES pathway regulated by the transcription factor *FOXG1* ( $P_{FDR} = 0.049$ ). The findings of post-GWAS analysis for PE indicate that there are 30 significant genes and 26 pleiotropic genes. Future studies are required to investigate the efficacy of targeting pleiotropic genes to reduce the risk of PE.

## OPEN ACCESS

**Citation:** Zhu W, Wang Q, Chen M (2025) Identifying genes and traits associated with pre-eclampsia using summary statistics. PLoS One 20(5): e0323683. <https://doi.org/10.1371/journal.pone.0323683>

**Editor:** Yun Li, University of North Carolina at Chapel Hill, UNITED STATES OF AMERICA

**Received:** November 3, 2024

**Accepted:** April 11, 2025

**Published:** May 30, 2025

**Copyright:** © 2025 Zhu et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data availability statement:** The data of pre-eclampsia GWAS summary statistic is publicly available from <https://www.ebi.ac.uk/gwas/studies/GCST90269903>. The data of tissue-specific eQTL summary statistics are publicly available from <https://ftp.ebi.ac.uk/pub/databases/spot/eQTL/imported/GTEXV8>. The data of summary statistic from Neale lab

is available through [https://nealelab.github.io/UKBB\\_ldsc/index.html](https://nealelab.github.io/UKBB_ldsc/index.html).

**Funding:** The author(s) received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist.

## Introduction

Pre-eclampsia (PE) is a pregnancy-specific syndrome characterized by the onset of hypertension after 20 weeks of gestation, accompanied by at least one of the following complications: proteinuria, maternal organ dysfunction, or uteroplacental dysfunction [1]. Pre-eclampsia (PE) is a complex multisystem disorder associated with long-term health risks for both mothers and their offspring. Evidence indicates that women with a history of PE exhibit reduced life expectancy and an increased risk of developing stroke, cardiovascular disease, and diabetes. Furthermore, offspring born to mothers with PE may face a higher likelihood of cardiovascular and metabolic disorders in later life [2–4]. Previous epidemiological studies indicate that PE affects approximately 3–5% of pregnancies [5], and around 5–10% of pregnancies are affected by hypertensive disorders [6]. When PE begins 28 weeks of gestation, the risk of maternal mortality increases by 24 times compared to when it starts at term [7]. Without timely treatment, PE can lead to serious complications for both the mother and the infant, potentially threatening their lives [8].

From the patterns of onset for eclampsia and PE, increasing evidence suggests that genetics plays at least a partial role in its causality [9]. Advancements in high-throughput sequencing technologies have facilitated the identification of numerous genomic loci associated with pre-eclampsia (PE) through genome-wide association studies (GWAS). For example, a recent GWAS on PE identified 19 genome-wide significant loci, including 13 novel loci. Among these, seven loci contain genes previously implicated in blood pressure regulation, such as *NPPA*, *NPR3*, *PLCE1*, *TNS2*, *FURIN*, *RGL3*, and *PREX1* [10]. Another two-sample Mendelian randomization study indicated a strong causal relationship between rheumatoid arthritis (RA) and PE [11]. PE GWAS shows that most single nucleotide polymorphisms (SNPs) identified do not meet genome-wide significance thresholds, indicating that PE GWAS are still in their early stages [12].

Genetic correlations have emerged as crucial indicators for elucidating the shared genetic architecture of complex traits and the etiology of complex diseases. Estimating genetic correlations is a critical step in deciphering the shared genetic underpinnings of complex traits and diseases. Recently developed methods, such as Linkage Disequilibrium Score Regression (LDSC), have significantly advanced this field [13], Genetic and Environmental Covariance estimation by composite-likelihood Optimization (GECKO) [14], and Leveraging Allelic Variants Analysis (LAVA) [15] can compute global correlations between traits, while large-scale quantitative trait locus (QTL) data have been generated to determine the relationship between genotypes and gene expression (eQTL) [16,17]. This has facilitated the development of statistical techniques that integrate multidimensional data more easily. By contrast, transcriptome-wide association studies (TWASs) have been used to study the relationship between phenotypes and gene expression [18]. Furthermore, the emergence of methods, such as coloc [19] and SMR [20], enables us to investigate whether two traits share causal variants. Overall, integrating multidimensional eQTL data into GWAS data will help explore the pathogenesis of PE. Despite multiple attempts, there is no

integration of multi-omics analyses based on GWAS, making it challenging to move from risk loci identification to therapeutic approaches.

Given the shortcomings of previous studies – such as the limited exploration of genetic correlations between PE and other complex traits, the lack of comprehensive multi-omics integration, and the reliance on European-centric data—this study proposes a series of post-GWAS analyses [21]. Specifically, we leverage the 1000 Genome Project (1000GP) as a reference panel for European individuals to ensure robust linkage disequilibrium (LD) estimation and SNP quality control. However, we acknowledge the need for future studies to include more diverse populations to enhance the generalizability of the findings. First, we estimated the heritability of pre-eclampsia (PE) and evaluated its genetic correlations with 4,347 other phenotypes from the UK Biobank (UKB). Next, by integrating expression quantitative trait loci (eQTL) summary statistics across 11 relevant tissues, we conducted transcriptome-wide association studies (TWAS) to identify associated genes and performed colocalization analysis to detect pleiotropic genes. Finally, we carried out enrichment analysis to explore the biological mechanisms underlying the significant genes.

## Methods

### GWAS summary statistics collection

We collected a PE summary statistic estimated by 296,824 participants in European (EUR) and Asian (ASA) from four cohorts, including 16,743 cases and 280,081 controls [10]. This PE summary statistic includes 14,447 European cases versus 278,022 European controls and 2,296 Asian cases versus 2,059 Asian controls. The ratio between case and control is 1:16.7, which might not be imbalanced. Data from the Finnish Genetics of Pre-eclampsia Consortium (FINNPEC, 1990–2011), the Finnish FinnGen project (1964–2019), the Estonian Biobank (1997–2019), and the previously published InterPregGen consortium [10]. This study included 12,556,951 SNPs [10]. Referring to the EUR individuals in 1000GP, we included SNPs: i) minor allele frequency (MAF) > 0.01; ii) Hardy Weinberg equilibrium (HWE) >  $10E-7$  [22–24]. After the stringent quality control for SNPs, we remained with 8,916,273 high-quality SNPs in the following analysis.

Additionally, for TWAS and genetic correlation analysis, we included the tissue-specific eQTL summary statistics and summary statistics for complex traits. For the eQTL, we also included eQTL data for 10 tissues with potential relationships with PE from FUSION (<http://gusevlab.org/projects/fusion/>), including Adipose-Subcutaneous, Adipose-Visceral, Adrenal Gland, Artery-Aorta, Artery-Coronary, Artery-Tibial, Heart-Atrial Appendage, Heart-Left Ventricle, Kidney-Cortex, and Uterus [20]. For the genetic correlation analysis, we included 4,347 phenotypes categorized under the ‘both\_sexes’ category from the UK Biobank (UKB), excluding features marked as ‘pending’ in the md5 file. All of the summary statistics are shared freely with the scientific community on the Neale Lab website ([https://nealelab.github.io/UKBB\\_ldsc/downloads.html/](https://nealelab.github.io/UKBB_ldsc/downloads.html/)) [25].

### Genetic correlation

To examine whether PE and different genes in various tissues are influenced by common causal variants, we conducted colocalization analysis using the coloc method [19] on the summary data for PE and the eQTL data published for different tissues from GTEx [26]. For heritability estimation, LDSC regresses the LD score to the chi-square for each SNP. For genetic correlation, LDS regresses the product of marginal  $z$ -scores for the two traits on LD scores of each SNP. The majority of the GWAS data used included EUR Europeans; therefore, we regarded 503 EUR individuals as the reference panel. The coloc method is based on Bayesian statistics and calculates the posterior probability that there is a causal variant within a region affecting both traits by assigning priors to the association of each SNP within the region with the two traits. For the posterior probabilities calculated by coloc, we set a threshold of  $PP\ H4 > 0.95$  to identify functional genes sharing genetic variants with the trait.

## Transcriptomics-wide association study

To explore functional genes associated with the trait in different tissues, we performed a transcriptome-wide association study (TWAS) using the Functional Summary-Based Imputation (FUSION) method [27]. This method estimates the association between predicted gene expression and a trait by calculating a weighted linear combination of SNP-trait standardized effect sizes, while accounting for linkage disequilibrium (LD) among SNPs. Using FUSION, we identified genes significantly associated with the trait across multiple tissues, and the Benjamini-Hochberg correction was applied to control the false discovery rate.

## Colocalization

Colocalization analysis is a popular technique for explaining how important signal sites impact the phenotype, which is required once GWAS analysis identifies significant signal sites [28]. We hypothesize that the loci identified by GWAS signals may modulate biological processes related to gene expression, thereby influencing phenotypic outcomes when eQTL colocalization and GWAS signals coincide. Colocalization analysis is commonly employed to determine whether two phenotypes are driven by the same genetic variant within a specific genomic region, thereby providing robust evidence for the previously observed association between the two phenotypes.

Specifically, when GWAS signaling and eQTL co-localization are detected, we assume that sites on the GWAS signaling may influence the phenotype by altering the biological processes of gene expression. SNP.PP H4 represents the posterior probability that the GWAS significant signal and eQTL site are the same site, ranging from 0 to 1. We treated genes with PP H4 > 0.95 as the pleiotropic genes. We used the *coloc* package (<https://github.com/chr1swallace/coloc>) to fit the colocalization model.

## Enrichment analysis

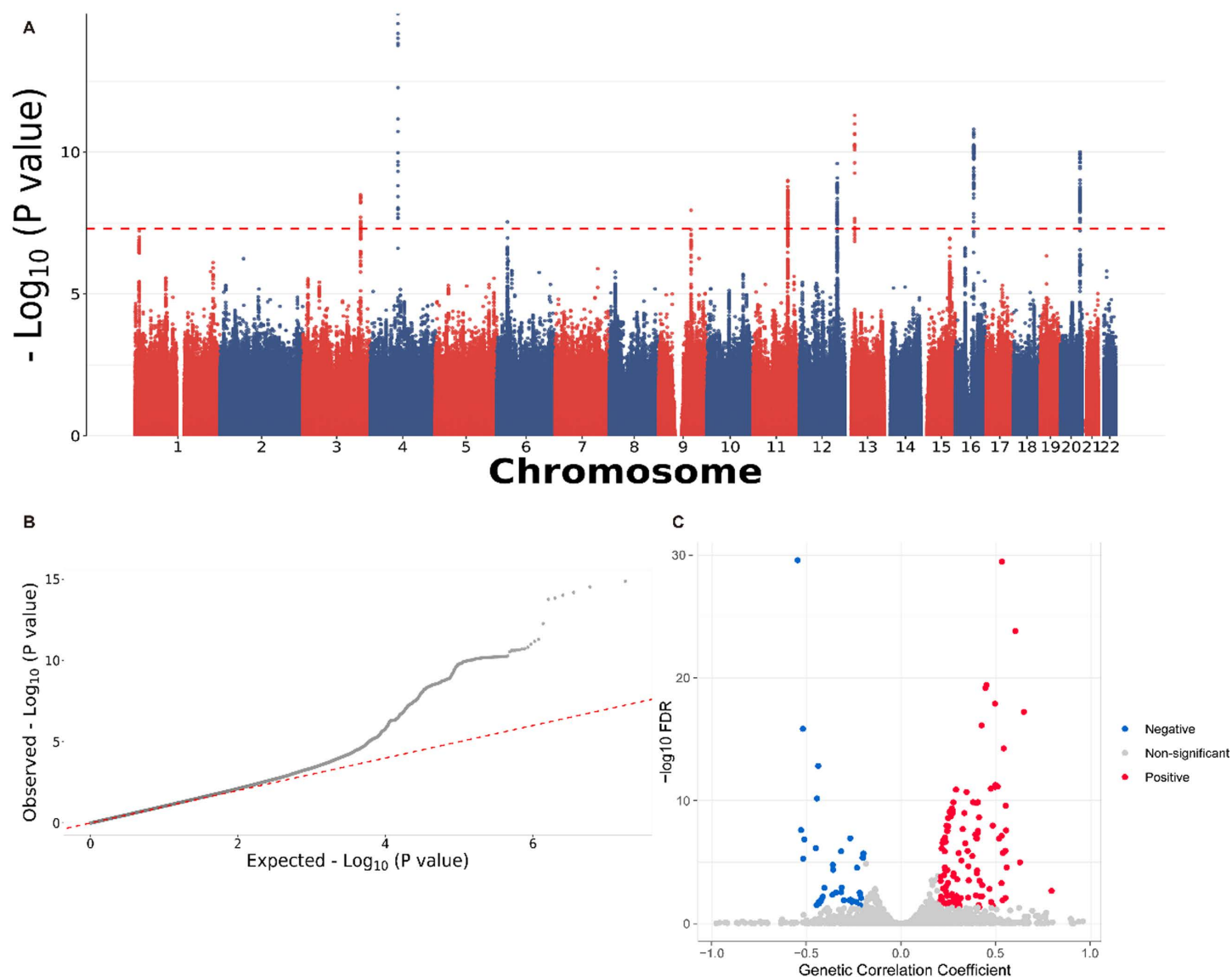
To investigate the biological function of pleiotropic genes and associated genes, we performed the enrichment analysis using Multi-marker Analysis of GenoMic Annotation (MAGMA), a tool for gene analysis and generalized gene-set analysis of GWAS data [29]. Based on multivariate regression models for gene and pathway analysis, it can be used to analyze both raw genotype data and summary statistics (<https://ctg.cncr.nl/software/magma>) [30]. Through MAGMA gene and gene set analysis, all relevant enrichment pathways containing the common genes found in TWAS and colocalization analysis were found, and significant pathways were selected by Benjamini-Hochberg correction. We used GWAS-associated genes to identify genes associated with specific traits or diseases from GWAS with the destination genome assembly of GRCh38/hg38.3. We conducted MAGMA analysis using the common genes defined by FUSION and colocalization in 11 tissues. We used the FUMA platform to provide functional enrichment results from MAGMA [25]. We set FDR < 0.05 as the threshold for the significant pathways.

## Results

### Genetic correlation between PE and complex traits

Based on LDSC, the observed heritability of PE was 0.0179 (standard error [SE] = 0.0021,  $P$ -value < 0.001; Fig 1A). Consistent with the previous paper [31], we defined 289 significant SNPs with  $P$ -value <  $1 \times 10^{-8}$ . The genetic inflation factor ( $\lambda$ ) was 1.071, indicating that there may be a small number of SNP mutations (Fig 1B).

Additionally, we investigated the genetic correlation between PE and 4,347 complex traits with LDSC (Fig 1C). The observed heritability of 4,347 traits, on average, was 0.065 (ranging from -2.23, for Job SOC coding: paramedics, to 2.69, for recent medication for chronic bronchitis), in which the heritability of, e.g., fat percentage (left) traits was significant (FDR < 0.05). The 238 significant genetic correlations between PE and the complex traits ranged from -0.550 ( $P$ -value =  $8.33 \times 10^{-34}$ , FDR =  $2.65 \times 10^{-30}$ ) for the other cardiovascular diseases to 0.7932 ( $P$ -value =  $8.64 \times 10^{-5}$ ,



**Fig 1. Summary for the GWAS analysis for PE.** (A) Manhattan plot for PE. (B) Q-Q plot for PE. (C) Volcano plot for genetic correlation between PE and 4,347 complex traits.

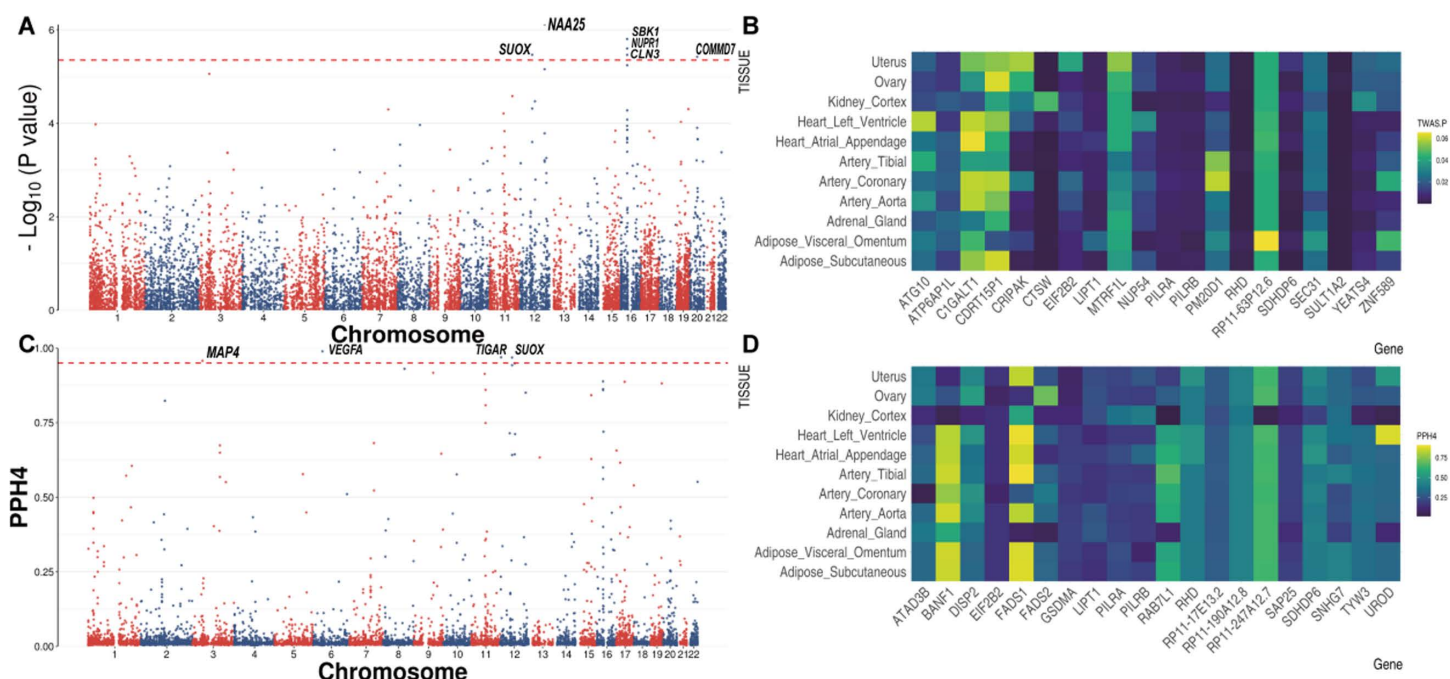
<https://doi.org/10.1371/journal.pone.0323683.g001>

FDR =  $2.16 \times 10^{-3}$ ) for receiving diuretic therapy. For example, furosemide is a potent diuretic that primarily works by increasing urine excretion to reduce fluid volume in the body, thereby alleviating symptoms such as edema. For pregnant women with PE who have complications related to hypertension, doctors might consider using furosemide to manage fluid retention associated with PE [32].

### TWAS and colocalization genes and pleiotropic genes

We performed TWAS for PE by integrating with summary statistics of eleven human-related tissues using the FUSION pipeline. We identified 30 significant genes using Bonferroni correction ( $0.05/\text{number of included genes}$ ). For thyroid, we defined 6 significant genes, the largest number of significant genes among the eleven tissues (Fig 2A, 2B and S1 Table).





**Fig 2. Summary for the TWAS and colocalization for PE.** (A) Manhattan plot for TWAS. (B) Heatmap for the top 20 significant genes in TWAS. (C) Manhattan plot for colocalization. (D) Heatmap for the top 20 pleiotropic genes in colocalization.

<https://doi.org/10.1371/journal.pone.0323683.g002>

Six significant genes in thyroid were *SUOX* (TWAS. *Z*: 4.64, TWAS. *P*:  $3.40 \times 10^{-6}$ ), *NAA25* (TWAS. *Z*: -4.94, TWAS. *P*:  $7.68 \times 10^{-7}$ ), *SBK1* (TWAS. *Z*: -4.80, TWAS. *P*:  $1.56 \times 10^{-6}$ ), *CLN3* (TWAS. *Z*: 4.64, TWAS. *P*:  $3.42 \times 10^{-6}$ ), *NUPR1* (TWAS. *Z*: 4.71, TWAS. *P*:  $2.49 \times 10^{-6}$ ), and *COMMD7* (TWAS. *Z*: -4.62, TWAS. *P*:  $3.77 \times 10^{-6}$ ) (Table 1).

Moreover, we conducted fine-mapping on the TWAS result to identify causal variants for pleiotropic genes in relevant tissues, such as artery tibial and thyroid. Using SuSiE, we identified 10 genes with at least one causal SNP (*ALDH2*, *CDC37P2*, *CLN3*, *NUPR1*, *LINC00484*, *SUOX*, *NAA25*, *SBK1*, *CLN3*, *NUPR1*, and *COMMD7*). *CLN3* had 12 and 16 causal SNPs in artery tibial and thyroid, respectively (Table 2). The *CLN3* gene is involved in basic physiological processes such as intracellular material transport and maintaining lysosomal homeostasis. In placental tissue, these cellular functions are crucial for the differentiation, proliferation of trophoblast cells, and the integrity of the placental barrier. In the

**Table 1. Summary for significant genes in Thyroid with FUSION.**

Gene	Chromosome	Position	Number of <i>cis</i> -SNP	Z	<i>P</i> -value
<i>SUOX</i>	12	55997179-55997180	293	4.64	$3.40 \times 10^{-6}$
<i>NAA25</i>	12	112109021-112109022	201	-4.94	$7.68 \times 10^{-6}$
<i>SBK1</i>	16	28292518-28292519	289	-4.80	$1.56 \times 10^{-6}$
<i>CLN3</i>	16	28495097-28495098	215	4.64	$3.42 \times 10^{-6}$
<i>NUPR1</i>	16	28539173-28539174	207	4.71	$2.49 \times 10^{-6}$
<i>COMMD7</i>	20	32743996-32743997	416	-4.62	$3.77 \times 10^{-6}$

1<sup>st</sup> column: gene symbol; 2<sup>nd</sup> column: chromosome; 3<sup>rd</sup> column: chromosome location of the gene; 4<sup>th</sup> column: the number of *cis*-SNP; 5<sup>th</sup> column: Z-value; 6<sup>th</sup> column: *P* value.

<https://doi.org/10.1371/journal.pone.0323683.t001>

placentas of patients with preeclampsia, abnormalities like insufficient invasion of trophoblast cells, placental ischemia, and hypoxia often occur [33].

In addition, we performed COLOC for PE by integrating with summary statistics of human eleven related tissues. We identified 26 pleiotropic genes with PP H4 > 0.95 (Fig 2C and 2D and S2 Table). For artery tibial, we identified six pleiotropic genes with PP H4 > 0.95, including *SUOX* (PP H4 = 0.98), *AGRN* (PP H4 = 0.98), *CLCN6* (PP H4 = 0.97), *TIGAR* (PP H4 = 0.97), *GDNF* (PP H4 = 0.95), and *GDNF-AS1* (PP H4 = 0.97) (Table 3). Previous studies have identified separate loci of *CLCN6* associated with PE, with roles in natriuretic peptide signaling, angiogenesis, glomerular function, trophoblast development, and immune dysregulation [34]. For thyroid, we defined 4 significant genes, the largest number of significant genes among the eleven tissues (Fig 2B). The four significant genes in thyroid were *VEGFA* (PP H4 = 0.99), *TIGAR* (PP H4 = 0.97), *SUOX* (PP H4 = 0.97), and *MAP4* (PP H4 = 0.96) (Table 3).

We found that three genes, including *RPS26*, *SULT1A2*, and *SUOX*, were defined with TWAS and colocalization analysis (Fig 3). rs1873914 was the lead SNP for *SUOX* with PP H4 = 0.98, TWAS  $P = 3.40 \times 10^{-6}$ , and GWAS  $P = 1.01 \times 10^{-5}$ . Sulfite oxidase is a homomeric enzyme situated in the intermembrane space of mitochondria. *SUOX* changes in the activity or expression of sulfate metabolites may lead to changes that may have an impact on blood vessel function

**Table 2. Summary for the fine-mapping for the associated genes.**

Gene	CHR	Start	End	P(min)	Tissue	NO.SNP
<i>ALDH2</i>	12	111,766,886	111,766,887	3.18E-15	Artery Tibial	3
<i>CDC37P2</i>	16	28,413,702	28,413,703	3.49E-06	Artery Tibial	20
<i>CLN3</i>	16	28,495,097	28,495,098	6.02E-06	Artery Tibial	12
<i>NUPR1</i>	16	28,539,173	28,539,174	1.95E-41	Artery Tibial	11
<i>LINC00484</i>	9	91,159,572	91,159,573	9.67E-08	Artery Tibial	9
<i>SUOX</i>	12	55,997,179	55,997,180	1.91E-76	Thyroid	7
<i>SBK1</i>	16	28,292,518	28,292,519	1.42E-06	Thyroid	12
<i>CLN3</i>	16	28,495,097	28,495,098	1.54E-09	Thyroid	29
<i>NUPR1</i>	16	28,539,173	28,539,174	3.94E-08	Thyroid	19
<i>COMMD7</i>	20	32,743,996	32,743,997	2.26E-17	Thyroid	21

1<sup>st</sup> column: gene symbol; 2<sup>nd</sup> column: chromosome; 3<sup>rd</sup> and 4<sup>th</sup> columns: the start and end position for gene; 5<sup>th</sup> column: minimum value for cis-SNP in summary statistics; 6<sup>th</sup> column: tissue; 7<sup>th</sup> column: the number of causal SNP.

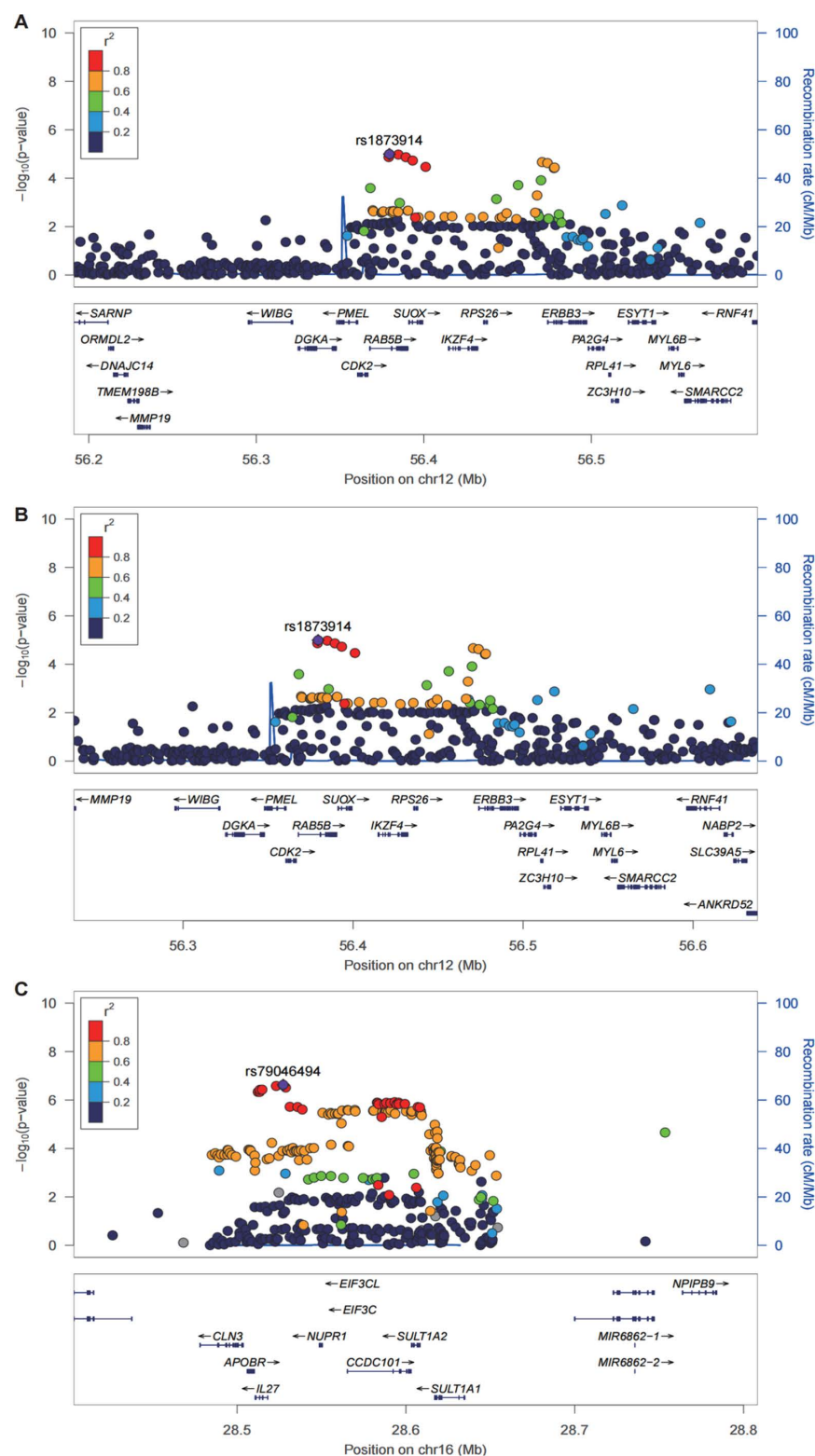
<https://doi.org/10.1371/journal.pone.0323683.t002>

**Table 3. Summary of pleiotropic genes in Artery Tibial and Thyroid.**

Gene	Tissue	Chromosome	Position	Lead SNP	PP H4
<i>SUOX</i>	Artery Tibial	12	56379427	rs1873914	0.98
<i>AGRN</i>	Artery Tibial	1	953678	rs113454255	0.98
<i>CLCN6</i>	Artery Tibial	1	11858957	rs45553335	0.97
<i>TIGAR</i>	Artery Tibial	12	4429331	rs191800458	0.97
<i>GDNF</i>	Artery Tibial	5	38060688	rs270594	0.95
<i>GDNF-AS1</i>	Artery Tibial	5	38056271	rs169187	0.97
<i>VEGFA</i>	Thyroid	6	43746169	rs3025000	0.99
<i>TIGAR</i>	Thyroid	12	4429331	rs191800458	0.97
<i>SUOX</i>	Thyroid	12	56393337	rs1081975	0.97
<i>MAP4</i>	Thyroid	3	48185266	rs9832957	0.96

1<sup>st</sup> column: gene symbol; 2<sup>nd</sup> column: tissues; 3<sup>rd</sup> column: chromosome; 4<sup>th</sup> column: chromosome location of the gene; 5<sup>th</sup> column: lead SNP; 6<sup>th</sup> column: PP H4 value.

<https://doi.org/10.1371/journal.pone.0323683.t003>



**Fig 3. LocusZoom of *SUOX* (A), *RPS26* (B), and *SULT1A2* (C) in the summary statistics of PE in thyroid.**

<https://doi.org/10.1371/journal.pone.0323683.g003>

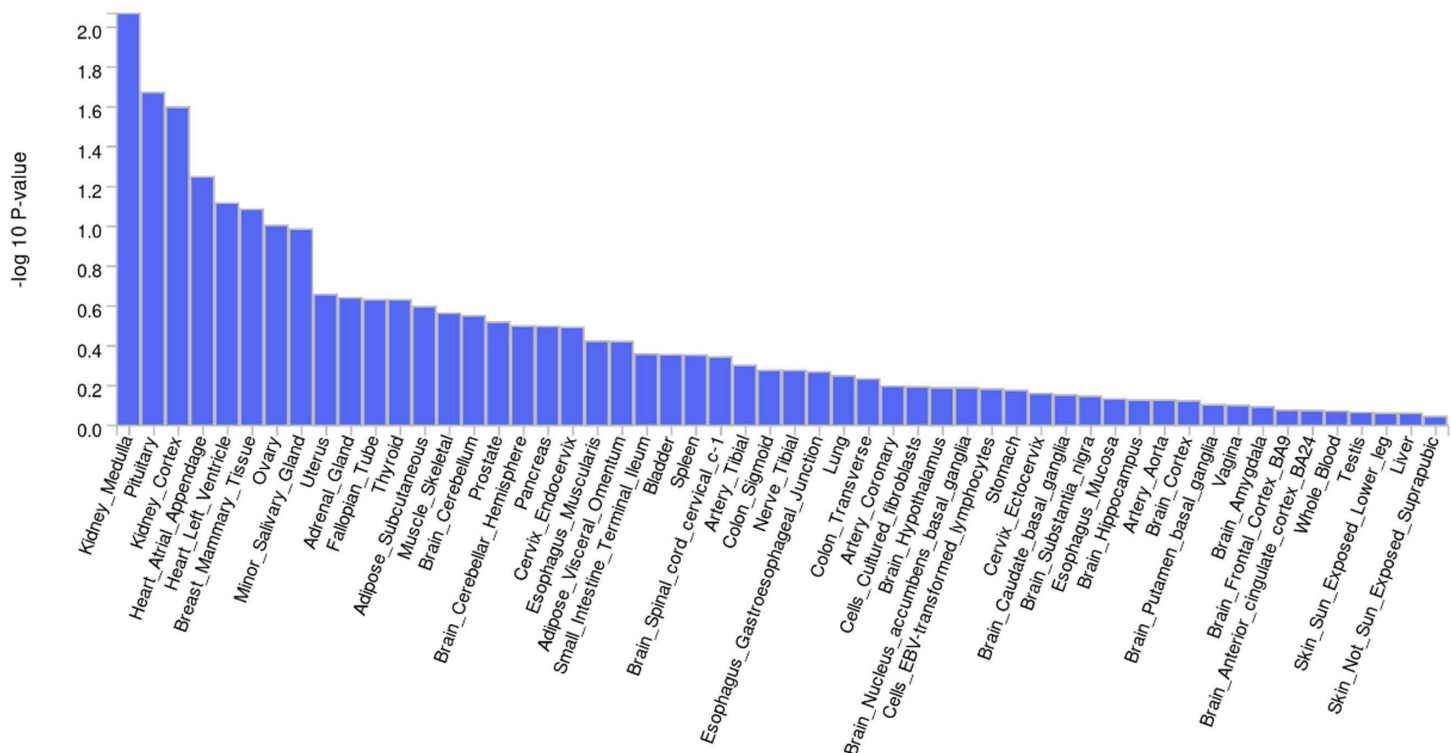


during pregnancy [35]. The rs1873914 was the lead SNP for *RPS26* with PP H4=0.96, TWAS  $P=2.29 \times 10^{-5}$ , and GWAS  $P=1.01 \times 10^{-5}$ . Vascular endothelial dysfunction in patients with PE is one of its typical manifestations. As a ribosomal protein, *RPS26* primarily functions in protein synthesis; however, alterations in its expression might influence intracellular metabolic pathways, which could indirectly impact angiogenesis and endothelial function [36]. Finally, rs79046494 was the lead SNP for *SULT1A2* with PP H4=0.95, TWAS  $P=4.77 \times 10^{-7}$ , and GWAS  $P=2.37 \times 10^{-7}$ . *SULT1A2* is involved in hormone sulfation metabolism, especially estrogen metabolism. PE is associated with hormonal imbalances, so changes in *SULT1A2* activity may affect hormone levels and thus the development of PE [37].

## Enrichment analysis

We used the summary statistics of PE to perform the enrichment analysis at the SNP level. Based on FDR, we obtained three significant pathways that have a role in the pathophysiology of PE, suggesting that the development and course of the illness may be significantly impacted by their dysregulation. Integrating the tissue information, the enrichment analysis from FUMA gives a more comprehensive picture for PE. The renal medulla and other organs might link to particular biochemical processes that facilitate the development of PE ( $P$ -value<0.05) (Fig 4). It is possible that the genes expressed in the renal medulla are enriched in pathways associated with electrolyte balance and renal function, both of which are vital for preserving blood pressure equilibrium [38]. The expression of genes in the pituitary gland and renal cortex may also be enriched in pathways related to the regulation of hormones and renal filtration processes,

Based on the gene enrichment analysis results for pre-eclampsia, the X-axis shows the results of 53 tissue types from GTEx v8, while the Y-axis represents the  $-\log_{10}P$  values.



**Fig 4. Summary for the results from MAGMA.** The x-axis shows the pathway names, and y-axis shows the  $-\log_{10}$  of the  $P$ -value. The kidney medulla displays the highest pathway enrichment score, suggesting its central role in PE pathogenesis. The pituitary gland and renal cortex also show significant enrichment, implicating their potential involvement in PE-associated gene regulatory mechanisms.

<https://doi.org/10.1371/journal.pone.0323683.g004>

respectively [39]. Moreover, according to MAGMA enrichment analysis, after correction for FDR of four genes, including *RPS26*, *SULT1A2*, *OBSCN-AS1*, and *SUOX*, were simultaneously defined by FUSION and colocalization, it was found that the significant enrichment was in a group of genes (*FOXG1\_TARGET\_GENES*) pathway regulated by transcription factor *FOXG1* ( $BH\_corrected\_P\_value = 0.049$ ). There is *SULT1A2* in four genes on this pathway. Genes regulated by *FOXG1* may be involved in the function of vascular endothelial cells, thereby affecting vascular health during pregnancy.

## Discussion

In our study, we performed a comprehensive post-GWAS analysis for PE, including genetic correlation, TWAS, colocalization and enrichment analysis. In this post-GWAS analysis, we confirmed the genetic associations of PE. We used LDSC to estimate heritability, which is calculated by regression LD score and SNP chi-square statistics. In contrast, Steinthorsdottir et al. reported higher heritability estimates using Genomic Relatedness Restricted Maximum Likelihood (GREML) [40]. Different methods might cause different estimation for heritability. Our study examines the genetic correlations between 4,347 traits and PE, in contrast to the 894 pairs analyzed by Tyrmi et al [10]. Additionally, we employed MAGMA to perform enrichment analyses, whereas Tyrmi et al. focused on identifying associated genes. In their research, Tyrmi et al. conducted extensive data analyses and provided insights into the genetic architecture of PE [10]. We explored and identified 25,185 potential SNPs located on pathogenic loci that may influence PE. As a supplementary analysis, MAGMA was used, which is flexible and computationally fast, with greater power and less reliance on linkage disequilibrium data. It can directly use GWAS summary statistics as input data. The aim of using MAGMA was to demonstrate the relative differences in gene expression levels across different tissues, identifying the tissues where gene expression changes are most likely to contribute to the occurrence of PE. Integrating the tissue information, we observed significant associations for kidney medulla, pituitary, and kidney cortex ( $P\text{-value} < 0.05$ ). In the past, Steinthorsdottir et al. reported the results of a meta-analysis of eight GWAS on preeclampsia in European and Central Asian women [40]. Based on a large meta-analysis, the study performed association analysis, estimated genetic correlation between 12 traits, and estimated the heritability using LDSC and GCTA. Based on summary statistics, our study examines the genetic correlations between 4,347 traits and PE and focuses on the post-GWAS analysis, such as colocalization to identify pleiotropy genes, TWAS to define significant genes, and enrichment analysis to identify relevant pathways [40].

PE is primarily caused by systemic endothelial dysfunction, which is characterized by reduced vascular integrity, endothelial damage, and impaired vasodilation [41,42]. Human studies support the potential importance of this heptapeptide during pregnancy, as an increase in urinary ANG-(1–7) has been observed throughout gestation. The initial rise in urinary ANG-(1–7) was detected at 12 weeks of gestation, but an increasing trend was observed as early as week 6 of pregnancy, suggesting that early adaptive changes in the renal RAS system may contribute to changes in fluid and electrolyte balance throughout pregnancy. Additionally, in normal late-term pregnant human subjects, plasma ANG-(1–7) levels increase, whereas in preeclamptic subjects, these levels decrease, indicating that a reduction in this peptide could be a significant factor in PE [43]. This information can help guide subsequent experimental design and drug-targeted therapeutic strategies, e.g., interventions targeting tissues such as the renal medulla, ventricles, breast, thyroid, etc. may be useful in improving symptoms of PE.

This work revealed a common genetic etiology and possible pleiotropy between PE and thyroid by examining the genome-wide and localized genetic connections between the two conditions. In studies integrating and associating multi-omics data, it has been demonstrated that the expression of thyroid-related genes is associated with PE. Experiments have indicated that, compared to hyperthyroidism, subclinical hypothyroidism during pregnancy is associated with a higher risk of PE. TSH (Thyroid-Stimulating Hormone) exhibits a U-shaped relationship with PE. These findings quantify the risk of hypertension or PE in women with abnormal thyroid function tests, adding to the overall evidence of the risk of adverse maternal and fetal outcomes associated with thyroid dysfunction during pregnancy [44].

## Limitation

There are many restrictions on this study. First, the majority of the GWAS data used included Europeans, which limited the applicability of our findings to other ethnic groups. More varied groups should be included in future studies to confirm and build on these findings. Second, an exact causal relationship cannot be derived from this figure alone, as further experimental validation and other evidence are needed. Third, although we employ a variety of sensitivity analyses to address the issues of pleiotropy and heterogeneity, the inherent limitations of coloc and FUSION may still affect the robustness of our causal inferences. Traits with complex genetic architectures, such as those influenced by multiple small-effect SNPs, pose challenges for coloc. The presence of pleiotropy (where one SNP affects multiple traits) can complicate the interpretation of results. Both coloc and FUSION analyses assume that genetic factors are the primary drivers of observed associations. However, environmental factors can also play significant roles in modulating gene expression and disease risk, which may not be captured by these methods.

## Supporting information

**S1 Table. Summary for significant genes in ten tissues with FUSION.**

(XLSX)

**S2 Table. Summary for pleiotropic genes in ten tissues with coloc.**

(XLSX)

## Acknowledgments

We extend our sincere gratitude to the participants and funding agencies for their invaluable support and contributions to this research endeavor. Their commitment has been instrumental in bringing our study to fruition. The authors thank Dr. Xingyu Zhang at School of Health and Rehabilitation Science, University of Pittsburg for his helpful editorial suggestions.

## Author contributions

**Conceptualization:** Wenyan Zhu.

**Data curation:** Wenyan Zhu.

**Formal analysis:** Wenyan Zhu, Qi Wang.

**Funding acquisition:** Min Chen.

**Investigation:** Wenyan Zhu, Qi Wang.

**Methodology:** Wenyan Zhu.

**Project administration:** Qi Wang.

**Resources:** Wenyan Zhu, Qi Wang.

**Software:** Wenyan Zhu, Qi Wang.

**Supervision:** Min Chen.

**Validation:** Qi Wang, Min Chen.

**Visualization:** Wenyan Zhu.

**Writing – original draft:** Wenyan Zhu.

**Writing – review & editing:** Qi Wang, Min Chen.

## References

1. Dimitriadis E, Rolnik DL, Zhou W, Estrada-Gutierrez G, Koga K, Francisco RPV, et al. Pre-eclampsia. *Nat Rev Dis Primers*. 2023;9(1):8. <https://doi.org/10.1038/s41572-023-00417-6> PMID: 36797292
2. Poon LC, Shennan A, Hyett JA, Kapur A, Hadar E, Divakar H, et al. The International Federation of Gynecology and Obstetrics (FIGO) initiative on pre-eclampsia: A pragmatic guide for first-trimester screening and prevention. *Int J Gynaecol Obstet*. 2019;145 Suppl 1(Suppl 1):1–33. <https://doi.org/10.1002/ijgo.12802> PMID: 31111484
3. Pittara T, Vyrides A, Lamniso D, Giannakou K. Pre-eclampsia and long-term health outcomes for mother and infant: an umbrella review. *BJOG*. 2021;128(9):1421–30. <https://doi.org/10.1111/1471-0528.16683> PMID: 33638891
4. Kvalvik LG, Wilcox AJ, Skjærven R, Østbye T, Harmon QE. Term complications and subsequent risk of preterm birth: registry based study. *BMJ*. 2020;369:m1007. <https://doi.org/10.1136/bmj.m1007> PMID: 32349968
5. Ananth C, Keyes K, Wapner R. Pre-eclampsia rates in the United States, 1980–2010: age-period-cohort analysis. *BMJ*. 2013;347(f6564).
6. Hutcheon JA, Lisonkova S, Joseph KS. Epidemiology of pre-eclampsia and the other hypertensive disorders of pregnancy. *Best Pract Res Clin Obstet Gynaecol*. 2011;25(4):391–403. <https://doi.org/10.1016/j.bpobgyn.2011.01.006> PMID: 21333604
7. Khan KS, Wojdyla D, Say L, Gülmezoglu AM, Van Look PF. WHO analysis of causes of maternal death: a systematic review. *Lancet*. 2006;367(9516):1066–74. [https://doi.org/10.1016/S0140-6736\(06\)68397-9](https://doi.org/10.1016/S0140-6736(06)68397-9) PMID: 16581405
8. Yang S, Li H, Ge Q, Guo L, Chen F. Deregulated microRNA species in the plasma and placenta of patients with preeclampsia. *Mol Med Rep*. 2015;12(1):527–34. <https://doi.org/10.3892/mmr.2015.3414> PMID: 25738738
9. Roberts JM, Cooper DW. Pathogenesis and genetics of pre-eclampsia. *Lancet*. 2001;357(9249):53–6. [https://doi.org/10.1016/s0140-6736\(00\)03577-7](https://doi.org/10.1016/s0140-6736(00)03577-7) PMID: 11197372
10. Tyrmi J, Kaartokallio T, Lokki A, Jääskeläinen T, Kortelainen E, Ruotsalainen S. Genetic risk factors associated with preeclampsia and hypertensive disorders of pregnancy. *JAMA Cardiol*. 2023;8(7):674–83.
11. Zhang D, Hu Y, Guo W, Song Y, Yang L, Yang S, et al. Mendelian randomization study reveals a causal relationship between rheumatoid arthritis and risk for pre-eclampsia. *Front Immunol*. 2022;13:1080980. <https://doi.org/10.3389/fimmu.2022.1080980> PMID: 36578485
12. Miller EC, Wilczek A, Bello NA, Tom S, Wapner R, Suh Y. Pregnancy, preeclampsia and maternal aging: From epidemiology to functional genomics. *Ageing Res Rev*. 2022;73:101535. <https://doi.org/10.1016/j.arr.2021.101535> PMID: 34871806
13. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh P-R, et al. An atlas of genetic correlations across human diseases and traits. *Nat Genet*. 2015;47(11):1236–41. <https://doi.org/10.1038/ng.3406> PMID: 26414676
14. Gao B, Yang C, Liu J, Zhou X. Accurate genetic and environmental covariance estimation with composite likelihood in genome-wide association studies. *PLoS Genet*. 2021;17(1):e1009293. <https://doi.org/10.1371/journal.pgen.1009293> PMID: 33395406
15. Werme J, van der Sluis S, Posthuma D, de Leeuw CA. An integrated framework for local genetic correlation analysis. *Nat Genet*. 2022;54(3):274–82. <https://doi.org/10.1038/s41588-022-01017-y> PMID: 35288712
16. Zhang J, Zhao H. eQTL studies: from bulk tissues to single cells. *J Genet Genomics*. 2023;50(12):925–33. <https://doi.org/10.1016/j.jgg.2023.05.003> PMID: 37207929
17. Cao C, Tian M, Li Z, Zhu W, Huang P, Yang S. Gwashug: a comprehensive platform for decoding the shared genetic basis between complex traits based on summary statistics. *Nucleic Acids Res*. 2024.
18. Wainberg M, Sinnott-Armstrong N, Mancuso N, Barbeira AN, Knowles DA, Golan D, et al. Opportunities and challenges for transcriptome-wide association studies. *Nat Genet*. 2019;51(4):592–9. <https://doi.org/10.1038/s41588-019-0385-z> PMID: 30926968
19. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet*. 2014;10(5):e1004383. <https://doi.org/10.1371/journal.pgen.1004383> PMID: 24830394
20. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet*. 2016;48(5):481–7. <https://doi.org/10.1038/ng.3538> PMID: 27019110
21. Farashi S, Kryza T, Clements J, Batra J. Post-GWAS in prostate cancer: from genetic association to biological contribution. *Nat Rev Cancer*. 2019;19(1):46–59. <https://doi.org/10.1038/s41588-018-0087-3> PMID: 30538273
22. Yang S, Zhou X. Accurate and Scalable Construction of Polygenic Scores in Large Biobank Data Sets. *Am J Hum Genet*. 2020;106(5):679–93. <https://doi.org/10.1016/j.ajhg.2020.03.013> PMID: 32330416
23. Yang S, Zhou X. PGS-server: accuracy, robustness and transferability of polygenic score methods for biobank scale studies. *Brief Bioinform*. 2022;23(2):bbac039. <https://doi.org/10.1093/bib/bbac039> PMID: 35193147
24. Cao C, Zhang S, Wang J, Tian M, Ji X, Huang D, et al. PGS-Depot: a comprehensive resource for polygenic scores constructed by summary statistics based methods [J]. *Nucleic Acids Res*. 2024, 52(D1): D963–d71.
25. Trubetskoy V, Pardiñas AF, Qi T, Panagiotaropoulou G, Awasthi S, Bigdeli TB, et al. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature*. 2022;604(7906):502–8. <https://doi.org/10.1038/s41586-022-04434-5> PMID: 35396580
26. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science*. 2015;348(6235):648–60. <https://doi.org/10.1126/science.1262110> PMID: 25954001

27. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BWJH, et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet.* 2016;48(3):245–52. <https://doi.org/10.1038/ng.3506> PMID: [26854917](#)
28. Rasooly D, Peloso GM, Giambartolomei C. Bayesian genetic colocalization test of two traits using coloc. *Curr Protoc.* 2022;2(12):e627.
29. Ni J, Wang P, Yin K, Yang X, Cen H, Sui C. Novel insight into the aetiology of rheumatoid arthritis gained by a cross-tissue transcriptome-wide association study. *RMD Open.* 2022;8(2).
30. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol.* 2015;11(4):e1004219. <https://doi.org/10.1371/journal.pcbi.1004219> PMID: [25885710](#)
31. McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JPA, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet.* 2008;9(5):356–69. <https://doi.org/10.1038/nrg2344> PMID: [18398418](#)
32. Tamás P, Kovács K, Várnagy Á, Farkas B, Alemu WG, Bódis J. Preeclampsia subtypes: Clinical aspects regarding pathogenesis, signs, and management with special attention to diuretic administration [J]. *Eur J Obstet Gynecol Reprod Biol.* 2022;274(175):81.
33. Zhou W, Dimitriadis E. Tripeptidyl peptidase 1 regulates human trophoblast cell proliferation implying a role in placentation. *Biomed Res Int.* 2022;2022(6856768).
34. Honigberg MC, Truong B, Khan RR, Xiao B, Bhatta L, Vy HMT, et al. Polygenic prediction of preeclampsia and gestational hypertension. *Nat Med.* 2023;29(6):1540–9. <https://doi.org/10.1038/s41591-023-02374-9> PMID: [37248299](#)
35. Zhang R, Hao Y, Xu Y, Qin J, Wang Y, Kumar Dey S, et al. Whole exome sequencing identified a homozygous novel mutation in SUOX gene causes extremely rare autosomal recessive isolated sulfite oxidase deficiency. *Clin Chim Acta.* 2022;532:115–22. <https://doi.org/10.1016/j.cca.2022.06.005> PMID: [35679912](#)
36. Ferretti MB, Ghalei H, Ward EA, Potts EL, Karbstein K. Rps26 directs mRNA-specific translation by recognition of Kozak sequence elements. *Nat Struct Mol Biol.* 2017;24(9):700–7. <https://doi.org/10.1038/nsmb.3442> PMID: [28759050](#)
37. Zhang P, Yang H, Feng Y, Wu W, Li S, Thompson B, et al. Polymorphisms in Sex Hormone Metabolism Genes and Risk of Preeclampsia in Taiyuan, China. *Gynecol Obstet Invest.* 2018;83(2):179–86. <https://doi.org/10.1159/000478931> PMID: [29059671](#)
38. Villalaín González C, Herraiz García I, Fernández-Friera L, Ruiz-Hurtado G, Morales E, Solís J, et al. Cardiovascular and renal health: Preeclampsia as a risk marker. *Nefrologia (Engl Ed).* 2023;43(3):269–80. <https://doi.org/10.1016/j.nefro.2022.04.009> PMID: [37635012](#)
39. Kuriya A, Morris DV, Dahan MH. Pituitary injury and persistent hypofunction resulting from a peripartum non-hemorrhagic, vaso-occlusive event. *Endocrinol Diabetes Metab Case Rep.* 2015;2015:150001. <https://doi.org/10.1530/EDM-15-0001> PMID: [26038692](#)
40. Steinthorsdottir V, McGinnis R, Williams NO, Stefansdottir L, Thorleifsson G, Shooter S, et al. Genetic predisposition to hypertension is associated with preeclampsia in European and Central Asian women. *Nat Commun.* 2020;11(1):5976. <https://doi.org/10.1038/s41467-020-19733-6> PMID: [33239696](#)
41. Boeldt D, Bird I. Vascular adaptation in pregnancy and endothelial dysfunction in preeclampsia. *J Endocrinol.* 2017;232(1):R27–44.
42. Daehn IS, Duffield JS. The glomerular filtration barrier: a structural target for novel kidney therapies. *Nat Rev Drug Discov.* 2021;20(10):770–88. <https://doi.org/10.1038/s41573-021-00242-0> PMID: [34262140](#)
43. Merrill DC, Karoly M, Chen K, Ferrario CM, Brosnihan KB. Angiotensin-(1-7) in normal and preeclamptic pregnancy. *Endocrine.* 2002;18(3):239–45. <https://doi.org/10.1385/ENDO:18:3:239> PMID: [12450315](#)
44. Toloza FJK, Derakhshan A, Männistö T, Bliddal S, Popova PV, Carty DM, et al. Association between maternal thyroid function and risk of gestational hypertension and pre-eclampsia: a systematic review and individual-participant data meta-analysis. *Lancet Diabetes Endocrinol.* 2022;10(4):243–52. [https://doi.org/10.1016/S2213-8587\(22\)00007-9](https://doi.org/10.1016/S2213-8587(22)00007-9) PMID: [35255260](#)