BMC Pregnancy and Childbirth

RESEARCH

Open Access

Cross-ancestry genome-wide association study identifies new susceptibility genes for preeclampsia

Yuping Shan[1], Hong Hu[2] and Yijing Chu[1*]

Abstract

Background Preeclampsia (PE) is a heterogeneous, multi-organ pregnancy disorder that poses a significant health burden globally, with its pathogenesis remaining unclear. This study aimed to identify novel susceptibility genes for PE through a cross-ancestry genome-wide association study (GWAS).

Methods We performed meta-analysis to summarize the PE GWAS data from the United Kingdom, Finland, and Japan. Subsequently, the multi-ancestry sum of the single-effects model was used to perform cross-ancestry fine-mapping. The functional mapping and annotation (FUMA)-expression quantitative trait loci (eQTL) mapping method, transcriptome-wide association study (TWAS)- functional summary-based imputation (FUSION) method, genome-wide complex trait analysis (GCTA)-multivariate set-based association test (mBAT)-combo method, and polygenic priority score (PoPS) method were employed to screen for candidate genes. We utilized biomarker expression level imputation using summary-level statistics (BLISS), based on summary-level protein quantitative trait loci (pQTL) data, to conduct a multi-ancestry proteome-wide association study (PWAS) analysis, followed by candidate drug prediction.

Results Six novel susceptibility genes associated with PE risk were identified: *NPPA*, *SWAP70*, *NPR3*, *FGF5*, *REPIN1*, and *ACAA1*. High expression of the *NPPA* and *SWAP70* and low expression of the remaining genes were associated with a reduced risk of PE. Furthermore, we identified drugs that target *NPPA*, *NPR3*, and *REPIN1*.

Conclusions Our study identified *NPPA*, *SWAP70*, *NPR3*, *FGF5*, *REPIN1*, and *ACAA1* as novel genes whose predicted expression was linked to the risk of PE, offering new insights into the genetic framework of this condition.

Keywords Drug targets, Genetics, Genome-wide association study, Pre-eclampsia

*Correspondence:
Yijing Chu
chuyijing@163.com
[1]Department of Obstetrics and Gynecology, The Affiliated Hospital of Qingdao University, Qingdao, China
[2]Clinical Medicine, Nantong University, Nantong, China

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Introduction

Preeclampsia (PE) is a pregnancy disorder characterized by a systolic blood pressure of approximately 140 mmHg and/or a diastolic blood pressure of at least 90 mmHg, measured at least twice, four hours apart in previously normotensive women after 20 weeks of gestation; it is accompanied by proteinuria or other maternal end-organ dysfunctions including acute kidney injury, liver involvement, neurological complications, hematological complications, pulmonary edema, and uteroplacental dysfunction [1, 2]. With improper treatment, PE can progress to eclampsia, stroke, and even death [3]. Currently, PE affects approximately 1.2–8% of pregnancies worldwide and accounts for more than 76,000 maternal and 500,000 perinatal deaths annually [1, 4]. Although PE poses a significant health burden worldwide, the specific underlying mechanisms remain unclear. Traditionally, PE has been recognized as a progressive disease characterized by impaired placentation resulting from inadequate remodeling of the spiral arteries, followed by oxidative stress (OS) and systemic maternal endothelial dysfunction [5]. Multiple factors, including genetic factors, immune-system modifications, abnormal metabolism, and an imbalance between angiogenic and anti-angiogenic factors, may contribute to PE [6, 7].

Although the exact mechanisms involved in the pathogenesis of PE are not yet fully understood, an increasing number of studies based on limited sample sizes have demonstrated that dysregulation of gene expression plays a significant role in its development [8, 9]. Currently, some genome-wide association studies (GWASs) investigating the risk of PE have been conducted; however, most of these studies focused on specific racial groups [10]. Considering the variations in allele frequency, linkage disequilibrium (LD), and effect size among different populations, it is crucial to conduct GWAS in ancestrally diverse groups. This approach will facilitate the identification of both ancestry-shared and ancestry-specific genetic associations related to PE, ensuring that the GWAS findings are broadly applicable [11].

To identify new susceptibility genes for PE across diverse ancestries, we conducted this study using a meta-analysis of GWAS to summarize GWAS data from the United Kingdom, Finland, and Japan. This study aims to identify ancestry-shared genetic associations related to PE, and our findings may contribute to developing more universally applicable strategies for the treatment of PE.

Materials and methods

Study design

The data analysis process is illustrated in Fig. 1. First, we collected the PE data from three publicly available GWAS databases. Next, we performed a meta-analysis of the acquired data to obtain the GWAS results across various racial groups. Subsequently, fine-mapping and screening of the candidate genes were conducted. Finally, we performed a preliminary exploration of potential therapeutic targets for these candidate genes.

Datasets

Publicly available data were used for this study. To obtain the PE GWAS data, we selected: GCST90301704 from the GWAS catalog, which includes 194,127 samples (1728 cases and 192,399 controls) from the United Kingdom; the FinnGen R11 dataset, which comprises 242,332 samples (8,185 cases and 234,147 controls) from Finland; and GCST90018686 from the GWAS catalog, which contains 82,085 samples (123 cases and 81,962 controls) from Japan [12–14]. Expression quantitative trait loci (eQTL) data were obtained from the Genotype Tissue Expression (GTEx) v8 dataset web portal and the eQTLGen Consortium. The GTEx v8 dataset comprises 17,382 samples from 52 tissues and 2 cell lines, whereas the eQTLGen Consortium includes 16,987 genes and 31,684 cis-eQTLs derived from blood samples of predominantly healthy European individuals [15–17]. The protein quantitative trait loci (pQTL) data were acquired from: the UK Biobank Pharma Proteomics Project (UKB-PPP), which provides information on 2,923 proteins from 54,219 UKB participants; deCODE Genetics, which includes 4,907 plasma proteins from 35,559 Icelandic participants; and the Atherosclerosis Risk in Communities (ARIC) study, which includes 4,483 plasma proteins from 7,213 individuals of European American descent [18–20].

Statistical analysis

We included the three summary statistical datasets described above to perform a GWAS meta-analysis using METAL [21]. METAL is a software tool designed for the meta-analysis of GWAS that allows researchers to combine data from different studies or databases to increase the power to detect genetic associations with traits of interest [22]. Single nucleotide polymorphisms (SNPs) with a minor allele frequency (MAF) > 0.01 and a *HetPVal* > 0.05 [23] were retained. The following analyses were based on GWAS data obtained from the meta-analysis. *HetPVal* is primarily used to assess the degree of heterogeneity in the association results across various subgroups or data sources. *HetPVal* < 0.05 is widely accepted as an indicator of significant heterogeneity in the association results among subgroups. This suggests that we cannot simply aggregate the association results from each subgroup to arrive at a universally applicable conclusion. Consequently, in line with a previous study [23], we included SNPs with a *HetPVal* > 0.05 in the present study to exclude those that exhibited significant heterogeneity. Unlike Mendelian randomization, current cross-ancestry GWASs cannot perform sensitivity

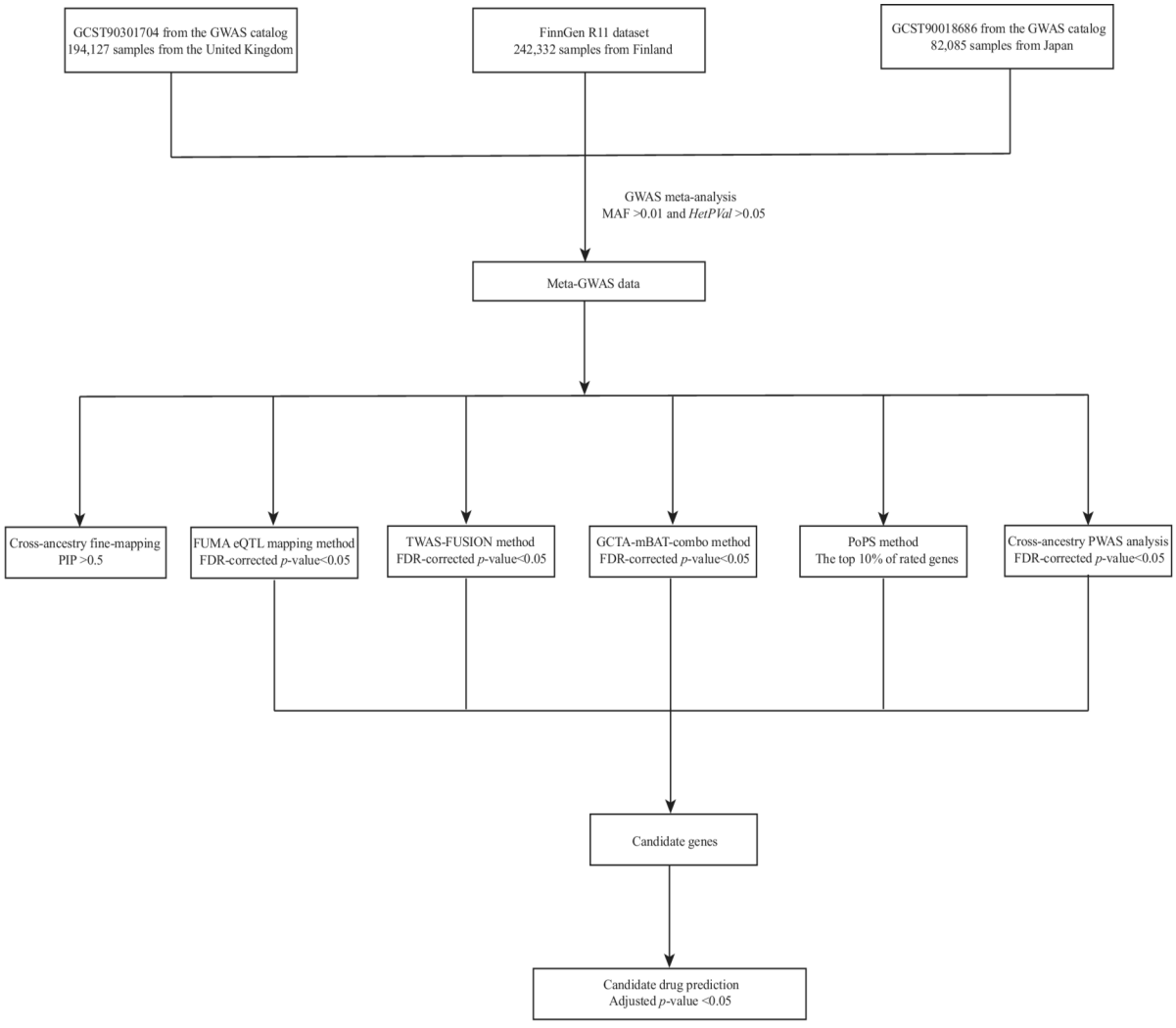

Fig. 1 Flowchart of the study. FDR, False discovery rate; FUMA, Functional Mapping and Annotation; FUSION, Functional summary-based imputation; GCTA, Genome-wide complex trait analysis; GWAS, Genome-wide association study; mBAT, Multivariate set-based association test; PoPS, Polygenic priority score; PWAS, Proteome-wide association study; TWAS, Transcriptome-wide association study

analyses. As a result, similar to other GWASs, we did not perform sensitivity analyses [24]. However, in our study, we employed various methods to identify reliable new susceptibility genes; these methods can be considered as indirect sensitivity analyses.

Probabilistic fine-mapping of causal gene sets

To perform fine-mapping, we utilized the multi-ancestry sum of the single effects model (MESuSiE), a probabilistic fine-mapping method that enhances the accuracy and resolution of fine-mapping by leveraging association information across different ancestries [25]. Probabilistic fine-mapping is a crucial approach in genomic studies aimed at identifying causal gene sets associated with complex traits and diseases. This method leverages statistical models to prioritize genetic variants and genes that are most likely to have a causal effect on a given phenotype [26]. MESuSiE improves the fine-mapping resolution compared to existing approaches. It uses summary data as input, considers various LD patterns observed across different ancestries, explicitly models both shared and ancestry-specific causal SNPs, and utilizes a variational inference approach for scalable computations. We performed MESuSiE using PLINK 2.0 and Rstudio 4.4.0 software [27]. SNPs with a posterior inclusion probability (PIP) > 0.5 were considered significant. PIP is used to evaluate the likelihood that a genetic variation, such as an SNP or gene, serves as a causal factor in a biological trait. PIP > 0.5 indicates a relatively high probability that the genetic variation is associated with the disease [24]; this is utilized in the present study.

Candidate gene selection

eQTL mapping is a method that links genetic variation to gene expression, enabling the identification of genomic loci that contain regulators of gene expression [28]. eQTL mapping can help elucidate the genetic architecture by linking genetic variants to changes in gene expression that may contribute to disease susceptibility [29]. We performed eQTL mapping using functional mapping and annotation (FUMA) web applications. FUMA eQTL mapping utilizes data from three repositories: GTEx, the Blood eQTL Browser, and the BIOS QTL Browser [30]. It maps SNPs to genes based on significant eQTL associations. We applied a false discovery rate (FDR) threshold of 0.05 in each analysis to identify significant eQTL associations.

Transcriptome-wide association studies (TWASs) can provide valuable insights into the relationships between genes and traits by integrating the effects of eQTLs into a single, robust predictor of gene expression [31]. It is a powerful tool for bridging the gap between genotypes and phenotypes, offering a deeper understanding of the genetic basis of complex traits and diseases across various organisms and conditions [32]. We used the functional summary-based imputation (FUSION) tool to examine the causal relationships between gene expression levels and PE using Rstudio 4.4.0 software. Candidate genes were identified as those with an FDR-corrected p-value < 0.05.

The genome-wide complex trait analysis (GCTA)-multivariate set-based association test (mBAT)-combo integrates various test statistics without requiring knowledge of the correlation structure by combining the mBAT and fastBAT test statistics through a Cauchy combination method [33–35]. This method is particularly advantageous in scenarios where multiple SNPs within a gene or genomic region collectively influence a trait, thus providing a more comprehensive understanding of genetic influences on complex traits [34]. Using the Linux operating system, we utilized data from the 1000 Genomes Project Phase 3 to perform a GCTA-mBAT-combo analysis for gene prioritization. Differences were considered statistically significant at an FDR-corrected p-value of < 0.05.

The polygenic priority score (PoPS) can prioritize causal genes identified in GWAS that predict polygenic genetic associations based on gene expression profiles, protein-protein interaction networks, and pathway databases [36]. By learning biologically relevant properties from various gene features, PoPS can effectively prioritize causal genes associated with complex traits and diseases [37]. This method has shown promise for identifying gene-trait pairs with high precision, not only confirming well-established relationships but also nominating new genes at unresolved loci [36]. We utilized PoPS to identify the top 10% rated genes as candidate genes.

The final candidate genes were identified by intersecting the genes identified using the four approaches described above.

Proteome-wide association study (PWAS) analysis

We utilized biomarker expression level imputation using summary-level statistics (BLISS), a novel method developed to create protein imputation models based on summary-level pQTL data, to conduct multi-ancestry PWAS analysis [38]. By integrating proteomic data with GWAS findings, PWAS provides a more direct understanding of how genetic variations influence diseases through changes in protein abundance and function [39]. We conducted a trans-ethnic PWAS analysis using RStudio 4.4.0 software, based on the three pQTL datasets mentioned above. Additionally, we applied an FDR-corrected p-value of < 0.05 to obtain the most significant results.

Candidate drug prediction

To determine whether the discovered genes could be effectively used as therapeutic targets, we used the Drug Signatures Database (DSigDB) to evaluate potential protein-drug interactions [40]. DSigDB contains 22,527 gene sets and 17,389 unique compounds associated with 19,531 genes, facilitating the identification of connections between drugs, chemicals, and their target genes. The identified target genes were uploaded to DSigDB, enabling the prediction of drug candidates to evaluate the medicinal activity of these target genes. Enrichr is a comprehensive web-based tool that features 180,184 annotated gene sets from 102 distinct gene set libraries [41–43]. Enrichr can analyze the interactions between transcription factors and their target genes after importing gene information for the key module. We used the Proteomics Drug Atlas module in Enrichr to analyze the expression of target genes, which may reveal their potential mechanisms as therapeutic targets. Differences were considered statistically significant at an adjusted p-value < 0.05.

Results

Overall, we identified 164 SNPs associated with the risk of PE from meta-analysis of the GWAS data (Fig. 2).

Fine-mapping results

One of the key aspects of probabilistic fine-mapping is its ability to incorporate functional annotations to improve the accuracy of identifying causal variants. For instance, a Bayesian framework can be used to systematically integrate functional annotations, which has been shown to increase the discovery power and fine-mapping accuracy in GWAS. This approach allows researchers to compute

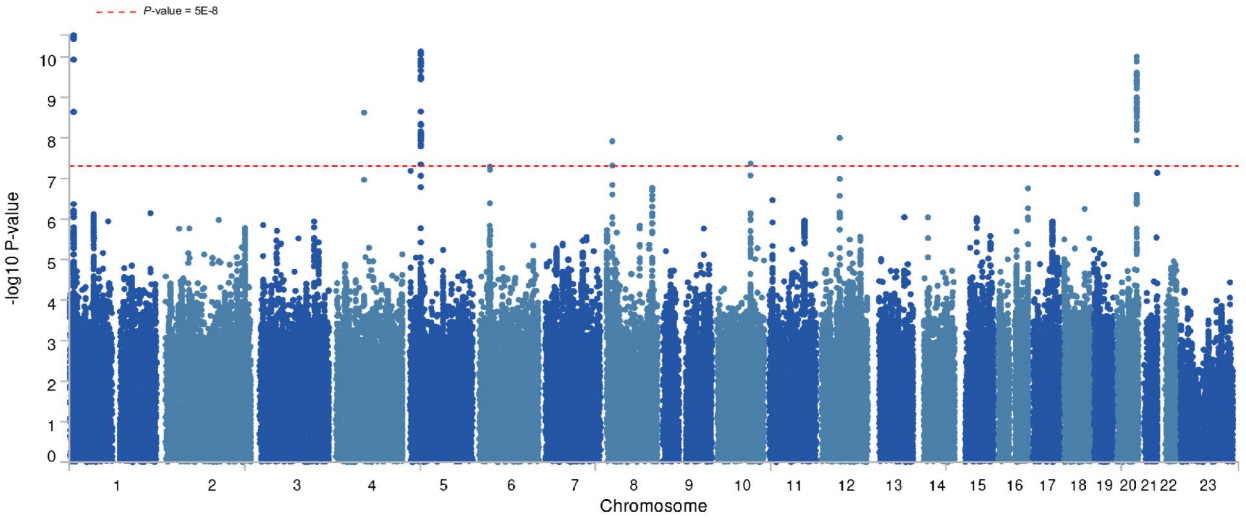

Fig. 2 Manhattan plot illustrating the meta-analysis of GWAS data

the maximum a posteriori solution and optimize penalty parameters through cross-validation, thereby improving the identification of causal variants in complex traits. In the present study, we identified multiple genetic loci associated with PE that were fine-mapped. We conducted a detailed analysis of their ancestral specificities and discovered that rs62368100 was shared among individuals of European and Asian ancestry within the upstream and downstream regions, specifically between positions 32,821,939 and 32,841,939 on chromosome 5 (PIP = 0.575). However, in the same region of chromosome 5, rs1173709 may have a specific effect on European populations (PIP = 0.718). Figure 3 presents the details of these results.

Candidate gene screening

Integrating eQTL data with GWAS enhances the ability to identify causal genes and pathways involved in complex traits. This integration can reveal pleiotropic eQTLs affecting multiple traits, thereby providing a more comprehensive understanding of the genetic basis of phenotypic variation. Additionally, the TWAS-FUSION method is instrumental in identifying candidate genes associated with complex traits and diseases. GCTA-mBAT-combo analysis is especially useful in the context of pleiotropy, where a single genetic variant can affect multiple phenotypic traits, a common occurrence in complex diseases. PoPS not only improves the accuracy of gene prioritization at GWAS loci, but also offers valuable insights into the complex interplay between genetic variants and phenotypic traits, ultimately contributing to the advancement of precision medicine and the identification of novel therapeutic targets. We used FUMA eQTL mapping, TWAS-FUSION, GCTA-mBAT-combo, and PoPS methods to identify 12, 145, 27, and 1691 candidate genes, respectively. Two candidate genes, *NPPA* and *NPR3*, met the stringent criteria after analyzing the intersection of the previously mentioned genes (Fig. 4). Furthermore, high expression levels of *NPPA* (Z = 6.449) and low expression levels of *NPR3* (Z= -5.372) both decrease the risk of PE.

PWAS analysis

PWAS represents a significant advancement in genetic research, offering a more nuanced understanding of the molecular mechanisms underlying complex diseases. By focusing on protein-level changes, PWAS analysis bridges the gap between genetic variants and phenotypic outcomes, paving the way for novel therapeutic strategies. Using BLISS, we identified four genes (*FGF5*, *REPIN1*, *ACAA1*, *SWAP70*) whose regulated protein levels were associated with PE at an FDR-corrected *p*-value of < 0.05. Table 1 summarizes the detailed results of the PWAS analysis. According to these results, low expression levels of *FGF5*, *REPIN1*, and *ACAA1* were associated with a decreased risk of PE. In contrast, high expression level of *SWAP70*, a protective gene against PE was associated with a lower risk of developing PE.

Potential therapeutic targets

The DSigDB is a valuable resource for predicting potentially effective interventional drugs by leveraging a comprehensive collection of drug signatures. This database facilitates the identification of drugs that potentially reverse or modulate the effects of specific gene expression profiles associated with various diseases. The DSigDB was used to predict potentially effective intervention drugs. According to the results presented in Table 2, the most significant drugs associated with *NPPA* included cyclic GMP (CTD 00006063), glycerol

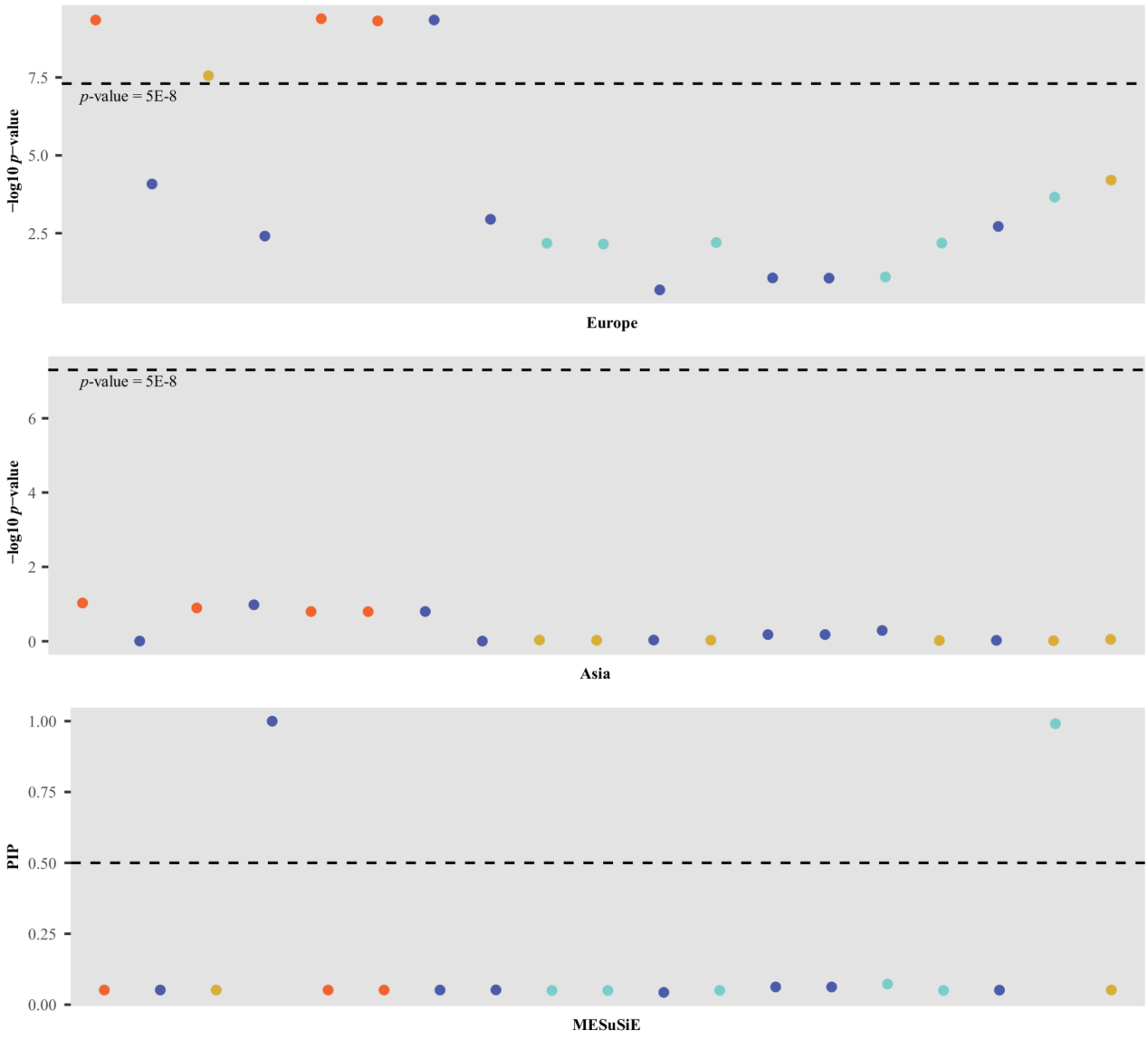

Fig. 3 The results of the cross-ancestry fine mapping. PIP, Posterior inclusion probability

(CTD 00006038), alprostadil (CTD 00005360), labetalol (CTD 00006196), felodipine (CTD 00007084), irbesartan (CTD 00002968), furosemide (CTD 00006012), bonuten (CTD 00005895), naloxone (CTD 00006373), epirubicin (CTD 00007057), phenylephrine (CTD 00006521), aldosterone (CTD 00005347), spironolactone (CTD 00006774), and atenolol (CTD 00005448). Methyl salicylate (CTD 00001586), potassium (CTD 00006595), and barium (CTD 00005464) were the three most significant drugs linked to *NPR3*. Additionally, progesterone (CTD 00006624) interacted with both genes. Additionally, we found that the drugs HX 531 UP, NNC 26-9100 UP, H-Arg(NO2)-OH UP, Bisphenol A UP, Crizotinib UP, Nisoldipine UP, Paliperidone UP, BEC Down, Rifaximin Down, Eptifibatide Down, TC-O 9311 Down, P005091 Down, LY411575 Down, GSK343 Down, and Torcetrapib Down were most significantly associated with *REPIN1*.

Discussion

Cross-ancestry GWASs have proven beneficial for identifying new genetic associations, enhancing the fine-mapping of causal variants and increasing their applicability to underrepresented populations [44, 45]. However, to date, no cross-ancestry GWASs have been conducted on PE. In the present study, we conducted cross-ancestry GWAS meta-analyses to identify new susceptibility genes and potential therapeutic targets for PE. Consistent with previous studies [46, 47], we utilized only cross-ancestry GWAS data without incorporating single-cell RNA sequencing (scRNA-seq) data to uncover the

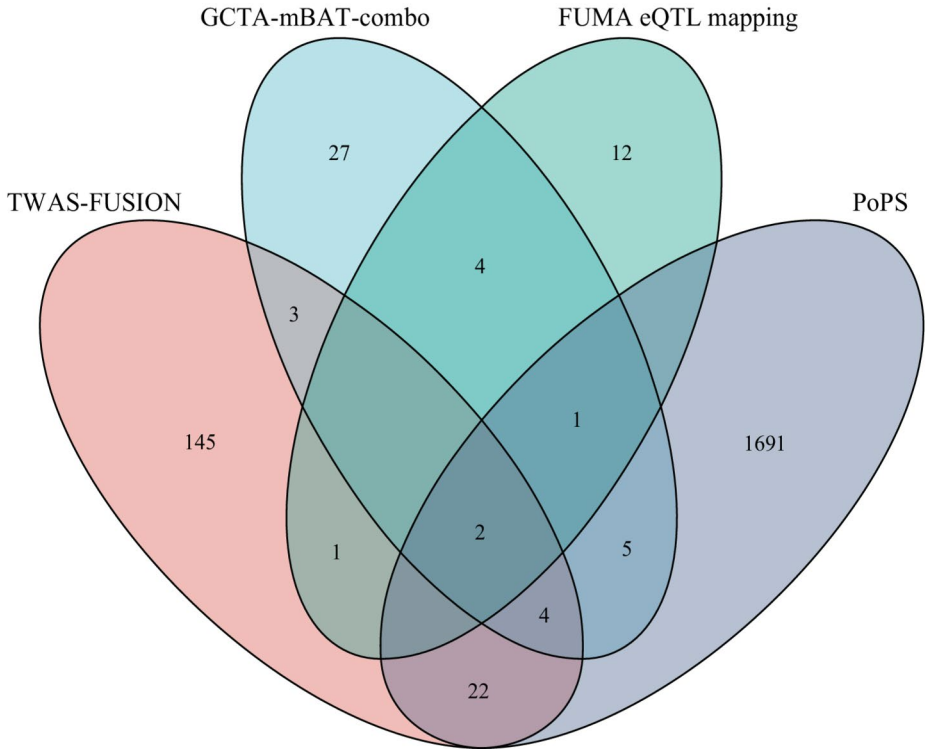

Fig. 4 The results of candidate gene selection. eQTL, Expression quantitative trait loci; FUMA, Functional Mapping and Annotation; FUSION, Functional Summary-based Imputation; GCTA, Genome-wide complex trait analysis; mBAT, Multivariate set-based association test; PoPS, Polygenic priority score; TWAS, Transcriptome-wide association study

Table 1 Cross-ancestry proteome-wide association study (PWAS) analysis

Gene	Chromosome	z-score	*p*-value	FDR-corrected *p*-value
SWAP70	11	-4.456	1.22E-05	4.08E-02
FGF5	4	5.068	7.50E-07	9.25E-03
REPIN1	7	4.652	5.12E-06	3.16E-02
ACAA1	3	4.436	1.32E-05	4.08E-02

FDR, False discovery rate

ancestry-shared and ancestry-specific genetic associations of PE and to ensure that the obtained findings are broadly applicable.

The results of the MESuSiE indicated that certain loci exhibited significant differences among various populations, whereas others demonstrated common genetic effects in both European and Asian populations. Additionally, some loci may have specific effects in either European or Asian populations. These findings provide new insights into the genetic heterogeneity of PE across racial groups. In addition, our study identified a correlation between elevated *NPPA* and *SWAP70* expression, along with a reduction in the expression of *NPR3*, *FGF5*, *REPIN1*, and *ACAA1*, all of which are associated with a decreased risk of PE.

The natriuretic peptide gene *NPPA* and its clearance receptor *NPR3* have been found to be associated with PE. Natriuretic peptide hormones are implicated in controlling the regulation of blood pressure and kidney function, among their several other effects [48]. Armstrong et al. found that mice lacking atrial natriuretic peptide expression developed gestational hypertension and proteinuria [49]. Similar to PE, these mice exhibit impaired trophoblast invasion and remodeling of uterine spiral arteries [8]. *SWAP70* is associated with calcium ion binding and is involved in hypertension [50]. However, its role in PE requires further investigation. *FGF5* has been identified as a susceptibility gene for PE in European and Central Asian women [51]. Xin et al. found that *FGF5* was significantly upregulated in placental tissues from patients with PE and in a PE mouse model, compared to their respective controls [52]. In addition, in vitro cell experiments confirmed that *FGF5* is involved in various biological processes related to cell survival, promoting apoptosis in HTR8/SVneo cells and inhibiting cell invasion [53]. *REPIN1* has not been thoroughly studied since its discovery; however, previous results suggest that this gene may play a role in glucose import, fatty acid transport, iron metabolism, and apoptosis-related pathways [54]. Abnormal iron metabolism in trophoblasts can result in lipid peroxidation and excessive generation of reactive oxygen species (ROS), which can impair extravillous trophoblast invasion and spiral artery remodeling, leading

Table 2 Candidate drug prediction

Drug names	*p*-value	Adjusted *p*-value	Genes
Cyclic gmp CTD 00006063	0.001	0.012	*NPPA*
Glycerol CTD 00006038	0.001	0.012	*NPPA*
Alprostadil CTD 00005360	0.002	0.012	*NPPA*
Labetalol CTD 00006196	0.002	0.012	*NPPA*
Felodipine CTD 00007084	0.002	0.012	*NPPA*
Irbesartan CTD 00002968	0.002	0.012	*NPPA*
Furosemide CTD 00006012	0.002	0.012	*NPPA*
Bonuten CTD 00005895	0.002	0.012	*NPPA*
Naloxone CTD 00006373	0.002	0.012	*NPPA*
Epirubicin CTD 00007057	0.003	0.012	*NPPA*
Phenylephrine CTD 00006521	0.003	0.012	*NPPA*
Aldosterone CTD 00005347	0.004	0.012	*NPPA*
Spironolactone CTD 00006774	0.004	0.012	*NPPA*
Atenolol CTD 00005448	0.004	0.012	*NPPA*
Verapamil CTD 00006985	0.005	0.014	*NPPA*
Carvedilol CTD 00001961	0.006	0.016	*NPPA*
Hydrocortisone CTD 00006117	0.010	0.022	*NPPA*
Manganese chloride CTD 00001187	0.020	0.035	*NPPA*
Manganese CTD 00006240	0.021	0.036	*NPPA*
Allococaine CTD 00005697	0.029	0.046	*NPPA*
methyl salicylate CTD 00001586	0.003	0.012	*NPR3*
Potassium CTD 00006595	0.004	0.012	*NPR3*
Barium CTD 00005464	0.004	0.012	*NPR3*
Buflomedil PC3 DOWN	0.011	0.023	*NPR3*
Acrolein CTD 00005313	0.014	0.029	*NPR3*
Chromium CTD 00005668	0.015	0.029	*NPR3*
Luteolin HL60 UP	0.016	0.030	*NPR3*
Apigenin HL60 UP	0.020	0.035	*NPR3*
1-Chloro-2,4-dinitrobenzene CTD 00005848	0.024	0.040	*NPR3*
Progesterone CTD 00006624	0.009	0.022	*NPPA*, *NPR3*
HX 531 Up	0.028	0.030	*REPIN1*
BEC Down	0.028	0.030	*REPIN1*
Rifaximin Down	0.028	0.030	*REPIN1*
Eptifibatide Down	0.028	0.030	*REPIN1*
TC-O 9311 Down	0.028	0.030	*REPIN1*
P005091 Down	0.029	0.030	*REPIN1*
NNC 26-9100 Up	0.029	0.030	*REPIN1*
H-Arg(NO2)-OH Up	0.029	0.030	*REPIN1*
Bisphenol A Up	0.029	0.030	*REPIN1*
Crizotinib Up	0.029	0.030	*REPIN1*
Nisoldipine Up	0.029	0.030	*REPIN1*
LY411575 Down	0.029	0.030	*REPIN1*
GSK343 Down	0.029	0.030	*REPIN1*
Torcetrapib Down	0.029	0.030	*REPIN1*
Paliperidone Up	0.030	0.030	*REPIN1*

to placental ischemia and hypoxia [55]. This sequence of events triggers vascular inflammation, endothelial dysfunction, and maternal vascular damage. Ultimately, the maternal tissues and organs are affected, resulting in PE symptoms [56]. *ACAA1* is associated with peroxisomal lipid and fatty acid metabolism [57]. Dysregulated peroxisomal lipid metabolism can lead to excessive production of ROS and OS, potentially resulting in insufficient remodeling of the uterine spiral arteries, which is a hallmark feature of PE. Changes in the function or expression levels of these genes, which significantly affect the key mechanisms of PE, may contribute to the multiorgan dysfunction characteristics of this pregnancy-related disorder.

Additionally, our study suggests that these genes, along with their associated RNAs and proteins, may serve as important markers for the diagnosis and prognosis of PE. For example, the loss of *REPIN1* leads to notable changes in downstream target molecules, such as peroxisome proliferator-activated receptor γ (PPARγ) and glucose transporter type 2 (GLUT2) protein, as well as protein kinase B (Akt) phosphorylation and the mRNA expression of lipoprotein transporters 2, vesicle-associated membrane protein 4 (VAMP4), and synaptosome-associated protein 23 (SNAP23) [58]. In addition to the levels of related proteins in the maternal blood that can assist in the diagnosis and prognosis of PE, some researchers have recently identified that cell-free DNAs (cfDNAs) and cell-free RNAs (cfRNAs) in maternal plasma may also serve as potential biomarkers for this condition [59]. In particular, non-coding RNAs (ncRNAs) have attracted increasing attention in the context of PE. ncRNAs are a class of functional RNA molecules that do not encode proteins but play crucial roles in both pathological and physiological processes throughout various life cycle activities [60]. They primarily consist of microRNAs (miRNAs), long noncoding RNAs (lncRNAs), and circular RNAs (circRNAs), which may contribute to PE through various mechanisms including the inhibition of trophoblast migration, invasion, and proliferation [61]. For example, a study found that miR-567 inhibited cell proliferation, migration, and invasion by targeting the 3' untranslated region of *FGF5* [62]. Additionally, lnc-APPAT can inhibit cell proliferation and migration by interacting with miR-647 and *FGF5* [63]. However, circ-0001715 functions as a sponge for miR-1249-3p, promoting the upregulation of *FGF5* [64]. Therefore, the corresponding cfDNAs or cfRNAs in the maternal peripheral blood may provide significant value for the diagnosis and prognosis of PE.

However, several limitations must be considered when interpreting our results. First, because the current GWAS data do not specifically differentiate between early- and late-onset PE [65], our study included patients with both types of PE. As the GWAS database continues to evolve, we will further differentiate between early- and late-onset PE samples to reveal more accurate research findings. Second, most of the data in this study were derived from individuals of European ancestry, with a small cohort of Japanese participants. This limits the applicability of our

findings to other populations, particularly those in Africa and South America. In the future, as more GWAS databases become developed and enhanced, we will include a broader diversity of patients with PE from various racial and geographic backgrounds to achieve more comprehensive and reliable research outcomes. Third, although the DSigDB collects drug-related gene expression data from various publicly available sources, some genes may not have known interactions. Moreover, it is currently not feasible to further validate potential therapeutic targets for these genes using existing databases. However, in the future, the integration and enhancement of the DSigDB and other resources may lead to verifying these therapeutic targets more effectively.

Conclusions

In summary, we analyzed GWAS data from Europe and Asia to conduct a cross-ancestry genome-wide association study, identifying six novel susceptibility genes: *NPPA*, *SWAP70*, *NPR3*, *FGF5*, *REPIN1*, and *ACAA1*. Expression of these genes was associated with the risk of PE, and targeted drugs were related to *NPPA*, *NPR3*, and *REPIN1*. These findings provide new insights into the genetic architecture of PE and advance personalized medical approaches for its treatment. Further research and clinical trials on drugs targeting these genes are warranted.

Abbreviations

Akt	Protein kinase B
ARIC	Atherosclerosis Risk in Communities
BLISS	Biomarker Expression Level Imputation using Summary-level Statistics
cfDNAs	Cell-free DNAs
cfRNAs	Cell-free RNAs
eQTL	Expression quantitative trait loci
FDR	False discovery rate
FUMA	Functional Mapping and Annotation
FUSION	Functional Summary-based Imputation
GCTA	Genome-wide complex trait analysis
GLUT2	Glucose transporter type 2
GTEx	Genotype Tissue Expression
GWAS	Genome-wide association study
LD	Linkage disequilibrium
MAF	Minor allele frequency
mBAT	Multivariate set-based association test
OS	Oxidative stress
PE	Preeclampsia
PoPS	Polygenic priority score
PPARγ	Peroxisome proliferator-activated receptor γ
pQTL	Protein quantitative trait loci
PWAS	Proteome-wide association study
scRNA-seq	Single-Cell RNA Sequencing
SNPs	Single nucleotide polymorphisms
SNAP23	Synaptosome-associated protein 23
TWAS	Transcriptome-wide association study
UKB-PPP	UK Biobank Pharma Proteomics Project
VAMP4	Vesicle-associated membrane protein 4

Acknowledgements

We gratefully acknowledge the authors and participants of all GWASs including the UK Biobank, FinnGen, Biobank Japan, the GTEx v8 dataset web portal, UKB-PPP, deCODE Genetics, and the ARIC study, from which we utilized summary statistics data.

Author contributions

YS: Data analysis, drawing figures and tables, writing original draft. HH: Data analysis, drawing figures and tables. YC: Conceptualization, project administration, supervision.

Funding

This work was supported by National Natural Science Foundation of Shandong Province (No.ZR2022QH101).

Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 3 January 2025 / Accepted: 26 March 2025
Published online: 01 April 2025

References

1. Dimitriadis E, Rolnik DL, Zhou W, Estrada-Gutierrez G, Koga K, Francisco RPV, et al. Pre-eclampsia. Nat Rev Dis Primers. 2023;9:8.
2. Wu P, Green M, Myers JE. Hypertensive disorders of pregnancy. BMJ. 2023;381:e071653.
3. Cuckle H. Pre-eclampsia screening studies - overcoming intervention bias. BJOG. 2022;129:1318.
4. Shennan AH, Hurrell A. The evolving definition of pre-eclampsia. BJOG. 2021;128:1383.
5. Hauspurg A, Jeyabalan A. Postpartum preeclampsia or eclampsia: defining its place and management among the hypertensive disorders of pregnancy. Am J Obstet Gynecol. 2022;226:S1211–21.
6. Hofmeyr GJ. Why does pre-eclampsia exist? An evolutionary theory. Pregnancy Hypertens. 2021;24:124–5.
7. Jordao H, Herink K, Ka E, McVicker L, Kearns C, McMenamin C. Pre-eclampsia during pregnancy and risk of endometrial cancer: a systematic review and meta-analysis. BMC Womens Health. 2023;23:259.
8. Tyrmi JS, Kaartokallio T, Lokki AI, Jääskeläinen T, Kortelainen E, Ruotsalainen S, et al. Genetic risk factors associated with preeclampsia and hypertensive disorders of pregnancy. JAMA Cardiol. 2023;8:674–83.
9. Wang Y, Li B, Zhao Y. Inflammation in preeclampsia: genetic biomarkers, mechanisms, and therapeutic strategies. Front Immunol. 2022;13:883404.
10. Honigberg MC, Truong B, Khan RR, Xiao B, Bhatta L, Vy HMT, et al. Polygenic prediction of preeclampsia and gestational hypertension. Nat Med. 2023;29:1540–9.
11. Peterson RE, Kuchenbaecker K, Walters RK, Chen CY, Popejoy AB, Periyasamy S, et al. Genome-wide association studies in ancestrally diverse populations: opportunities, methods, pitfalls, and recommendations. Cell. 2019;179:589–603.
12. Stefanucci L, Collins J, Sims MC, Barrio-Hernandez I, Sun L, Burren OS, et al. The effects of pathogenic and likely pathogenic variants for inherited haemostasis disorders in 140 214 UK biobank participants. Blood. 2023;142:2055–68.
13. Kurki MI, Karjalainen J, Palta P, Sipilä TP, Kristiansson K, Donner KM, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. Nature. 2023;613:508–18.
14. Sakaue S, Kanai M, Tanigawa Y, Karjalainen J, Kurki M, Koshiba S, et al. A cross-population atlas of genetic associations for 220 human phenotypes. Nat Genet. 2021;53:1415–24.
15. The Genotype-Tissue. Expression (GTEx) project. Nat Genet. 2013;45:580–5.

16. Võsa U, Claringbould A, Westra HJ, Bonder MJ, Deelen P, Zeng B, et al. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. Nat Genet. 2021;53:1300–10.
17. Lin PW, Lin ZR, Wang WW, Guo AS, Chen YX. Identification of immune-inflammation targets for intracranial aneurysms: A multiomics and epigenome-wide study integrating summary-data-based Mendelian randomization, single-cell-type expression analysis, and DNA methylation regulation. Int J Surg. 2025;111:346–59.
18. Sun BB, Chiou J, Traylor M, Benner C, Hsu YH, Richardson TG, et al. Plasma proteomic associations with genetics and health in the UK biobank. Nature. 2023;622:329–38.
19. Ferkingstad E, Sulem P, Atlason BA, Sveinbjornsson G, Magnusson MI, Styrmisdottir EL, et al. Large-scale integration of the plasma proteome with genetics and disease. Nat Genet. 2021;53:1712–21.
20. Zhang J, Dutta D, Köttgen A, Tin A, Schlosser P, Grams ME, et al. Plasma proteome analyses in individuals of European and African ancestry identify cis-pQTLs and models for proteome-wide association studies. Nat Genet. 2022;54:593–602.
21. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010;26:2190–1.
22. Fu S, Wheeler W, Wang X, Hua X, Godbole D, Duan J, et al. A comprehensive framework for trans-ancestry pathway analysis using GWAS summary data from diverse populations. PLoS Genet. 2024;20:e1011322.
23. Lee Y, Yoon JW, Kim YA, Choi HJ, Yoon BW, Seo JH. A Genome-Wide association study of genetic variants of Apolipoprotein A1 levels and their association with vitamin D in Korean cohorts. Genes (Basel). 2022;13:1553.
24. Li Y, Dang X, Chen R, Teng Z, Wang J, Li S et al. Cross-ancestry genome-wide association study and systems-level integrative analyses implicate new risk genes and therapeutic targets for depression. Nat Hum Behav. 2025. https://doi.org/10.1038/s41562-024-02073-6
25. Gao B, Zhou X. MESuSiE enables scalable and powerful multi-ancestry fine-mapping of causal variants in genome-wide association studies. Nat Genet. 2024;56:170–9.
26. Yuan K, Longchamps RJ, Pardiñas AF, Yu M, Chen TT, Lin SC, et al. Fine-mapping across diverse ancestries drives the discovery of putative causal variants underlying human complex traits and diseases. Nat Genet. 2024;56:1841–50.
27. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4:7.
28. Yazar S, Alquicira-Hernandez J, Wing K, Senabouth A, Gordon MG, Andersen S, et al. Single-cell eQTL mapping identifies cell type-specific genetic control of autoimmune disease. Science. 2022;376:eabf3041.
29. Peng S, Deyssenroth MA, Narzo AFD, Lambertini L, Marsit CJ, Chen J, et al. Expression quantitative trait loci (eQTLs) in human placentas suggest developmental origins of complex diseases. Hum Mol Genet. 2017;26:3432–41.
30. Farries G, Bryan K, McGivney CL, McGettigan PA, Gough KF, Browne JA, et al. Expression quantitative trait loci in equine skeletal muscle reveals heritable variation in metabolism and the training responsive transcriptome. Front Genet. 2019;10:1215.
31. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BW, et al. Integrative approaches for large-scale transcriptome-wide association studies. Nat Genet. 2016;48:245–52.
32. Wei K, Lu Y, Ma X, Duan A, Lu X, Abdel-Shafy H et al. Transcriptome-Wide association study reveals potentially candidate genes responsible for milk production traits in Buffalo. Int J Mol Sci 2024;25:2626.
33. Liu Y, Chen S, Li Z, Morrison AC, Boerwinkle E, Lin X. A fast and powerful P value combination method for Rare-Variant analysis in sequencing studies. Am J Hum Genet. 2019;104:410–21.
34. Li A, Liu S, Bakshi A, Jiang L, Chen W, Zheng Z, et al. mBAT-combo: A more powerful test to detect gene-trait associations from GWAS data. Am J Hum Genet. 2023;110:30–43.
35. Diaz-Torres S, Lee SS, Ogonowski NS, Mackey DA, MacGregor S, Gharahkhani P, et al. Macular structural integrity estimates are associated with Parkinson's disease genetic risk. Acta Neuropathol Commun. 2024;12:130.
36. Weeks EM, Ulirsch JC, Cheng NY, Trippe BL, Fine RS, Miao J, et al. Leveraging polygenic enrichments of gene features to predict genes underlying complex traits and diseases. Nat Genet. 2023;55:1267–76.
37. Hemerich D, Svenstrup V, Obrero VD, Preuss M, Moscati A, Hirschhorn JN, et al. An integrative framework to prioritize genes in more than 500 loci associated with body mass index. Am J Hum Genet. 2024;111:1035–46.
38. Wu C, Zhang Z, Yang X, Zhao B. Large-scale imputation models for multi-ancestry proteome-wide association analysis. BioRxiv. 2023;2023(2010):2005–561120.
39. Gui J, Meng L, Huang D, Wang L, Yang X, Ding R, et al. Identification of novel proteins for sleep apnea by integrating genome-wide association data and human brain proteomes. Sleep Med. 2024;114:92–9.
40. Yoo M, Shin J, Kim J, Ryall KA, Lee K, Lee S, et al. DSigDB: drug signatures database for gene set analysis. Bioinformatics. 2015;31:3069–71.
41. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics. 2013;14:128.
42. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 2016;44:W90–97.
43. Xie Z, Bailey A, Kuleshov MV, Clarke DJB, Evangelista JE, Jenkins SL, et al. Gene set knowledge discovery with enrichr. Curr Protoc. 2021;1:e90.
44. Chen J, Spracklen CN, Marenne G, Varshney A, Corbin LJ, Luan J, et al. The trans-ancestral genomic architecture of glycemic traits. Nat Genet. 2021;53:840–60.
45. Chen MH, Raffield LM, Mousas A, Sakaue S, Huffman JE, Moscati A, et al. Trans-ethnic and Ancestry-Specific Blood-Cell genetics in 746,667 individuals from 5 global populations. Cell. 2020;182:1198–e12131114.
46. Fu J, Zhang Q, Wang J, Wang M, Zhang B, Zhu W, et al. Cross-ancestry genome-wide association studies of brain imaging phenotypes. Nat Genet. 2024;56:1110–20.
47. Liu N, Zhang L, Tian T, Cheng J, Zhang B, Qiu S, et al. Cross-ancestry genome-wide association meta-analyses of hippocampal and subfield volumes. Nat Genet. 2023;55:1126–37.
48. Volpe M. Natriuretic peptides and cardio-renal disease. Int J Cardiol. 2014;176:630–9.
49. Armstrong DW, Tse MY, O'Tierney-Ginn PF, Wong PG, Ventura NM, Janzen-Pang JJ, et al. Gestational hypertension in atrial natriuretic peptide knockout mice and the developmental origins of salt-sensitivity and cardiac hypertrophy. Regul Pept. 2013;186:108–15.
50. Zou X, Wang L, Wang S, Zhang Y, Ma J, Chen L, et al. Promising therapeutic targets for ischemic stroke identified from plasma and cerebrospinal fluid proteomes: a multicenter Mendelian randomization study. Int J Surg. 2024;110:766–76.
51. 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;11:5976.
52. Xin Q, Han Y, Jiang W, Wang J, Luan Y, Ji Q, et al. Genetic susceptibility analysis of FGF5 polymorphism to preeclampsia in Chinese Han population. Mol Genet Genomics. 2022;297:791–800.
53. Changalidis AI, Maksiutenko EM, Barbitoff YA, Tkachenko AA, Vashukova ES, Pachuliia OV et al. Aggregation of Genome-Wide association data from FinnGen and UK biobank replicates multiple risk loci for pregnancy complications. Genes (Basel). 2022;13:2255.
54. Xia Y, Ge G, Xiao H, Wu M, Wang T, Gu C, et al. REPIN1 regulates iron metabolism and osteoblast apoptosis in osteoporosis. Cell Death Dis. 2023;14:631.
55. Gumilar KE, Priangga B, Lu CH, Dachlan EG, Tan M. Iron metabolism and ferroptosis: A pathway for Understanding preeclampsia. Biomed Pharmacother. 2023;167:115565.
56. Zhang Y, Lu Y, Jin L. Iron metabolism and ferroptosis in physiological and pathological pregnancy. Int J Mol Sci 2022;23:9395.
57. Luo R, Fan Y, Yang J, Ye M, Zhang DF, Guo K, et al. A novel missense variant in ACAA1 contributes to early-onset Alzheimer's disease, impairs lysosomal function, and facilitates amyloid-β pathology and cognitive decline. Signal Transduct Target Ther. 2021;6:325.
58. Kern M, Kosacka J, Hesselbarth N, Brückner J, Heiker JT, Flehmig G, et al. Liver-restricted Repin1 deficiency improves whole-body insulin sensitivity, alters lipid metabolism, and causes secondary changes in adipose tissue in mice. Diabetes. 2014;63:3295–309.
59. León-Martínez D, Lynn T, Abrahams VM. Cell-free fetal DNA impairs trophoblast migration in a TLR9-dependent manner and can be reversed by hydroxychloroquine. J Reprod Immunol. 2023;157:103945.
60. Munjas J, Sopić M, Stefanović A, Košir R, Ninić A, Joksić I et al. Non-Coding RNAs in Preeclampsia-Molecular mechanisms and diagnostic potential. Int J Mol Sci 2021;22:10652.

61. Ogoyama M, Takahashi H, Suzuki H, Ohkuchi A, Fujiwara H, Takizawa T. Non-Coding RNAs and Prediction of Preeclampsia in the First Trimester of Pregnancy. Cells. 2022; 11.
62. Liu D, Zhang C, Li X, Zhang H, Pang Q, Wan A. MicroRNA-567 inhibits cell proliferation, migration and invasion by targeting FGF5 in osteosarcoma. Excli J. 2018;17:102–12.
63. Meng F, Han L, Liang Q, Lu S, Huang Y, Liu J. The Lnc-RNA APPAT suppresses human aortic smooth muscle cell proliferation and migration by interacting with MiR-647 and FGF5 in atherosclerosis. J Endovasc Ther. 2023;30:937–50.
64. Zhang Q, Ding F, Zhang C, Han X, Cheng H. Circ_0001715 functions as a miR-1249-3p sponge to accelerate the progression of Non-small cell lung cancer via upregulating the level of FGF5. Biochem Genet. 2023;61:1807–26.
65. 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.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.