

Genome-wide association study reveals HSF2, GJA1 and TRIM36 as susceptibility genes for preeclampsia: a community-based population study in Tianjin, China

Zhenhua Cao, Li Ma, Wei Cai, Xiulong Niu, Ning Yang, Jianmei Ni, Xiaojing Wang, Maoti Wei, Shaobo Chen & Yuming Li

To cite this article: Zhenhua Cao, Li Ma, Wei Cai, Xiulong Niu, Ning Yang, Jianmei Ni, Xiaojing Wang, Maoti Wei, Shaobo Chen & Yuming Li (2023) Genome-wide association study reveals HSF2, GJA1 and TRIM36 as susceptibility genes for preeclampsia: a community-based population study in Tianjin, China, *Hypertension in Pregnancy*, 42:1, 2256863, DOI: [10.1080/10641955.2023.2256863](https://doi.org/10.1080/10641955.2023.2256863)

To link to this article: <https://doi.org/10.1080/10641955.2023.2256863>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 21 Sep 2023.



Submit your article to this journal [↗](#)



Article views: 840



View related articles [↗](#)



View Crossmark data [↗](#)

Genome-wide association study reveals HSF2, GJA1 and TRIM36 as susceptibility genes for preeclampsia: a community-based population study in Tianjin, China

Zhenhua Cao^{a,b}, Li Ma^c, Wei Cai^d, Xiulong Niu^d, Ning Yang^e, Jianmei Ni^b, Xiaojing Wang^d, Maoti Wei^e, Shaobo Chen^d, and Yuming Li^e

^aNankai Hospital, Tianjin Medical University, Tianjin, China; ^bDepartment of cardiology, Nankai Hospital, Tianjin, China; ^cDepartment of Obstetrics, The First Teaching Hospital of Tianjin University of Traditional Chinese Medicine, Tianjin, China; ^dInstitute of Cardiovascular Disease and Heart Center, Characteristic Medical Center of the Chinese People's Armed Police Force, Tianjin, China; ^eDepartment of Cardiology, Tianjin Economic-Technological Development Area (TEDA) International Cardiovascular Hospital, Tianjin, China

ABSTRACT

Background: Preeclampsia (PE) mainly occurs in pregnant women and is hereditary. Several genome-wide association studies (GWAS) on Caucasian samples have reported some gene loci that are associated with preeclampsia. However, these studies have not reached consistent conclusions. No previous GWAS has examined preeclampsia in the Chinese Han population.

Method: This study aimed to identify common genetic variations associated with preeclampsia in the Chinese Han population through two-stage case-control studies. The discovery cohort included 92 patients with severe preeclampsia and 187 healthy controls. The validation cohort included 52 patients with preeclampsia and 104 controls. A genome-wide association study was performed to identify putative preeclampsia genes in the discovery cohort, with validation in the validation cohort.

Results: In the discovery cohort, GWAS demonstrated that 19 single-nucleotide polymorphisms (SNPs) were associated with preeclampsia ($P < 10^{-5}$). The pathway analysis revealed that these 19 SNP representative genes were mainly enriched in the adenylyl cyclase-inhibiting G-protein coupled receptor signaling pathway. After validation in the validation cohort, rs13176432 and rs13210237 remained closely related to preeclampsia ($P < 0.05$). In the combined data set, the frequency of the G allele in rs13176432 was significantly higher in cases with preeclampsia than in controls ($P = 5 \times 10^{-6}$). The frequency of the A allele in rs13210237 was higher in the preeclampsia group ($P = 8 \times 10^{-6}$). The rs13210237 representative genes include HSF2 and GJA1, while the rs13176432 representative gene is TRIM36. There were no differences in genotype distribution between the early-onset and late-onset preeclampsia groups ($P > 0.05$). Furthermore, rs13210237 and rs13176432 were related to preeclampsia in the adjusted regression model ($P < 0.000$).

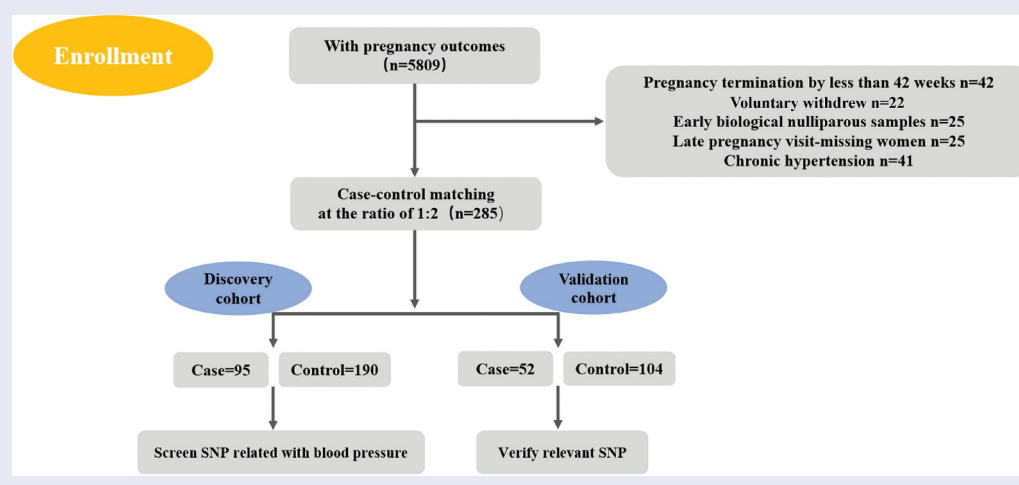
Conclusion: In this study of two independent cohorts, we found that rs13210237 and rs13176432 might be novel preeclampsia-susceptible genetic factors in the Han population in China. However, there was no association between the onset of preeclampsia and these genotypes.

ARTICLE HISTORY

Received 11 November 2022
Accepted 4 September 2023

KEYWORDS

Genome-wide association study (GWAS); single-nucleotide polymorphisms (SNPs); preeclampsia (PE)



CONTACT Zhenhua Cao  czh_man@126.com  Nankai Hospital, Tianjin Medical University, No. 6, Changjiang Road, Tianjin, Nankai District, China; Yuming Li  cardiolab@live.com  Department of Cardiology, Tianjin Economic-Technological Development Area (TEDA) International Cardiovascular Hospital, No. 61, Third Street, Binhai New Area, Tianjin, China

*Zhenhua Cao and Li Ma are contributed equally to this study.

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

Preeclampsia is a devastating disease for pregnant women, with an incidence rate of approximately 5% (1). This disease severely impacts maternal health and could cause renal failure, liver injury, intracranial hemorrhage, pulmonary edema and even death. For perinatal infants, preeclampsia may also cause premature delivery or intrauterine growth restriction.

To date, the pathogenesis of preeclampsia remains unclear. Previous studies have shown that preeclampsia has a characteristic of familial aggregation (2,3), indicating that genetic factors may play an important role in preeclampsia. Numerous genome-wide association studies (GWAS) in preeclampsia have found some candidate risk genes, such as *STOX1* (4) and *ACVR2A* (5). However, GWAS results are not replicable due to the distinction of study design and ethnic background (6,7). More importantly, no GWAS data are available for Chinese patients with preeclampsia. Several studies have verified specific preeclampsia loci in Western populations in Chinese populations, but very few susceptible sites have been verified (8,9), suggesting that Chinese preeclampsia patients may have their own genetic features.

In this project, we enrolled patients with preeclampsia from a large-scale community-based Chinese Han population and performed two-stage case-control studies to identify novel genetic risk loci for preeclampsia.

Subject and method

Study subjects

All participants were recruited from 19 community hospitals in Tianjin, a city in northern China. From 1 November 2016, to 30 May 2018, we enrolled 5809 pregnant women with pregnancy outcomes and identified 147 preeclampsia patients. These preeclampsia patients were divided into “discovery” and “validation” cohorts. Specifically, the discovery cohort included 95 severe preeclampsia patients and 190 normal pregnant women, case-control matching was performed by age, height and weight at a ratio of 1:2. The validation cohort included 52 preeclampsia patients and 104 normal pregnant women same as before. Each of the participants signed an informed consent form, and this study was approved by the Ethics Committee of Armed Police Logistics College.

All pregnant women were recruited during the 11 ~ 13 weeks after pregnancy, and simultaneously collecting venous blood. The subjects were excluded if they refused to sign the informed consent, planned to terminate the pregnancy, had primary hypertension or communication difficulty or neurological disorder, or had congenital heart disease or connective tissue disease. preeclampsia is

determined by increased blood pressure (gestational blood pressure elevation) accompanied by proteinuria after 20 weeks of gestation. High gestational blood pressure was defined as a blood pressure ≥ 140 mmHg systolic or ≥ 90 mmHg diastolic in a woman who was normotensive before 20 weeks of gestation. Severe preeclampsia was diagnosed based on the following criteria: blood pressure ≥ 160 mmHg systolic or ≥ 110 mmHg diastolic; proteinuria of ≥ 3.0 g in 24 hours (2+ or 3+ on qualitative examination), which should occur for the first time during pregnancy and regress after delivery; increased serum creatinine levels (>1.2 mg/dL unless known to be previously elevated); platelet count $< 100,000$ cells/mm³; evidence of microangiopathic hemolytic anemia (with increased lactic acid dehydrogenase concentration); elevated hepatic enzyme levels (either alanine aminotransferase, aspartate aminotransferase, or both); patient reports of persistent headache or other cerebral or visual disturbances; and patient reports of persistent epigastric pain.

GWAS for exploring genotypes of preeclampsia

Collection and extraction of samples and DNA preparation

A total of 5 mL peripheral venous blood was drawn from each participant, placed in a pro-coagulation tube, centrifuged for 10 min at 2500 rpm and stored in a -80°C refrigerator until use. The procedure was completed within one hour.

DNA extraction

Genomic DNA was extracted using an OMEGA Blood DNA Kit (Cat# D3392-02, OMEGA) following the manufacturer's instructions. Potential RNA contamination was removed by RNaseA (Cat#19101, QIAGEN, GmBH, Germany). We used a Nano Drop (Cat#ND-2000, Thermo Fisher, Waltham, MA, US) and performed 1% agarose gel electrophoresis to check the quantity and quality of purified DNA.

DNA amplification, fragmentation and precipitation, drying, resuspension and QC

Total genomic DNA (200 ng) was amplified and randomly fragmented into 25 to 125 base pair (bp) fragments. These fragments are purified, resuspended, and hybridized to Axiom Array Plates. Quantitation and fragmentation QC checks were performed by NanoDrop-2000 and E-Gels.

Chip hybridization

Hybridization buffers were prepared, and array hybridization was performed according to the Affymetrix Axiom 2.0

Assay Manual Workflow (Cat# 901758, Affymetrix, Santa Clara, CA, US). Manual Preparation of Ligation, Staining, and Stabilization Reagent Trays for the GeneTitan MC Instrument. The instrument will automatically complete the array hybridization and washing steps.

Data acquisition

The CEL file is added to the Axiom™ Analysis Suite software after processing the DNA chips. The Summary window tab displays a summary snapshot of the analysis, including detailed threshold values and tables based on the analysis.

Data quality control

Before analysis, the following quality control and filtration options were available: (1) removal of sites with a successful typing rate of less than 90% (the call rate of sites < 90%), the removal of sites with a minimal gene frequency of less than 0.001, the removal of sites inconsistent with Hardy-Weinberg equilibrium (1×10^{-6}), or the removal of SNPs on X, Y and linear chromosomes; (2) Plink 1.07 software was used to calculate the identity by state (IBS) between two sites; and (3) Plink 1.07 software was used for principal component analysis (PCA) to ensure good case-control matching without any population stratification.

Validation method

Multiplex PCR and sequencing

A panel that contains 19 target SNP sites was designed. Library preparation was performed by two-step PCR. The first round of PCR was set up as follows: DNA (10 ng/μl) 2 μl; amplicon PCR forwards primer mix (10 μM) 1 μl; amplicon PCR reverse primer mix (10 μM) 1 μl; 2×PCR Ready Mix 15 μl (total 25 μl) (Kapa HiFi Ready Mix). The plate was sealed, and PCR was performed in a thermal instrument (Bio-Rad, T100™) using the following program: 1 cycle of denaturing at 98°C for 5 min; the first 8 cycles of denaturing at 98°C for 30 s, annealing at 50°C for 30 s, and elongation at 72°C for 30 s; then, 25 cycles of denaturing at 98°C for 30 s, annealing at 66°C for 30 s, and elongation at 72°C for 30 s; and a final extension at 72°C for 5 min. Finally, the PCR products were checked using electrophoresis in 1% (w/v) agarose gels in TBE buffer (Tris, boric acid, EDTA) stained with ethidium bromide (EB) and visualized under UV light. Then, we used AMPure XP beads to purify the amplicon product. Afterwards, the second round of PCR was performed. PCR was set up as follows: DNA (10 ng/μl) 2 μl; universal P7 primer with barcode (10 μM) 1 μl; universal P5 primer (10 μM) 1 μl; 2× PCR Ready Mix 15 μl (total 30 μl) (Kapa HiFi Ready Mix). The plate was sealed, and PCR was performed in a thermal instrument (Bio-Rad, T100™) using the following program: 1 cycle of denaturing at

98°C for 3 min, then 5 cycles of denaturing at 94°C for 30 s, annealing at 55°C for 20 s, elongation at 72°C for 30 s, and a final extension at 72°C for 5 min. Then, we used AMPure XP beads to purify the amplicon product. The libraries were then quantified and pooled. Paired-end sequencing of the library was performed on HiSeq XTen sequencers (Illumina, San Diego, CA).

Data QC and SNP calling

Raw reads were filtered according to two steps: 1) adaptor sequences were removed if reads were contained by cutadapt (v 1.2.1); 2) low-quality bases were removed from reads 3' to 5' ($Q < 20$) by PRINSEQ-lite (v 0.20.3); and the remaining clean data were mapped to the reference genome by BWA (version 0.7.13-r1126) with default parameters. Samtools (Version: 0.1.18) was used to calculate each genotype of the target site. ANNOVAR (16 April 2018) was used to detect genetic variants.

Statistical analysis

GWAS statistical analysis and data drawing: SNP numbers, chromosome numbers, physical locations and *P* values in the genome-wide Cochran-Armitage test for trend were entered in R software (version 4.1.0, <https://www.R-project.org>) to generate chromosome maps and quantile-quantile plots for genome-wide association analysis; the chi-square test results for trend of SNPs in all regions generated chromosome maps for corresponding regions; a statistical analysis was made for test data by SPSS (version 26.00, Armonk, NY: IBM Corp), and the R x C contingency table chi-square test was conducted to compare the genotype and allele frequency, which was expressed by odds ratio (OR) and 95% confidence interval (CI). The obtained data were checked to determine whether they conformed to the Hardy-Weinberg genetic equilibrium law. The analysis revealed that these data were all significantly different, with *P* values less than 0.05. The GO analysis was completed using KOBAS software.

Results

Entry process and clinical characteristics

Figure 1 shows the entry process of all cases and controls for this project. We enrolled 5809 pregnant women with pregnancy outcomes from the community population, of whom 42 women terminated the pregnancy at less than 20 weeks and 22 women voluntarily withdrew. Twenty-five early biological nulliparous samples were available, and 25 late-pregnancy women missed the follow-ups. For the discovery

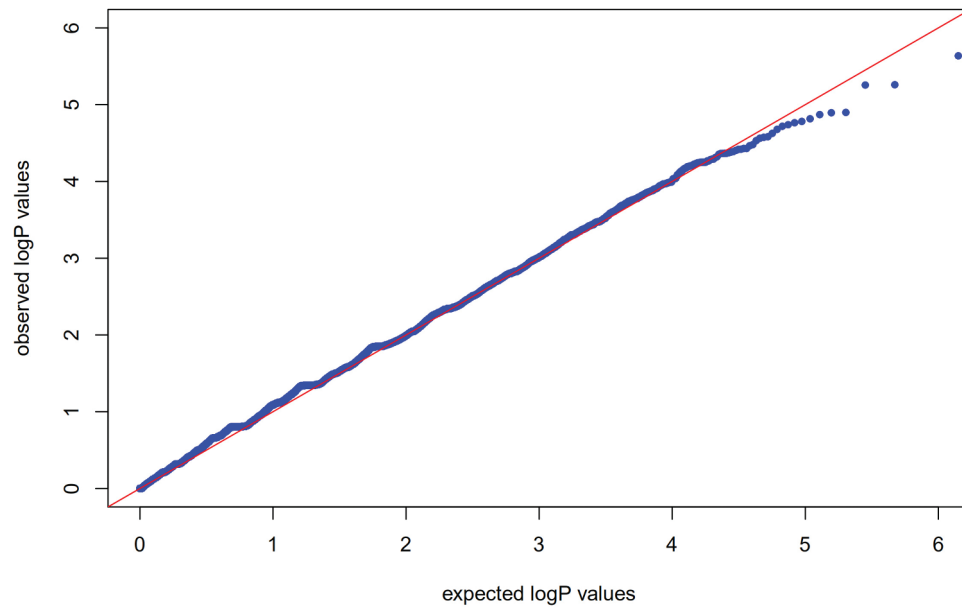


Figure 1. Flow chart of the study subject recruitment process.

Table 1. Summary of patient characteristics in discovery cohort.

	case(n=92)	control(n=187)	<i>p</i>
Age	31.25 ± 3.87	31.62 ± 3.79	0.45
Height (cm)	162.65 ± 5.95	163.13 ± 4.93	0.48
Weight (kg)	65.58 ± 13.95	63.05 ± 11.10	0.13
Gravidity	1.65 ± 1.02	1.84 ± 1.10	0.16
Parity	0.20 ± 0.43	0.29 ± 0.47	0.09
Initial systolic blood pressure (mmHg)	113.89 ± 15.29	111.91 ± 10.82	0.27
Initial diastolic blood pressure (mmHg)	72.75 ± 9.96	70.51 ± 8.50	0.07
Red blood cell ($\times 10^{12}/L$)	4.43 ± 0.55	4.25 ± 0.35	0.00
White blood cell ($\times 10^9/L$)	8.60 ± 2.04	8.24 ± 1.86	0.15
Hemoglobin (g/L)	128.64 ± 11.69	126.62 ± 9.94	0.13
Platelet ($\times 10^9/L$)	247.48 ± 61.89	244.27 ± 59.16	0.68
Fasting blood glucose (mmol/L)	4.91 ± 0.64	4.54 ± 0.80	0.00
Total protein (g/L)	69.75 ± 3.73	67.41 ± 3.79	0.00
Albumin (g/L)	43.03 ± 3.16	42.02 ± 2.47	0.03
Alanine aminotransferase	27.50 ± 28.75	23.01 ± 14.83	0.16
Urea nitrogen (mmol/L)	3.17 ± 0.70	3.25 ± 2.42	0.74
Creatinine ($\mu\text{mol/L}$)	58.35 ± 11.40	63.38 ± 16.08	0.00
Uric acid ($\mu\text{mol/L}$)	229.16 ± 61.96	218.34 ± 52.79	0.13
Gestational age of delivery	38.31 ± 2.14	39.39 ± 1.25	0.00
Weight before delivery	77.28 ± 13.46	75.71 ± 11.90	0.33
Neonatal weight	3008.02 ± 766.22	3260.58 ± 507.94	0.01
Length of newborn	48.04 ± 6.01	49.57 ± 1.95	0.02
42 days postpartum systolic blood pressure (mmHg)	117.34 ± 14.16	113.63 ± 10.14	0.04
42 days postpartum diastolic blood pressure (mmHg)	75.99 ± 11.39	74.65 ± 8.17	0.35

cohort, cases and controls were matched by age, height and weight at a ratio of 1:2 (95 cases and 190 controls). During the subsequent GWAS analysis, there were three patients and three controls with a call ratio less than 90%. After removal of these six samples, 92 patients and 187 controls were finally included in the discovery cohort. Case-control matching was performed by age, height and weight at a ratio of 1:2, and 52 cases were included in the case group and 104 controls in the control group.

In the discovery cohort, there were no obvious differences between the case group and control group regarding age, height or weight. The initial red blood cell count, total protein, albumin, fasting blood glucose and systolic blood pressure on the 42nd day after delivery in the case group were all higher than those in the control group, while creatinine, gestational weeks before delivery, newborn weight and newborn length were lower than those in the control group (Table 1).

Table 2. Summary of patient characteristics in validation cohort.

	case(n=52)	control(n=104)	<i>p</i>
Age	30.83 ± 3.99	30.73 ± 3.34	0.87
Height(cm)	163.39 ± 5.49	161.82 ± 16.60	0.51
Weight(kg)	64.55 ± 11.27	63.13 ± 10.12	0.43
Gravidity	1.54 ± 0.92	1.93 ± 0.96	0.02
Parity	0.12 ± 0.32	0.43 ± 0.54	0.00
Initial systolic blood pressure(mmHg)	111.62 ± 14.48	105.46 ± 9.31	0.01
Initial diastolic blood pressure(mmHg)	72.69 ± 9.67	66.56 ± 7.34	0.00
Red blood cell ($\times 10^{12}/L$)	4.29 ± 0.44	4.24 ± 0.35	0.53
White blood cell ($\times 10^9/L$)	8.06 ± 1.57	7.76 ± 1.67	0.28
Hemoglobin (g/L)	129.81 ± 11.26	127.80 ± 10.49	0.27
Platelet ($\times 10^9/L$)	231.63 ± 64.84	229.34 ± 52.19	0.81
Fasting blood glucose (mmol/L)	4.95 ± 0.65	4.80 ± 0.44	0.10
Total protein (g/L)	69.23 ± 3.73	69.90 ± 4.60	0.48
Albumin (g/L)	43.79 ± 2.38	44.09 ± 3.16	0.64
Alanine aminotransferase	28.65 ± 29.54	21.74 ± 26.62	0.14
Urea nitrogen (mmol/L)	2.95 ± 0.57	3.03 ± 0.62	0.47
Creatinine ($\mu\text{mol/L}$)	53.19 ± 10.62	51.46 ± 11.93	0.38
Uric acid($\mu\text{mol/L}$)	219.48 ± 48.97	210.04 ± 50.08	0.27
Gestational age of delivery	38.31 ± 2.14	39.39 ± 1.25	0.00
Weight before delivery	75.88 ± 10.46	77.57 ± 22.10	0.60
Neonatal weight	3055.77 ± 606.84	3418.21 ± 443.57	0.00
Length of newborn	48.21 ± 7.07	50.27 ± 1.89	0.05
42 days postpartum systolic blood pressure(mmHg)	115.24 ± 10.19	109.40 ± 10.04	0.00
42 days postpartum diastolic blood pressure (mmHg)	74.72 ± 7.81	70.48 ± 7.42	0.00

In the validation cohort, there were no obvious differences between the case group and the control group regarding age, height or weight. The gravidity, parity, gestational weeks before delivery, and newborn weight in the case group were lower than those in the control group, while the initial systolic blood pressure, initial diastolic blood pressure, systolic blood pressure and diastolic blood pressure on the 42nd day after delivery were higher than those in the control group (Table 2).

GWAS results

The 95 preeclampsia patients and 190 controls in the discovery cohort were genotyped by an Affymetrix Axiom 2.0 Assay chip. After quality control, the call rate of three patients and three controls was less than 90%. After removing these six samples, GWAS was performed for the remaining 92 patients and 187 controls. PCA showed that there was no population stratification between the case group and the control group (Fig. S1). Q-Q plot analysis indicated that the observed distribution of *p* values for genotyped SNPs exhibited minimal deviation from the expected distribution (Figure 2). In total, the frequencies of 19 SNPs were significantly different between the cases and controls ($P < 10^{-5}$ Figure 3 and Table 3). GO enrichment analysis indicated that preeclampsia-related genes were mainly enriched in the adenylate cyclase-inhibiting G-protein coupled receptor signaling pathway (Figure 4).

Validation cohort

To verify the initial screening result of genome-wide genotyping, 19 SNPs discovered in the discovery cohort were verified in the independent validation cohort. The minor allele frequency of rs886043316 was 0 in both the case group and the control group, so this SNP was removed. The verified results of the remaining 18 SNPs are shown in Table 4. We found that two SNPs (rs13176432 and rs13210237) were still associated with preeclampsia at the validation stage ($P < 0.05$). We combined the results of the discovery cohort and validation cohort, and two SNP sites were more significantly different in the combined cohort, namely, rs13176432 ($P = 5 \times 10^{-6}$) and rs13210237 ($P = 8 \times 10^{-6}$). Therefore, these two sites are newly discovered genetic susceptibility sites associated with the morbidity of preeclampsia in the Chinese Han population.

After two-stage verification, rs13176432 and rs13210237 were significantly associated with preeclampsia in the Chinese Han population. The mutation site of rs13176432 is G, and the morbidity risk of preeclampsia with the G allele is 2.054 times that of the control group (95% CI: 1.315–3.209). The mutation site of rs13210237 is A, and the morbidity risk of preeclampsia with the A allele is 1.431 times that of the control group (95% CI: 1.179–1.736) (Table 5).

Genotype distribution with preeclampsia onset

Preeclampsia was defined as early or late onset based on the onset time; onset before 34 weeks was defined as

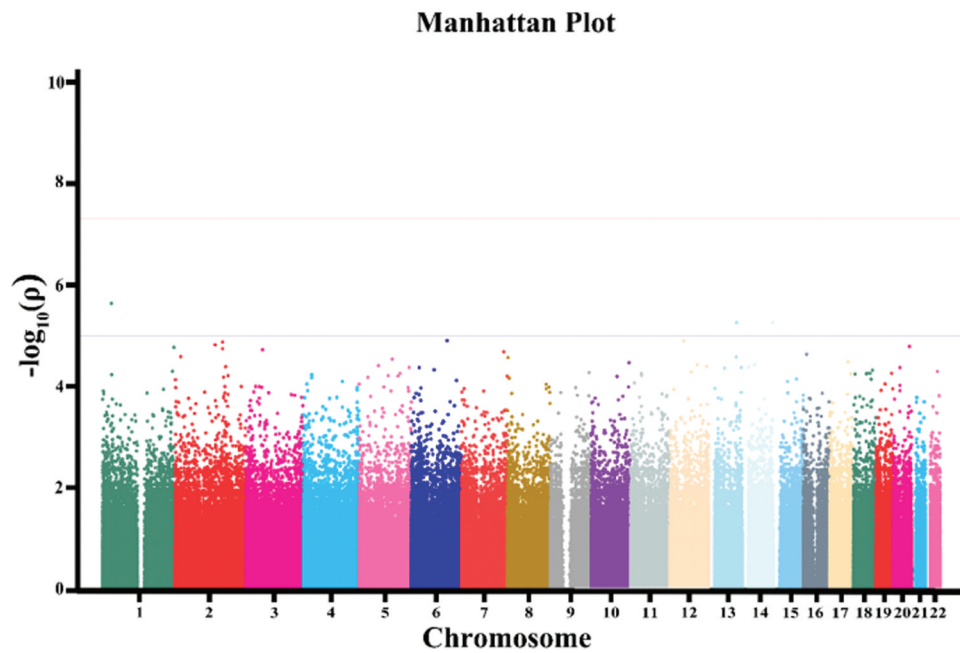


Figure 2. Q-Q plots with P values under the expected zero distribution relative to theoretical P values.

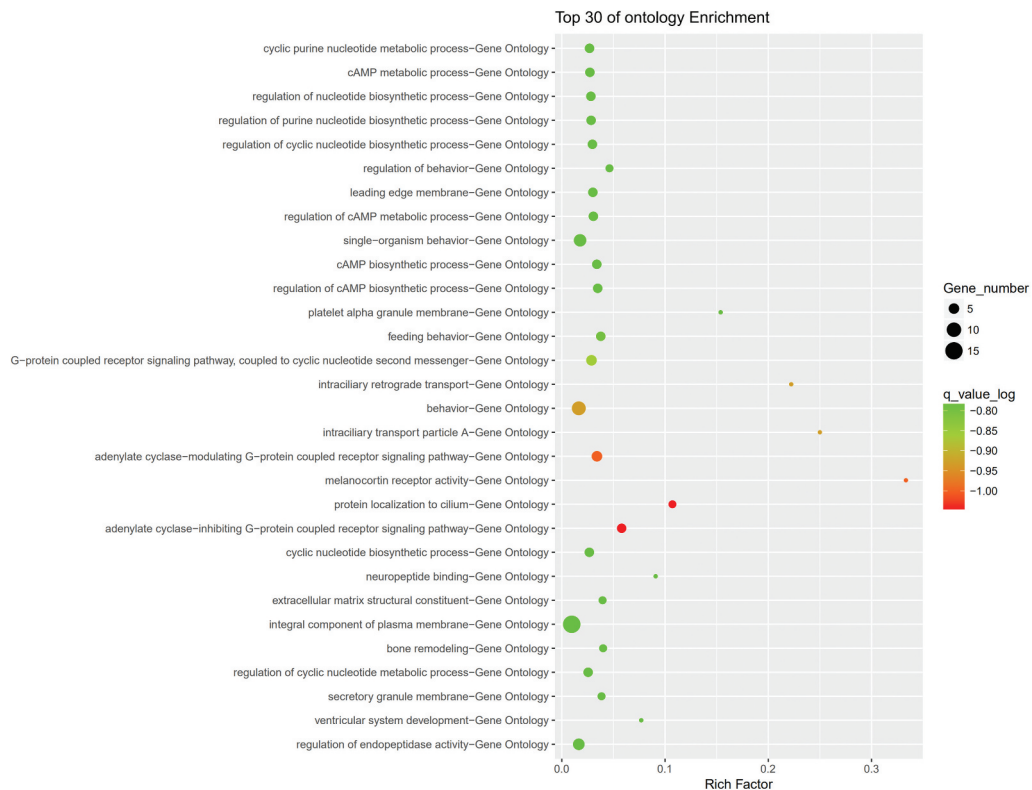


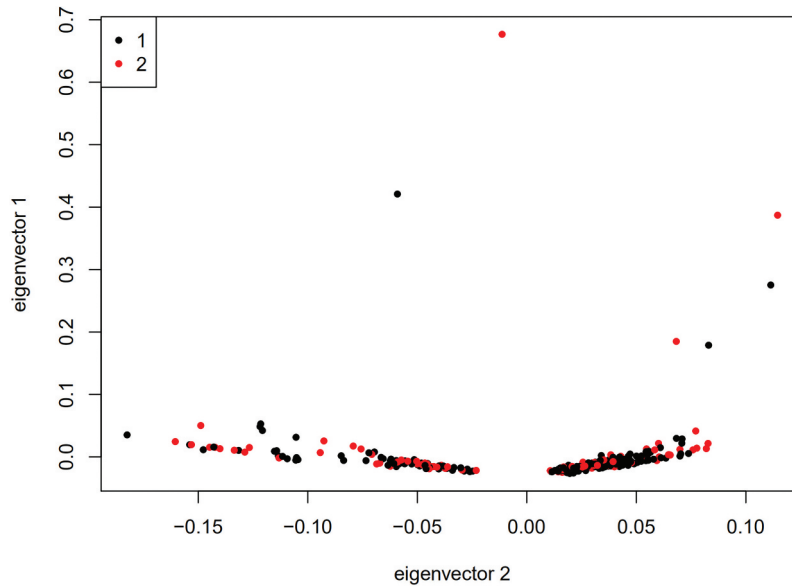
Figure 3. Manhattan chart at the preliminary GWAS discovery stage of severe preeclampsia.

early-onset preeclampsia ($n = 33$), and onset after 34 weeks was defined as late-onset preeclampsia ($n = 111$). Comparing genotypes between early-onset and late-

onset preeclampsia, there was no difference between the early-onset and late-onset preeclampsia genotypes. (Table 6)

Table 3. The initial GWAS result in the discovery stage.

CHR	dbSNPID	POSITION	gene	gene_relationship	A1	A2	MAF		OR	P (fisher)
							case	control		
13	rs9523993	93930403	GPC6	genic	G	A	0.40	0.23	2.219	3.59×10^{-5}
2	rs886043316	20130280	WDR35	genic	A	C	0.10	0.02	5.216	7.07×10^{-5}
2	rs78613705	163507406	KCNH7	genic	A	G	0.19	0.07	3.165	2.69×10^{-5}
1	rs77356340	245842272	KIF26B	genic	A	G	0.09	0.01	7.052	4.42×10^{-5}
2	rs76289759	163460529	KCNH7	genic	A	G	0.27	0.13	2.575	3.30×10^{-5}
13	rs75154137	95626009	LINC00557	intergenic	T	C	0.09	0.01	7.604	1.61×10^{-5}
20	rs73299501	56168784	ZBP1,PCK1	intergenic	A	G	0.15	0.05	3.622	3.61×10^{-5}
12	rs73095909	47679298	PCED1B,MIR4494	intergenic	G	A	0.13	0.03	4.41	3.13×10^{-5}
10	rs7068679	130318498	LINC01163,MGMT	intergenic	A	G	0.13	0.03	4.171	6.99×10^{-5}
17	rs56309431	63645289	CEP112	genic	T	G	0.60	0.42	2.108	4.07×10^{-5}
14	rs148761074	104502811	TDRD9	genic	T	C	0.08	0.01	10.67	1.49×10^{-5}
1	rs147968844	30829270	LOC101929406,MATN1	intergenic	T	C	0.06	0.00	NA	4.82×10^{-5}
3	rs146895271	57582163	ARF4	genic	G	T	0.11	0.02	5.34	6.15×10^{-5}
2	rs146718393	137647360	THSD7B	genic	A	G	0.09	0.01	7.105	4.07×10^{-5}
6	rs13210237	122346667	HSF2,GJA1	intergenic	A	G	0.24	0.10	2.793	2.54×10^{-5}
5	rs13176432	114369723	TRIM36	intergenic	G	T	0.11	0.02	4.85	6.67×10^{-5}
8	rs117487811	1647942	DLGAP2	genic	A	C	0.09	0.02	5.576	9.11×10^{-5}
16	rs117222557	11147978	CLEC16A	genic	A	G	0.10	0.02	5.642	8.75×10^{-5}
7	rs10229442	146503040	CNTNAP2	genic	T	C	0.05	0.00	NA	4.76×10^{-5}

**Figure 4.** GO bubble analysis regarding preeclampsia-related pathways.**Table 4.** SNP screened from validation cohort.

chr	SNP	BP	A1	A2	MAF		CHISQ	P	OR	95% Confidence Interval	
					case	control				Lower	Upper
13	rs9523993	93930403	G	A	0.28	0.27	0.01	0.929	1.02	0.61	1.73
2	rs886043316	20130280	C	C	0.00	0.00			NS		
2	rs78613705	163507406	A	G	0.07	0.13	2.44	0.118	0.51	0.21	1.21
1	rs77356340	245842272	A	G	0.01	0.02		0.668	0.39	0.05	3.42
2	rs76289759	163460529	A	G	0.13	0.19	1.95	0.163	0.62	0.31	1.22
13	rs75154137	95626009	T	C	0.01	0.03	1.17	0.431	0.33	0.04	2.75
20	rs73299501	56168784	A	G	0.05	0.04	0.16	0.766	1.26	0.40	3.96
12	rs73095909	47679298	G	A	0.04	0.05	0.32	0.575	0.72	0.22	2.31
10	rs7068679	130318498	A	G	0.06	0.03	1.00	0.371	1.76	0.58	5.37
17	rs56309431	63645289	T	G	0.50	0.51		0.936	0.98	0.61	1.57
14	rs148761074	104502811	T	C	0.02	0.04		0.505	0.49	0.10	2.35
1	rs147968844	30829270	T	C	0.01	0.03		0.431	0.33	0.04	2.75
3	rs146895271	57582163	G	T	1.00	0.99		0.553	0.00		
2	rs146718393	137647360	A	G	0.01	0.01		1.000	0.66	0.07	6.46
6	rs13210237	122346667	A	G	0.16	0.09	4.12	0.042	2.06	1.01	4.19
5	rs13176432	114369723	G	T	0.07	0.02		0.047	0.27	0.08	0.95
8	rs117487811	1647942	A	C	0.04	0.04		1.000	0.88	0.27	2.94
16	rs117222557	11147978	A	G	0.05	0.06	0.12	0.724	0.83	0.28	2.41
7	rs10229442	146503040	T	C	0.02	0.01		0.259	4.06	0.36	45.29

Table 5. Consolidation calculation of two stages.

Chr	SNP	BP	A1	A2	case (n=144)	control (n=291)	CHISQ	P	OR	95% Confidence Interval	
										Lower	Upper
5	rs13176432	114369723	G	T	0.090	0.022		5×10^{-6}	2.054	1.315	3.209
6	rs13210237	122346667	A	G	0.205	0.096	19.823	8×10^{-6}	1.431	1.179	1.736

Table 6. Comparison of genotypes between early-onset and late-onset preeclampsia.

Genotype and model	Early-onset (n=33)	Late-onset (n=111)	χ^2	P
rs13176432				
GG	0	2	Fisher exact	.872
TG	4	18		
TT	29	91		
Dominant model GG + TG vs TT			0.6369	.425
GG+TG	4	20		
TT	29	91		
Recessive model GG vs TT+ TG			Fisher exact	>.999
GG	0	2		
TT+TG	33	109		
rs13210237				
AA	1	4	Fisher exact	.142
GA	16	33		
GG	16	72		
Dominant model AA + GA vs GG			3.318	.068
AA+GA	17	37		
GG	16	72		
Recessive model AA vs GA + GG			$\chi^2_{Yates}=0.132$	.716
AA	1	4		
GA+GG	32	105		

Table 7. The demographic and clinical characteristics of PE patients and controls.

	PE N=144		non-PE N=291		χ^2 / Z	P
Maternal age(year)	31	(29.33)	30	(29.34)	-0.284	0.776
height(cm)	162	(159,166)	163	(160,166)	-0.977	0.329
BMI (Kg/m ²)	24.06	(21.34,28.51)	21.02	(19.13,24.23)	2.563	0.000
Weight gain(kg)	16	(11.1,20)	15	(11,20)	-1.379	0.168
Diastolic BP (mm Hg)	100	(90,120)	70	(60,70)	-4.457	0.000
Systolic BP (mm Hg)	158	(130,180)	100	(100,110)	-5.167	0.000
diabetes [n(%)]	41	(28.47)	37	(12.71)	16.254	0.000
Multiple pregnancy [n(%)]	9	(6.25)	5	(1.72)	4.980	0.026
gravidity ≥ 2 [n(%)]	56	(38.89)	143	(49.14)	4.079	0.043
History of miscarriage [n(%)]	41	(28.47)	91	(31.27)	0.357	0.550
family history of hypertension [n(%)]	39	(27.08)	82	(28.18)	0.058	0.810
in vitro fertilization [n(%)]	8	(5.56)	16	(5.50)	0.001	0.980
Education background						
≤ 9 Year (%)	11	(7.64)	20	(6.87)		
9-12 Year (%)	16	(11.11)	27	(9.28)	-0.007	0.994
12-16 Year (%)	100	(69.44)	214	(73.54)		
>16 Year (%)	17	(11.81)	30	(10.31)		
Monthly household income (RMB)						
<5000/case (%)	88	(61.11)	64	(21.99)		
5000-10000/case (%)	138	(95.83)	55	(18.90)	-2.662	0.008
≥ 10000 /case(%)	65	(45.14)	25	(8.59)		
rs13176432 Dominant model [n(%)]	24	(16.67)	13	(4.47)	18.422	0.000
rs13176432 Recessive model [n(%)]	2	(1.39)	0	(0.00)	1.593	0.207
rs13210237 Dominant model [n(%)]	54	(37.50)	54	(18.56)	18.522	0.000
rs13210237 Recessive model [n(%)]	5	(3.47)	2	(0.69)	3.124	0.077
red blood cell ($\times 10^{12}/L$)	4.21	(3.755,4.5)	4.1	(3.7, 4.36)	-2.015	0.044
White blood cell ($\times 10^9/L$)	8.685	(7.325,9.975)	7.8	(6.4, 9.2)	-3.969	0.000
hemoglobin (g/L)	131	(124,136)	128	(122,135)	-2.013	0.044
platelet ($\times 10^9/L$)	245	(201.25,275.7)	218	(184,256)	-3.398	0.001
Fasting blood glucose (mmol/L)	4.895	(4.56,5.26)	4.84	(4.56,5.1)	-1.554	0.120
Alanine aminotransferase (U/L)	18.2	(13.15,26.37)	16	(11,24)	-2.632	0.008
urea nitrogen (mmol/L)	3	(2.575,3.5575)	3.02	(2.6,3.37)	-0.382	0.702
creatinine (μ mol/L)	52	(46.61,2625)	49.2	(44,57.6)	-1.968	0.049
uric acid(μ mol/L)	222.55	(190.25,268.8)	203.4	(171.8,232.1)	-3.99	0.000

The results showed that BMI before pregnancy, systolic blood pressure, gestational diabetes, multiple pregnancy, pregnancy times ≥ 2 , rs13210237 dominant mode, rs13176432 dominant mode, platelet count in early pregnancy may be related to preeclampsia.

Table 8. Multiple logistic regression analysis.

	B	S.E.	Wald	P	Exp(B)	95% CI	
						Lower	Upper
BMI	0.097	0.031	9.687	0.002	1.102	1.037	1.172
Diastolic BP	0.036	0.011	11.223	0.001	1.037	1.015	1.059
Diabetes mellitus	0.752	0.291	6.696	0.01	2.121		
Multiple pregnancy	1.821	0.65	7.846	0.005	6.177	1.728	22.083
Gravidity $\geq$ 2	-0.526	0.241	4.774	0.029	0.591	0.369	0.947
Rs13210237 Dominant model	1.144	0.263	18.871	0.000	3.138	1.873	5.256
Rs13176432 Dominant model	1.769	0.403	19.254	0.000	5.867	2.662	12.932
Platelet	0.006	0.002	9.496	0.002	1.006	1.002	1.009
	-10.492	1.499	49.008	0	0		

The role of rs13210237 and rs13176432 in preeclampsia

According to multivariate logistic regression including the demographic and clinical characteristics (Table 7), rs13210237 (dominant mode) and rs13176432 (dominant mode) were related to preeclampsia (Table 8), with ORs of 3.138 (95% CI 1.873–5.256) and 5.867 (95% CI 2.662–12.932), respectively.

Discussion

By GWAS technology, genetic mutation sites highly associated with the morbidity of preeclampsia were screened from preeclampsia women among the Han population of China (refer to Table 2 for details). On this basis, enrichment analysis was conducted for the genes at the above conformed preeclampsia-associated SNP sites to elucidate the mechanism of action of the genes at SNP sites during the morbidity of preeclampsia. Multiple RCR technologies were adopted to verify the positive SNP sites from the population during GWAS (rs13176432, rs13210237) to obtain statistically significant preeclampsia-associated SNP sites; after enrichment analysis, adenylate cyclase-inhibiting G-protein coupled receptor protein signaling pathways could possibly participate in preeclampsia pathogenesis (10).

Based on this study, we found two SNPs associated with preeclampsia in the Han population in China. rs13176432 is located in area 1 and zone 3 of the No. 5 short chromosome arm and belongs to an intron. The associated gene is a ubiquitin ligase gene (tripartite motif-containing 36, TRIM36), which plays an important role in trophoblast autophagy (11). During autophagy extension and closure, for example, Atg8 (autophagy proteins, Atg8 protein family in mammals: LC3A, LC3B, LC3C, etc.) is required for the extension of phagosomes and the closure of autophagic phagosomes. The terminal C of LC3 (microtubule-associated protein

1 light chain 3 family protein) is combined with the Atg12-Atg5-Atg16L complex, which participates in the formation of autophagosomes (Atg12 is first activated by ubiquitin ligase and then forms a complex with Atg5; subsequently, Atg5-Atg12 is combined with Atg16 containing the coiled-coil domain protein to form a multisubunit protein complex) (12). Atg8 protein participates in the extension and closure of the phagosome membrane and finally forms a double-membrane autophagic phagosome (12). Therefore, ubiquitin ligase abnormalities resulting from genetic factors will directly affect autophagy inhibition or excessive autophagy in cells. These two factors have been reported to induce the occurrence of preeclampsia. For example, when cell autophagy occurs, the infiltration capacity of extravillous trophoblasts cells (EVTs) decreases, resulting in the shallow implantation of the placenta and then inducing preeclampsia (13). It has been reported that preeclampsia can also occur when the placental microenvironment, such as oxidative stress, induces excessive autophagy of trophoblast cells or endothelial cells (14).

rs13210237 is located in area 1 and zone 2 of the No. 6 short chromosome arm, and the corresponding gene is heat shock transcription factor 2 (HSF2). Placental trophoblast cells produce a series of stress proteins under oxidative stress, of which heat shock protein (HSP), as a molecular chaperone, is synthesized by biological cells after the destructive stimuli of pathogens and physical and chemical factors (15). In a previous report, the expression of several HSPs was too high or low in the preeclampsia patient's placental or maternal blood (16). For example, HSP70 shows a high expression status in preeclampsia patients' maternal blood, while HSP90 shows low expression. These abnormal expressions are possibly caused by the abnormality of heat shock transcription factors (16).

Another gene associated with the rs13210237 site is gap junction protein alpha 1 (GJA1), which corresponds to the gap junction protein. The cell gap junction, also called the gap junction, is the only pathway among adjacent cells and makes the cytoplasm couple

adjacent cells to regulate the differentiation and development of cells (17). In recent years, studies have shown that the gap junction plays an important role in cell apoptosis and survival (18). The balance of placental cell proliferation and apoptosis is the basis for the placenta, while the disorder of cell apoptosis will lead to pathological pregnancy, such as hypertension during pregnancy (19). Studies have shown that the quantity of apoptotic placental trophoblasts, syncytiotrophoblast cells and decidual cells is higher than that in the normal pregnancy group (20). Thus, the gap junction protein abnormality resulting from genetic factors may participate in the formation of preeclampsia by impacting the apoptosis of placental cells (20).

Based on the enrichment analysis, furthermore, the adenylate cyclase-inhibiting G-protein coupled receptor signaling pathway possibly participates in the pathogenesis of preeclampsia. Some reports show that the genes GPR124 and MRGPRF, which are related to the G-protein coupled receptor signaling pathway, are reduced in early-onset preeclampsia, showing the relationship between this pathway and the morbidity of preeclampsia (10).

Our studies about susceptibility genes for preeclampsia are restrictive. First, an important limitation of this manuscript is the limited number of participants for this kind of study and the lack of power calculation. Second, even though there are new findings during this study, our GWAS case-control study results have not been proven in other population samples. Because of the significant ethnic differences and population heterogeneity of preeclampsia patients worldwide, our study may not be representative.

Conclusions

In summary, herein, we performed the first GWAS on preeclampsia patients in the Tianjin Han population, China. This study included a moderate sample size, and we reported several preeclampsia-related genetic variations and gene regions. These new findings will be beneficial to further research on relevant biological functions to prevent preeclampsia. At present, it remains unclear how the underlying mechanism of preeclampsia and these genetic variations interact with other factors to affect the incidence of this condition. It is necessary to replicate our findings in other independent samples in China and in other Asian countries.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This project was supported by Tianjin Municipal Science and Technology Committee project No. 15ZXJZSY00010, and Tianjin Science and Technology project No. 16ZXMJSY00130.

ORCID

Zhenhua Cao  <http://orcid.org/0000-0001-5089-4303>

References

- [1] Souza JP, Gülmezoglu AM, Vogel J, et al. Moving beyond essential interventions for reduction of maternal mortality (the WHO multicountry survey on maternal and newborn health): a cross-sectional study. *Lancet*. 2013;381(9879):1747–1755. doi: 10.1016/S0140-6736(13)60686-8
- [2] Dawson LM, Parfrey PS, Hefferton D, et al. Familial risk of preeclampsia in Newfoundland: a population-based study. *J Am Soc Nephrol*. 2002;13(7):1901–1906. doi: 10.1097/01.asn.0000017224.24670.82
- [3] Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet*. 2005;6(2):95–108. doi: 10.1038/nrg1521
- [4] van Dijk M, Mulders J, Poutsma A, et al. Maternal segregation of the Dutch preeclampsia locus at 10q22 with a new member of the winged helix gene family. *Nat Genet*. 2005;37(5):514–519. doi: 10.1038/ng1541
- [5] Ferreira LC, Gomes CE, Araújo AC, et al. Association between ACVR2A and early-onset preeclampsia: replication study in a Northeastern Brazilian population. *Placenta*. 2015;36(2):186–190. doi: 10.1016/j.placenta.2014.11.007
- [6] Zhao L, Triche EW, Walsh KM, et al. Genome-wide association study identifies a maternal copy-number deletion in PSG11 enriched among preeclampsia patients. *BMC Pregnancy Childbirth*. 2012;12(1):61. doi: 10.1186/1471-2393-12-61
- [7] Melton PE, Johnson MP, Gokhale-Agashe D, et al. Whole-exome sequencing in multiplex preeclampsia families identifies novel candidate susceptibility genes. *J Hypertens*. 2019;37(5):997–1011. doi: 10.1097/HJH.0000000000002023
- [8] Guo LF, Wang ZH, Wang YF. Common variant rs7579169 is associated with preeclampsia in Han Chinese women. *Genet Mol Res*. 2016;15(2). doi: 10.4238/gmr.15028114
- [9] Wang Q, Wang G, Guo C, et al. Single nucleotide polymorphisms near the inhibin beta B gene on 2q14 are associated with pre-eclampsia in Han Chinese women. *Eur J Obstet Gynecol Reprod Biol*. 2015;193:127–131. doi: 10.1016/j.ejogrb.2015.04.001
- [10] Liang M, Niu J, Zhang L, et al. Gene expression profiling reveals different molecular patterns in G-protein coupled receptor signaling pathways between early- and late-onset preeclampsia. *Placenta*. 2016;40:52–59. doi: 10.1016/j.placenta.2016.02.015

- [11] Singh N, Kumble Bhat V, Tiwari A, et al. A homozygous mutation in TRIM36 causes autosomal recessive anencephaly in an Indian family. *Hum Mol Genet.* 2017;26(6):1104–1114. doi: [10.1093/hmg/ddx020](https://doi.org/10.1093/hmg/ddx020)
- [12] He C, Baba M, Cao Y, et al. Self-interaction is critical for Atg9 transport and function at the phagophore assembly site during autophagy. *Mol Biol Cell.* 2008;19(12):5506–5516. doi: [10.1091/mbc.e08-05-0544](https://doi.org/10.1091/mbc.e08-05-0544)
- [13] Yamanaka-Tatematsu M, Nakashima A, Fujita N, et al. Autophagy induced by HIF1 α overexpression supports trophoblast invasion by supplying cellular energy. *PLoS One.* 2013;8(10):e76605. doi: [10.1371/journal.pone.0076605](https://doi.org/10.1371/journal.pone.0076605)
- [14] Gao L, Qi HB, Kamana KC, et al. Excessive autophagy induces the failure of trophoblast invasion and vasculature: possible relevance to the pathogenesis of preeclampsia. *J Hypertens.* 2015;33(1):106–117. doi: [10.1097/HJH.0000000000000366](https://doi.org/10.1097/HJH.0000000000000366)
- [15] Gomez-Pastor R, Burchfiel ET, Thiele DJ. Regulation of heat shock transcription factors and their roles in physiology and disease. *Nat Rev Mol Cell Biol.* 2018;19(1):4–19. doi: [10.1038/nrm.2017.73](https://doi.org/10.1038/nrm.2017.73)
- [16] Molvarec A, Prohászka Z, Nagy B, et al. Association of elevated serum heat-shock protein 70 concentration with transient hypertension of pregnancy, preeclampsia and superimposed preeclampsia: a case-control study. *J Hum Hypertens.* 2006;20(10):780–786. doi: [10.1038/sj.jhh.1002060](https://doi.org/10.1038/sj.jhh.1002060)
- [17] Lampe PD, Lau AF. The effects of connexin phosphorylation on gap junctional communication. *Int J Biochem Cell Biol.* 2004;36(7):1171–1186. doi: [10.1016/S1357-2725\(03\)00264-4](https://doi.org/10.1016/S1357-2725(03)00264-4)
- [18] Zhang BX, Ma X, Shu Z, et al. Differential regulation of intracellular calcium oscillations by mitochondria and gap junctions. *Cell Biochem Biophys.* 2006;44(2):187–203. doi: [10.1385/CBB:44:2:187](https://doi.org/10.1385/CBB:44:2:187)
- [19] Aoki A, Nakashima A, Kusabiraki T, et al. Trophoblast-specific conditional Atg7 knockout mice develop gestational hypertension. *Am J Pathol.* 2018;188(11):2474–2486. doi: [10.1016/j.ajpath.2018.07.021](https://doi.org/10.1016/j.ajpath.2018.07.021)
- [20] DiFederico E, Genbacev O, Fisher SJ. Preeclampsia is associated with widespread apoptosis of placental cytotrophoblasts within the uterine wall. *Am J Pathol.* 1999;155(1):293–301. doi: [10.1016/S0002-9440\(10\)65123-1](https://doi.org/10.1016/S0002-9440(10)65123-1)