The Dynamic Genetic Atlas of 122 Gestational Phenotypes

Siyang Liu[1*#], Hao Zheng[1,2,3*], Yuqin Gu[1,3*], Zijing Yang[1,2*], Yanhong Liu[1], Yuandan Wei[1,3], Xinxin Guo[1], Yanchao Chen[1], Liang Hu[2], Xiaohang Chen[2], Quanfu Zhang[3], Guo-Bo Chen[4], Xiu Qiu[5], Shujia Huang[5#], Jianxin Zhen[3#], Fengxiang Wei[2,6#]

1. School of Public Health (Shenzhen), Sun Yat-sen University, Shenzhen, Guangdong 518107, China
2. The Genetics Laboratory, Longgang District Maternity and Child Healthcare Hospital of Shenzhen City, Shenzhen, Guangdong, 518172, China
3. Central Laboratory, Shenzhen Baoan Women's and Children's Hospital, Shenzhen, Guangdong 518102, China
4. Center for Productive Medicine, Department of Genetic and Genomic Medicine, Clinical Research Institute, Zhejiang Provincial People’s Hospital, People’s Hospital of Hangzhou Medical College, Hangzhou 310014, Zhejiang, China
5. Division of Birth Cohort Study, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, 510623, China
6. Longgang Maternity and Child Institute of Shantou University Medical College, Shenzhen, Guangdong, 518172, China

*: Those authors contribute equally as co-first authors

#: Correspondence can be addressed to

Siyang Liu liusy99@mail.sysu.edu.cn

Shujia Huang shujia.huang@bigcs.org

Jianxin Zhen jxzhen@qq.com

Fengxiang Wei haowei727499@163.com

Abstract

The gestational period, spanning approximately 40 weeks from fertilization to birth, is fundamental to human reproduction. Health monitoring during this period involves systematic prenatal and postpartum examinations, guided by indicators collectively termed gestational phenotypes under the national medical insurance framework. Despite their clinical importance, the genetic basis of these phenotypes and their links to later-life health outcomes remain poorly understood. In this large-scale genetic study, we analyzed 122 gestational phenotypes in 121,579 Chinese pregnancies, encompassing anthropometric metrics, blood biomarkers, and common gestational complications and outcomes. We identified 3,845 genetic loci, including 1,893 novel loci, and uncovered gestation-specific genetic effects in 23 phenotypes, with proportions ranging from 0% to 100%. These loci were enriched in pathways related to hormonal regulation, cell growth and immune function. Longitudinal genome-wide association analyses of repeated measures across 24 complete blood cell phenotypes revealed significant gene-by-gestational timing interactions for 17.8% of loci across five gestational and postpartum periods. Mendelian Randomization of 220 mid- and late-life phenotypes identified 73 causal associations between gestational phenotypes and chronic diseases risks. These findings provide critical insights into the genetic architecture of human gestational phenotypes and their implications for long-term health, providing a foundation for advancing population health during gestation. Visualization of results is available at https://monn.pheweb.com/ .

Introduction

Women's reproductive, maternal, newborn, child, and adolescent health are fundamental pillars of human health and key drivers of future population and social progress[1,2]. In the context of declining fertility rates, increasing early-life disorders and global population aging[3,4], maternal and child health has become a pressing global public health concern. The gestational period, spanning approximately 40 weeks from fertilization to birth, is particularly critical. During this time, the developing organism undergoes substantial growth and differentiation, while the pregnant individual experiences profound physiological changes to support fetal development[5,6]. Recognizing the critical nature of this period, maternal and newborn health is routinely monitored through systematic, timed prenatal and postpartum examinations supported by national medical insurance systems. These assessments encompass anthropometric metrics, a range of blood and urine tests, and prenatal screenings, collectively referred to as the *gestational phenotypes*, which serve as key indicators for diagnosing pregnancy-related disorders and forecasting birth outcomes (**Fig. 1 and Supplementary Table 1**).

Despite the recognized importance of the gestation period, the genetic underpinnings of gestational phenotypes and their potential connections to later-life health outcomes remain poorly understood. Of the 122 gestational phenotypes examined in this study, only 22 have been previously investigated in GWAS reported in the GWAS Catalog (**Supplementary Table 1**), with 10 included in our prior research[7-9]. Other studies that have not reported findings in the GWAS Catalog are limited by insufficient sample sizes[10,11]. Consequently, while large-scale genetic studies have explored various quantitative and qualitative phenotypes in the general population[12,13], few have addressed the unique genetic architecture of gestational phenotypes[8,9,14], particularly the time-dependent genetic effects during pregnancy. Furthermore, although observational studies have identified associations between gestational

conditions and maternal health outcomes later in life[15-18], causal evaluations of the spectrum of gestational phenotypes- typically collected in early adulthood- on later-life health outcomes remain scare. These gaps limit the potential for early diseases screening in women aged 20 to 40.

To address these limitations, we collected data of 122 prenatal and postnatal screening indicators alongside non-invasive prenatal test (NIPT) sequencing data from 121,579 unrelated pregnant women (mean age 30 ± 4 years). Using a large-scale GWAS with advanced analytical protocol[7,19], we constructed an atlas of quantitative and qualitative trait loci, and validated our findings in external cohorts. We identified genetic loci with differential effects during pregnancy compared to non-pregnancy using a Bayesian classification algorithm, and investigated gene-by-time ($G \times T$) interactions through repeated measurements of 24 phenotypes during pregnancy. Lastly, we leveraged this atlas to estimate both direct and indirect causal relationships between gestational phenotypes and 220 phenotypes in BioBank Japan (mean age 63 years)[12], modeling the long-term health impacts of gestational phenotypes (**Fig. 1**).

Results

Study design

We collected NIPT sequencing data alongside 122 comprehensive pregnancy screening metrics, including biomarkers and electronic medical records, from 121,579 pregnancies undergoing routine obstetric examinations at two hospitals in Shenzhen. Among these metrics, 88 gestational phenotypes were classified into nine molecular test categories: blood lipid (N=2), glycemic markers (N=6), blood routine (N=32), infection (N=14), kidney function (N=3), liver function (N=8), thyroid function (N=4), Tang screening (N=16), and NIPT screening (N=3) (**Fig.1, Supplementary Table 2**). Additionally, 34 phenotypes were grouped into three categories derived from electronic medical records: maternal baseline characteristics, such as height,

weight, and blood pressure (N=7); gestational disorders (N=21); and birth outcomes (N=6) (**Fig. 1, Supplementary Table 3**). Sample sizes for each phenotype ranged from 12,024 (gestational diabetes mellitus with medication in the glycemic category) to 115,872 (trisomy 18 risk score in the NIPT screening category). The distribution of all quantitative gestational phenotypes and the prevalence of qualitative gestational phenotypes, including common gestational comorbidities, were highly consistent across the two hospitals (**Supplementary Table 4-5, Supplementary Fig. 1-2**).

We applied our comprehensive analytical protocol to infer and impute genotypes[7,20], estimate family-relatedness, assess population structure, and conduct GWAS using the NIPT data. We replicated the GWAS findings in two independent cohorts (NIPT PLUS N= 5,733 and Baoan 20K cohort N=20,891). Detailed descriptions of the methodology are provided in the **Methods** section.

3,845 genetic associations with 122 gestational phenotypes

We performed GWAS for the 122 gestational phenotypes across the two hospitals, followed by a fixed-effect meta-analysis of 12.6 million genetic markers (**Supplementary Table 6**). Linkage disequilibrium (LD) score intercepts ranged from 0.94 to 1.10, indicating minimal statistical inflation (**Supplementary Table 7**). We report results from the conventional genome-wide significance threshold ($P < 5\times10^{-8}$), and a stricter study-wide significance threshold ($P = 4.1\times10^{-10}$) after Bonferroni correction for 122 phenotypes.

Using GCTA conditional and joint analysis[21], we identified 4,309 independent signals across 3,242 loci associated with at least one of the nine categories of molecular test phenotypes, with 2,942 signals in 1,999 loci reaching study-wide significance ($P < 4.1\times10^{-10}$) (**Fig. 2, Supplementary Table 8- 9**). The blood routine category yielded the highest number of loci, with platelet-large cell ratio (P_LCR) being the phenotype

with the most loci (**Fig. 2, Supplementary Table 9**). Across the three categories of basic information, birth outcomes and gestational disorders, we identified 718 independent signals in 603 loci ($P < 5\times10^{-8}$) and 367 independent signals in 279 loci ($P < 4.1\times10^{-10}$) (**Supplementary Table 10-11**). Notably, 5,483 and 5,194 functional variants, including missense and nonsense variants, were identified for the genome-wide and study-wide significant loci, respectively (**Supplementary Table 12-13**).

Genetic effect estimates were highly consistent across the two hospitals, with squared Pearson's correlation coefficient (R^2) ranging from 0.76 to 0.97 for the molecular test phenotypes and from 0.12 to 0.94 for the medical record phenotypes (**Supplementary Fig. 3**). Only 8.5% (369/4,309) of loci associated with the nine molecular test categories and 10.3% (74/718) of loci associated with the three medical record categories exhibited statistical heterogeneity ($P_{\text{het}}<0.05$) (**Supplementary Table 8-11**).

To validate our findings, we conducted two replication studies. First, assuming maternal height is not influenced by gestational status, genetic effect estimates for maternal height-associated variants in our study were compared to those from the Taiwan Biobank[22]. Estimates were highly consistent (R^2=0.93), with 121 variants surpassing Bonferroni correction ($P < 2.79\times10^{-4}$, 66.9%), and 166 variants achieving $P < 0.05$ (91.7%) (**Supplementary Fig. 4, Supplementary Table 14**). Second, we compared genetic effects with two independent cohorts (NIPT PLUS and Baoan 20K). Among 4,959 variants examined for 107 phenotypes in the replication cohorts, 4,716 (95.1%) demonstrated consistent genetic effect sizes (two-sample *t-test* $P >$ 0.05) (**Supplementary Fig. 5-7**, **Supplementary Table 15**). For study-wide significant variants (3,293), 3,156 (95.8%) showed consistent effects (**Supplementary Table 15**). These results indicate the robustness and fidelity of the GWAS findings and genetic effect estimates in this study.

1,893 novel genetic associations

Through comparison with the GWAS Catalog[23] of previously reported associations for the same or similar phenotypes in either pregnant or general populations (**Methods**), we identified 1,594 novel loci among the 3,242 genome-wide significant loci (**Supplementary Table 8**), and 830 novel loci among the 1,999 study-wide significant loci across the nine molecular test categories (**Supplementary Table 9**). Similarly, among the three medical record categories, 299 of the 603 genome-wide significant loci, and 109 of the 279 study-wide significant loci were novel (**Supplementary Table 10-11**). The 315 newly discovered loci with study-wide significance and at least one functional variant are illustrated in **Fig. 2**.

Among the 830 study-wide significant novel loci associated with the nine molecular test categories, 156 were linked to multiple gestational phenotypes (**Supplementary Table 16**). Genes mapped to these 156 pleiotropic loci were significantly enriched in Gene Ontology (GO) pathways such as responses to hormones, in contrast to enrichment for more fundamental cellular processes, such as hemostasis and regulation of cell activation, among genes without pleiotropy (**Supplementary Fig. 8**). Protein-protein interaction network (PPIN) analysis identified key hub genes, including *IL6* (involved in immune responses), *ESR1* (related to hormone response) and *IGF1* and *EGF* (both play roles in cell growth and development) (**Supplementary Fig. 9, Supplementary Table 17**). *ESR1*, associated with nine gestational phenotypes, is one of the loci with the broadest pleiotropy (**Supplementary Fig. 9, Supplementary Table 16**). This gene, encoding estrogen receptor 1, plays a critical role in embryo implantation by affecting uterine natural killer cell motility and vascular remodeling in early pregnancy and the placenta [24-26]. Consistent with this role, we found that *ESR1* variants were significantly associated with leukocyte-related traits, including

lymphocyte percentage (LYM_P), monocyte absolute value (MON), neutrophils absolute value (NEUT), and white blood cell count (WBC)) during pregnancy (**Supplementary Fig. 10, Supplementary Table 16**). In addition, our study first revealed strong associations of *ESR1* with triglyceride (TG) levels, oral glucose tolerance test 0-hour values (OGTT0H), red blood cell count (RBC) and hematocrit (Hct). Other highly pleiotropic loci include *ELOVL5* (encoding ELOL fatty acid elongase 5), associated with beta human chorionic gonadotropin (FHCG), pappalysin (PAPP), and Tang screening scores for trisomies 21 and 18. Additionally, *SLC35A1* (solute carrier family 35 member A1) was associated with FHCG, Tang screening scores for trisomy 21, and free thyroxine (FT4), all of are critical for pregnancy (**Supplementary Table 16**).

We particularly examined phenotypes that are previously underexplored, such as screening phenotypes for fetal trisomy. Combined maternal serum Tang and NIPT screening, along with fetal ultrasound, are commonly used to assess the risk of fetal trisomy[27,28], making these phenotypes crucial during gestation. For NIPT screening scores for trisomy 21 (T21), trisomy 18 (T18) and trisomy 13 (T13), we identified 26, 35 and 31 quantitative trait loci ($P < 4.1\times10^{-10}$), respectively (**Supplementary Table 9, Supplementary Fig. 11**). Notably, intronic variants and expression quantitative trait loci (eQTLs) for *PANX1* (Pannexin 1) on chromosome 11, were significantly associated with all three NIPT screening scores. *PANX1* encodes a cell membrane factor highly expressed in human oocytes, eight-cell embryos, and the brain, has been linked to a form of infertility known as “oocyte death” in family studies[29,30]. In our analysis, variants associated with decreased *PANX1* expression correlated with increased NIPT screening scores, indicating higher trisomy risk. Similarly, variants near and within *PADI4* (peptidyl arginine deiminase 4) were associated with these three phenotypes. *PADI4*, which plays a role in neutrophil extracellular trap formation (NETosis), is critical for immune regulation during

pregnancy[31-33]. Furthermore, *TNFSF13B (*TNF superfamily member 13b) and *USP3 (*ubiquitin specific peptidase 3) were associated with two of these scores, both genes being integral to the innate immune response[34,35].

Heritability and correlation of gestational phenotypes

To estimate the heritability of 122 gestational phenotypes, we employed LD Score regression (LDSC)[36]. Among these phenotypes, maternal serum vitamin B12 levels exhibited the highest SNP heritability, at 43.97% (**Supplementary Table 7, Supplementary Fig. 12**). In contrast, gestational thrombocytopenia had the lowest SNP heritability, at 0.0045% (**Supplementary Table 7, Supplementary Fig. 12**). While the SNP heritability of vitamin B12 levels in the UK Biobank (UKB) is reported to be low (h^2=0.00953, se=0.0143)[37], a study of white British twin women estimated a much higher heritability of 56% for serum vitamin B12[38]. We hypothesize that this discrepancy may be attributable to differences in race, gender and physiological status between the study populations.

Phenotypic and genetic correlations were computed for 88 quantitative phenotypes across nine categories (**Supplementary Fig. 13, Supplementary Table 18-19,**). The resulting heatmaps were largely symmetric along the red diagonal line, indicating that substantial phenotypic variability is derived from genetic variability (**Supplementary Fig. 13**). Phenotypic domains with strong correlations were typically clustered within categories, although significant and robust genetic correlations were also observed across categories.

Gestation-specific genetic effects

Gestational phenotypes may exhibit gestation-specific genetic effects distinct from those observed in the same phenotypes outside of pregnancy. For 23 of the 122 phenotypes investigated, including maternal height, we utilized GWAS data from a

non-pregnant cohort in the Taiwan Biobank (n = 92,615, age 30-70, southern Chinese ancestry) [22]. Genetic correlations between the two datasets for the same phenotypes were all significant, ranging from 0.18 to 0.94 (**Supplementary Table 20**). The genetic effects of lead SNPs from NIPT were highly correlated between this study and the Taiwan Biobank, with R^2 ranging from 0.001 to 0.99 (**Supplementary Fig. 14, Supplementary Table 21**). Among the 900 study-wide significant genetic signals identified for the 23 phenotypes, 792 had genetic effect information available in both datasets (**Supplementary Table 22**).

To identify genetic effects predominantly influenced by pregnancy compared to non-pregnancy populations, we applied a Bayesian clustering algorithm[39] combined with colocalization analysis. Detailed model parameters are provided in **Supplementary Table 23 and the Methods.** Genetic loci associated with each gestational trait were classified into three groups: general (shared effects with non-gestational phenotypes), gestation-specific (effects dominated by gestational phenotypes), and unclassified. In total, 138 genetic signals (17.4%) were identified as exhibiting gestation-specific effects, while 381 signals (48.1%) showed shared effects, and 273 (34.5%) remained unclassified (**Supplementary Fig. 15, Supplementary Table 24**). When ranking the 23 phenotypes by the proportion of gestation-specific loci, alpha-fetoprotein levels (AFP_S2) and their adjusted form, alpha-fetoprotein multiples-of-median (AFPMOM_S2), from second-trimester screening exhibited the highest proportion of gestation-specific variants. These were followed by two liver function biomarkers: alanine transaminase (ALT) and albumin (ALB) (**Fig. 3A, Supplementary Table 24**). In contrast, maternal height and first systolic blood pressure measurements had the lowest proportion of gestation-specific genetic loci, aligning with the expectation that maternal height remains largely unchanged during pregnancy, and blood pressure remains stable during early pregnancy.

Pathway enrichment analyses for the gestation-specific, general, and overall genetic loci are summarized in **Supplementary Fig. 14**. Further pathway analyses of genes from the gestation-specific and general effect clusters revealed significant differences in pathway composition (two-tailed chi-square or Fisher's exact test, $P < 0.05$) (**Fig. 3B, Supplementary Table 25)**. Gestation-specific genetic loci were enriched in pathways such as retinoid metabolism and transport, Wnt signaling, pluripotency, and regulation of protein catabolic processes. Additionally, nominally significant pathways include hormone response, glucose homeostasis, lymphocyte proliferation (GO terms), breast cancer (KEGG), cancer and adipogenesis (WikiPathways), and signaling by nuclear receptors (Reactome) (**Supplementary Table 25**). PPI network analyses weighted by gene degrees identified hub genes consistent with those associated with multiple gestational phenotypes, including the estrogen receptor (*ESR1*), insulin like growth factor 1 (*IGF1*) and interleukin 6 (*IL6*) (**Supplementary Fig. 17, Supplementary Table 26**). These findings highlight the critical roles of hormonal regulation, growth mediators, and immune processes in initiating and maintaining gestation.

G (genetics) by T (time) interaction across gestation and the postpartum period

Among the 122 gestational phenotypes studied, 24 phenotypes in the blood routine category, assessed using complete blood count (CBC) assays[40,41] , were repeatedly measured at a median of 5 to 6 time points in across the two hospitals (**Supplementary Fig. 18-19, Supplementary Table 27**). These data provided a unique opportunity to investigate longitudinal, time-varying genetic effects during gestation.

The temporal trends for these 24 phenotypes during gestation were largely consistent between the two hospitals (**Supplementary Fig. 20-21**). For leukocyte-related

phenotypes, white blood cell count (WBC), neutrophil count (NEUT), and neutrophil percentage (NEU_P) increased significantly during pregnancy, while lymphocyte count (LYM) decreased. Monocyte count (MON) increased, eosinophil count (EOS) remained stable, and basophil count (BASO) fluctuated around baseline values. Red blood cell count (RBC), hemoglobin concentration (HGB), hematocrit (Hct), and platelet count (PLT) exhibited significant declines during pregnancy. All CBC phenotypes underwent substantial changes at delivery and during the postpartum period, trending back toward pre-pregnancy levels.

To investigated longitudinal genetic effect, we employed a two-step approach. First, we used TrajGWAS[42] to identify genetic variants associated with phenotypic mean variability across the 24 CBC phenotypes, separately for each hospital. TrajGWAS[42] also identified variants influencing the within-subject variability. Variants affecting phenotypic mean (β_g) and within-subject variability (τ_g) at genome-wide significance levels are presented in **Supplementary Fig. 22-29**. In total, 645 variants reached genome-wide significance for phenotypic mean effect in both hospitals ($P < 5\times10^{-8}$). Second, we extended the TrajGWAS model to include an interaction term for genotype dosage and gestational day, enabling a locus-based genetic-by-time ($G \times T$) analysis (**Methods**). A meta-analysis integrating results from both hospitals using an inverse- variance weighted fixed-effect model identified 115 out of the 645 variants (17.8%) with significant genetics by gestational days interaction effects ($P < 5\times10^{-8}$) (**Supplementary Table 28**). The remaining 530 variants influenced phenotypic mean levels but did not show significant interaction effects (**Supplementary Table 29**). The proportion of variants with significant $G \times T$ effects varied across phenotypes, ranging from 5.4% to 47.8% (**Supplementary Table 30**).

To explore the temporal dynamics of these 115 $G \times T$ variants, we conducted localized association studies across five periods: first trimester, second trimester, third

trimester, delivery, and postpartum (**Methods, Supplementary Table 31**). Of these, 79 variants (68.7%) exhibited non-overlapping 95% confidence intervals for genetic effects between at least two periods (**Supplementary Fig. 30**). The top 25 variants with the most significant $G \times T$ interactions were presented in **Fig. 4A,** of which 22 demonstrated strong expression quantitative trait loci (eQTL) associations (**Supplementary Table 31**). Using gestational WBC count as an example, we observed a steady increase in WBC levels beginning in the first trimester, peaking during the 3rd trimester and delivery (**Fig. 4B**). Of the 27 variants associated with WBC, eight (29.6%) displayed significant $G \times T$ interaction. Among these, rs2270401 (17q21.1), a strong eQTL for *GSDMA*, exhibited the most pronounced $G \times T$ interaction effect, with substantial variation in genetic effects across the five periods: first trimester (β=0.270, P=6.44×10^{-491}), second trimester (β=0.396, P=1.62×10^{-1223}), third trimester (β=0.388, P=2.27×10^{-1021}), delivery (β=0.257, P=2.06×10^{-152}), and postpartum (β=0.122, P=6.48×10^{-60}) (**Fig. 4C, Supplementary Table 31**). The genetic effect of the rs2270401-A allele in the second trimester was 3.33 times greater than that observed during the postpartum period. The remaining seven loci with significant $G \times T$ interactions demonstrated U-shaped changes in genetic effects across the five period (**Supplementary Fig. 30**). These patterns in WBC phenotypic and genetic effect changes suggest that these loci may play pivotal roles in the rapid leukocyte proliferation observed during gestation and postpartum.

We further categorized genetic variants based on the presence or absence of significant genetic and environmental interactions and performed enrichment analyses. Gene enrichment results for each category are summarized in **Supplementary Fig. 31**. Using chi-squared and Fisher's exact tests, we identified pathways significantly enriched among variants with interaction effects compared to those without. These included pathways involved in the regulation of phagocytosis, engulfment, prefoldin mediated transfer of substrate to CCT/TriC, Heparan

sulfate/heparin (HS-GAG) metabolism, Gap junction, nerve development, and lipid modification (**Fig. 4D, Supplementary Table 32**).

Different from genes harboring gestation-specific genetic effects, PPI network hub gene analyses identified Epidermal Growth Factor (*EGF*), which is critical for cell growth and development, and GATA binding protein 2 (*GATA2*), a key regulator of hematopoietic development, as the most prominent hub genes (**Supplementary Fig. 32, Supplementary Table 33**). These findings suggest that changes in the genetic effects of variants influencing blood cell levels during pregnancy are predominantly associated with alterations in cell activation and cytokine activity in pregnant women.

Causal relationships with older-age phenotypes

To investigate the relationship of gestational phenotypes with long-term health, we examined causal relationships between the 122 gestational phenotypes and 220 older-age phenotypes in the BioBank Japan (BBJ) cohort[12]. Using two-sample, inverse-variance-weighted, bidirectional MR and four additional sensitivity methods, we identified 73 significant causal associations after excluding outliners with MR-PRESSO[43] and adjusting for pleiotropy (false discovery rate [FDR] $q < 0.01$). These associations were robust to reverse MR, indicating unidirectional causality (**Methods**). The causal effect estimates are summarized in a network (**Fig. 5A),** with detailed forward MR results for 88 molecular phenotypes in **Fig. 5B and Supplementary Table 34,** and 34 medical record phenotypes in **Supplementary Fig. 33 and Supplementary Table 35**.

We identified several previously unreported causal associations. Genetically predicated higher gestational BMI was found to contribute significantly to a range of older-age disorders and phenotypes examined in our study. It increased the risk for 13 older-age disorders or phenotypes, with the strongest associations observed for

chronic renal failure (*OR* [95%CI]: 1.68 [1.42-2.00]). Other significant associations included cardiovascular disorders such as myocardial infarction, peripheral artery disease, chronic heart failure, and unstable angina pectoris, as well as type 1 diabetes, cholelithiasis, the use of five medication classes, and three physical phenotypes. Conversely, higher gestational BMI was associated with a slight reduction in the risk of two disorders: pneumothorax and pulmonary tuberculosis (**Fig. 5A, Supplementary Fig. 33 and Supplementary Table 34**).

Additionally, MR analyses revealed that gestational blood lipid and glycemic biomarkers were causally associated with older-age disorders. For instance, genetically elevated triglyceride levels during pregnancy were associated with an increased risk of specific medication use, such as antithrombotic agents (*OR* [95%CI]: 1.09 [1.06-1.12], $P=1.24\times10^{-11}$) (**Fig. 5C, Supplementary Table 34**). Similarly, genetically elevated 1-hour glucose levels from oral glucose tolerance tests (OGTT1H) during pregnancy increased the risk of peripheral arterial disease (PAD) (*OR* [95%CI]: 1.24 [1.13-1.37], $P=1.24\times10^{-5}$) (**Fig. 5D, Supplementary Table 34**), consistent with prior reports linking glucose abnormalities to PAD[44]. Moreover, genetically elevated OGTT1H were associated with an increased risk of cataracts (*OR* [95%CI]: 1.15 [1.09-1.21], $P=3.21\times10^{-8}$) (**Supplementary Table 34**), corroborating findings from animal models[45].

Our analyses also highlighted the impact of gestational infections on disease liability. For instance, genetically increased positive hepatitis B core antibody (HBcAb) status during pregnancy increased the risk of gastric cancer (*OR* [95%CI]: 1.12 [1.06-1.19], $P=1.06\times10^{-4}$), chronic hepatitis B infection (*OR* [95%CI]: 1.31 [1.15-1.49], $P=5.27\times10^{-5}$), and hepatic cancer (*OR* [95%CI]: 1.24 [1.13-1.37], $P=5.02\times10^{-6}$) (**Fig. 5A, Supplementary Table 34**), suggesting a potential causal link between gestational hepatitis B infection and extrahepatic cancers. Moreover, genetically predicted

cytomegalovirus IgG (CMV_IgG_quan) levels were associated with increased risk of cardiac diseases such as angina pectoris (*OR* [95%CI]: 1.31 [1.16-1.49], $P=1.28\times10^{-5}$) and stable angina pectoris (*OR* [95%CI]: 1.27 [1.16-1.40]), as well as medication use of vasodilators for cardiac conditions (*OR* [95%CI]: 1.26 [1.15-1.38]). These findings align with prior studies linking cytomegalovirus infection to coronary artery and vascular atherosclerosis[46,47].

To investigate the direct causal pathways linking gestational phenotypes (collected from pregnant women, mean age 30) to later-age phenotypes (collected from BBJ participants, aged 30-70), we performed multivariable MR (MVMR) analyses for 23 phenotypes, incorporating GWAS data from the Taiwan Biobank as a secondary exposure to account for intermediate phenotypes. Using an FDR threshold ($q < 0.01$), we identified 45 significant causal effects of gestational phenotypes on later-life outcomes, after adjusting for genetic effects from the Taiwan Biobank. Among these, 40 effects (q value > 0.05 in the Taiwan Biobank, or with opposite genetic effects) suggested direct causal effects on the outcomes. The remaining five effects, with nominal significance and consistent effect direction in the Taiwan Biobank, indicated weaker direct effects (**Supplementary Table 36, Supplementary Fig. 34**). Interestingly, we also identified 97 effects with $q < 0.01$ in the Taiwan Biobank but with opposite beta directions or $q > 0.05$ in the NIPT dataset, indicating that the NIPT effects on later-life outcomes are predominantly mediated through middle- and older-age phenotypes. Notable direct effects included genetically predicted second-trimester AFP and AFPMOM levels, which were negatively associated with atopic dermatitis and all types of white blood cell counts but positively associated with thyroid disorders (e.g., Graves' disease, Hashimoto's thyroiditis), medication use for thyroid conditions, chronic sinusitis, type 2 diabetes, and diabetes medications (**Supplementary Fig. 34 A-B, Supplementary Table 36**). Additionally, genetically predicted albumin levels during pregnancy were inversely associated with chronic

hepatitis B infection and cirrhosis. Gestational AST and RBC levels were also causally linked to corresponding blood cell traits in the BBJ (**Supplementary Fig. 34 C, Supplementary Table 36**).

A comparison of univariate MR results for the 23 gestational phenotypes and Taiwanese phenotype demonstrated high concordance ($R^2 = 0.79$), supporting the robustness of our findings. Collectively, these results suggest that the majority of gestational phenotype effects on later-life outcomes are mediated through intermediate-age phenotypes, while a substantial subset reflects direct causal relationships (**Supplementary Fig. 35, Supplementary Table 37-38**).

Discussion

This study represents the largest genetic investigation of gestational phenotypes to date, providing novel insights into the dynamic genetic atlas of human gestation and its causal links to long-term health outcomes. Leveraging obstetric and NIPT sequencing data from 121,579 pregnant women, we identified 3,845 genome-wide significant loci, including 1,893 novel loci, the majority of which contain at least one functional variant. Over 95% of the genetic associations were replicated in independent cohorts. Genes associated with multiple gestational phenotypes were enriched in pathways related to hormonal and immune responses, compound biosynthesis, and multi-organism reproductive processes. PPI network analysis highlighted *IL6* (involved in immune responses), *ESR1* (related to hormone responses), and *IGF1* and *EGF* (associated with cell growth and development) as central hub genes, underscoring the critical roles of these biological processes in gestation. These findings provide unprecedented insights into the genetic regulation of gestational processes in humans, extending earlier insights derived from metabolomic studies in primates[5].

The genetic atlas presented in this study is dynamic, reflecting gestation-specific gene-environment ($G \times E$) and gene-time ($G \times T$) interactions. Using data from the Taiwan Biobank and Bayesian classification analysis, we found that approximately 17.4% of genetic associations (138/792) across 23 gestational phenotypes demonstrated gestation-specific genetic effects. These effects ranged from nearly 0% for early pregnancy systolic blood pressure to 100% for second trimester alpha-fetoprotein levels. Genes with gestation-specific effects were predominantly enriched in pathways related to cancer, hormonal responses, lipid, vitamin, and glucose metabolisms, and cell growth and development. PPI network analysis identified *ESR1*, *IGF1* (but not *EGF*) and *IL6* as key hub genes. Additionally, longitudinal trajectory GWAS of repeated measures for 24 complete blood count (CBC) phenotypes revealed significant $G \times T$ interactions in 17.8% of genetic variants (115/645) over five gestational periods, with PPI network analysis identifying *EGF* and *CATA2* as key hub genes enriched in pathways related to cell growth regulation, nerve development and lipid modification.

Gestational conditions present unique opportunities for early screening and prediction of late-onset chronic disease risks in women[48]. Our findings revealed novel causal relationships between gestational phenotypes and long-term health outcomes. For instance, genetically determined BMI was associated with renal and cardiovascular disorders, while blood glucose abnormalities during pregnancy were linked to increased risk of peripheral artery disease and cataracts. Prior studies have demonstrated that pregnant women with gestational diabetes mellitus, even in the absence of type 2 diabetes, are at an elevated risk of cataracts[49]. Furthermore, multivariable Mendelian randomization analyses suggested that elevated AFP levels during gestation have direct effects on atopic dermatitis and thyroid disorders. These findings emphasize the long-term implications of gestation conditions for women's health.

While this study provides significant contributions, several limitations should be addressed in future research. First, although this is the largest GWAS of gestational phenotypes conducted to date, the participants were exclusively of East-Asian ancestry. As a result, some genetic discoveries may be population-specific and may not fully translate to non-East-Asian populations. However, the methodologies developed in this study offer a robust foundation for extending research to diverse ancestries, which is increasingly feasible as NIPT and pregnancy screening programs become globally widespread[50]. Future cross-population studies could elucidate ethnic differences in gestational phenotypes and their relations with offspring health[51]. Cross-ancestral analyses will also aid in fine-mapping and validating the functional variants identified in this study.

Second, while we identified gestation-specific and time-varying genetic effects for approximately 17% of the genetic associations, the precise roles of novel genes in immune and hormonal responses, and cell-growth and development during gestation, remain to be fully understood. Future mechanistic studies are required to explore how these genes influence the initiation and progression of gestation and contribute to adverse outcomes, conditioning on the findings in this study. Third, our investigations primarily focused on the causal relationships between gestational phenotypes and chronic diseases in later life. Future studies should prioritize efforts to integrate our findings into expanding birth cohort genome studies, which serves unique opportunities to dissecting the intergenerational impacts of gestational genetics to provide insights into developmental origins of health and diseases[16].

This study offers foundational insights into the dynamic genetic architecture of the human gestational period. To facilitate future research, we have made visualization of the GWAS results available via the MONN PheWeb Websever

(https://monn.pheweb.com/). As genomic datasets and analytical tools within MONN continue to expand, future efforts will focus on leveraging genetic information to explore the developmental origins of childhood and adult disorders. Additionally, integrating these findings into health and risk prediction models for women during gestation will be a priority. These endeavors will require coordinated global efforts to maximize the potential of these discoveries.

Methods

Study population

Between 2017 and 2022, we recruited 121,579 pregnant women who attended maternity checkups at Shenzhen Baoan Women's and Children's Hospital and Longgang District Maternity and Child Healthcare Hospital in Shenzhen, China. Participants underwent prenatal and postnatal serum biomarkers testing and non-invasive prenatal testing (NIPT) covered by their medical insurance program. Maternal genotypes were inferred from the low-depth whole-genome sequencing data generated from NIPT utilizing advanced methods developed in our laboratory[7,20]. In addition, we collected basic physical measurements of the participants, as well as gestational complications and birth outcomes from their electronic medical records. The details of participant numbers for each test are provided in **Supplementary Table 2 and 3**. The age distribution of the pregnant women and the prevalence of certain gestational comorbidities at the two hospitals are shown in **Supplementary Table 4-5, and Supplementary Fig. 1-2**.

We integrated the obstetrical data, including both genotypic and phenotypic information, to explore the genetic and molecular risk factors underlying the phenotypic spectrum of gestational biomarkers. This study was approved by the Medical Ethics Committee of the School of Public Health (Shenzhen), Sun Yat-sen University, Longgang District Maternity and Child Healthcare Hospital of Shenzhen City, and

Shenzhen Baoan Women's and Children's Hospital. Data collection was authorized by the Human Genetic Resources Administration of China (HGRAC).

Phenotype definition

We classified the 122 gestational phenotypes in nine molecular test categories: Blood lipid (N=2), Glycemic (N=6), Blood routine (N=32), Infection (N=14), Kidney function (N=3), Liver function (N=8), Thyroid function (N=4), Tang screening (N=16) and NIPT screening (N=3), as well as three medical record categories including Basic information (N=7) , Gestational disorders (N=21) and Birth outcomes (N=6) (**Supplementary Table 2-3**). 24 complete blood count phenotypes have an average five time point measurements while the rest of the phenotypes usually have a single time-point measurement.

Genotype imputation from NIPT data and quality control

NIPT for fetal trisomy by sequencing maternal plasma free DNA (cfDNA) is an advanced genomic sequencing technology[52]. Large-scale low-pass and mediate-coverage sequencing data are readily available through NIPT. In our previous study, we demonstrated that NIPT data can be effectively utilized for GWAS analysis, generating highly accurate genetic discoveries and effect estimates[7,20]. In brief, we performed genotype imputation from NIPT data using genotype likelihoods imputation and phasing method with the GLIMPSE software (version 1.1.1)[53] , employing a reference panel of high-depth sequencing genotypes from 10,000 Chinese individuals. Genotype imputation accuracy was evaluated using 100 participants with high-depth sequencing, revealing an average $R^2 > 0.84$ for variants with a minor allele frequency MAF > 0.001.

Family relatedness and principal component analyses

PLINK[54] (version 2.0) was used to select SNPs with an MAF of at least 0.5%. Kinship coefficient was calculated using the KING-robust kinship estimator, with a cutoff of 0.177 to exclude monozygotic twins, duplicate samples and first-degree relatives. Principal component analysis (PCA) was conducted using PLINK[54] (version 2.0) after genotype imputation.

Variant annotation

Variant annotation was performed using the Ensembl Variant Effect Predictor (VEP) (version 101)[55], with indexed GRCh38 cache files (version 109). HGVS notations were generated from the primary reference FASTA files for *Homo sapiens*. Since a variant may overlap with multiple transcripts, the --pick option was used to assign a single result to each variant based on the default criteria set by VEP. For variants located in intergenic regions, the --nearest option was applied to identify the nearest gene to the protein-coding transcription start site (TSS).

Genome-wide association analysis and conditional analysis

After quality control of the sequencing data and genotype imputation from the NIPT data, we conducted a genome-wide association analysis (GWAS) using PLINK[54] (version 2.0) to examine the associations between 122 phenotypes, assuming additive genetic effects. For quantitative traits, we employed a linear regression model, while for qualitative traits, we applied a logistic regression model. To account for population stratification, we included relevant covariates in the GWAS for all phenotypes, such as maternal age, body mass index (BMI), gestational week at the time of phenotype measurement, and the top ten principal components. A list of all GWAS covariates is provided in **Supplementary Table 6**.

Meta-analyses were conducted using GWAS summary statistics from both hospitals, utilizing the METAL software[56] (version 2011-03-25), with an inverse-variance

weighted (IVW) fixed-effect model. For visualization of the meta-analyses results, we generated fuji plot using R script and Circos software[57,58]. The GWAS meta-summary statistics were subsequently used for all further analyses.

To identify independent genome-wide significant signals, we performed multi-SNP-based conditional & joint association analysis (COJO[21]) using GCTA[59] with a stepwise model selection approach (--cojo-slct). Variants with MAF < 0.01 were excluded, and a colinearity threshold of 0.2 was applied. A 500kb region upstream and downstream of each genome-wide significant signal was considered a separate locus, with the SNP exhibiting the smallest *P*-value selected as the lead SNP for each locus. Variants with *P*-value less than 4.1×10^{-10} were considered to have reached study-wide significance.

Replication of independent genome-wide significant signals

To assess whether the genetic effects of the same phenotype were consistent across the two cohorts, we performed both internal replication and external comparison of the independent genome-wide significant SNPs identified from conditional analysis using two methods. In the internal replication, we compared the direction of the beta coefficients and *P*-value for each SNP across the two hospitals. A SNP was considered replicated if it exhibited consistent beta direction, and the *P*-value was below the Bonferroni-corrected significance threshold (0.05/number of independent SNPs per phenotype). In the second approach, based on the central limit theorem, we employed a *two-sample t-test* with the following hypotheses:

$$Null\ hypothesis\ H_0: \beta_1 = \beta_2$$

$$Alternative\ hypothesis\ H_1: \beta_1 \neq \beta_2$$

The test statistic (*T)* was calculated as:

$$T = \frac{\beta_1 - \beta_2}{\sqrt{SE_1^2 + SE_2^2}} \sim t(v') \quad (1)$$

The degrees of freedom v' were calculated by the formula:

$$v' = \frac{(SE_1^2 + SE_2^2)^2}{\frac{SE_1^4}{n_1 - 1} + \frac{SE_2^4}{n_2 - 1}} \quad (2)$$

In these formulas, β represents the estimate of the genetic effect, *SE* represents the standard error of the genetic effect estimate, and *n* represents the sample size for the phenotype in each cohort. The *P*-value was derived from the test statistic (*T)* and the degrees of freedom (v'). A SNP was considered replicated if the *P*-value was greater than 0.05. Additionally, we used METAL to perform heterogeneity tests for the genetic effects of SNPs between the two hospitals, with *P* value ≥ 0.05 indicating no heterogeneity in genetic effects across the hospitals.

For external replication, we utilized the NIPT PLUS pregnancy cohort and the Baoan 20K pregnancy cohort, which consists of data from 26,624 participants, who sought maternal check-ups at Shenzhen Baoan Women's and Children's Hospital (Shenzhen, China) throughout the entire 40-week gestational period after the year of 2022. Potential duplicate samples with the discovery sets were excluded based on their unique identity number. Among the 122 phenotypes analyzed in this study, 107 phenotypes are shared with the replication cohort. We applied the same methods used for internal replication to perform external replication of the variants for these 107 phenotypes.

Comparison of genetic effects on maternal height and Taiwanese population height

Given that maternal height remains largely unchanged during gestation, we validate the effectiveness of our GWAS results by comparing the genetic effects of maternal height with those observed for height in the Taiwanese population[22]. We extracted genome-wide significant independent variants associated with height in the Taiwanese population and identified corresponding genetic effects in our GWAS summary

statistics. Comparisons were made using the Bonferroni correction threshold method, and the replication rate was calculated.

Identification of novel loci and signal

We downloaded the complete association dataset (v1.0.2_e109_r2023-02-15) from the GWAS Catalog[23] , and incorporated significant variants from three recently published pregnancy-related studies[10,11,60], to identify novel associated SNPs. A locus was considered novel if the ±500kb region surrounding the lead SNP did not overlap with previously reported associations for the same trait in the GWAS Catalog. Additionally, an association signal was classified as novel if the LD R^2 value between SNPs reported in GWAS Catalog and our identified SNP was less than 0.2. LD R^2 calculations were performed using the LDpair Tool through LDlink[61] (version 5.5.1), based on the GRCh38 1000G genome build for both the East Asian (EAS) and European (EUR) populations.

Protein-protein interaction network analysis

Genes containing loci identified in our study were selected for protein-protein interaction (PPI) analysis. Genes with PPI score > 0.4 were included in the network construction, which was visualized using Cytoscape (Version:3.10.1)[62]. The degree algorithm, implemented in CytoHubba[63] , was used to identify the top 10 hub genes. Default software parameters were applied for all analyses.

Genetic heritability and correlation estimates

We standardized the phenotypic data and computed the Spearman correlation coefficients to generate the phenotypic correlation matrix. Meta-GWAS summary statistics were then utilized to estimate SNP-based heritability and genetic correlation (r_g) for all phenotypes using LD Score Regression[36,64]. LD score files were calculated with GCTA[65] , using high-depth sequencing genotypes from the same 10,000 Chinese

reference panel for genotype imputation. The genome was divided into 1000kb segments (--ld-wind 1000), and an LD R^2 cutoff of 0.01 was applied (--ld-rsq-cutoff 0.01). The phenotypic correlation matrix and genetic association matrix for both hospitals were extracted to create heat maps for visualization.

Detection of gestation-specific effects

Considering the influence of population structure, we used GWAS summary statistics for 23 shared phenotypes from the Taiwan Biobank[22], where participants are also of southern Chinese ancestry. Independent SNPs were eligible for comparison if they met the following criteria: 1) the SNP was present in the external GWAS summary statistics; 2) the effect allele and the direction of the genetic effect matched those in the external data; and 3) the *P*-value was below the Bonferroni-corrected threshold (0.05/number of loci).

We applied the linemodels package[39] (https://github.com/mjpirinen/linemodels) to analyze GWAS summary statistics from gestational phenotypes and the same phenotypes in the general, non-pregnant Taiwan Biobanks. This analysis included 792 lead variants of independent signals across 23 phenotypes (**Supplementary Table 22**). The lead variants were classified into two clusters based on their effect sizes: gestation-specific variants and general variants. Variant with significantly different genetic effects compared to the general population were classified as gestation-specific (posterior probability > 0.95), while those with consistent genetic effects were classified as general variants (posterior probability > 0.95). The remaining variants were classified as uncertain. These clusters were represented by line models where the slope of the gestational cluster was fixed at 0, correlations were constrained to 0.999 in both clusters, and other parameters were estimated using the EM algorithm. Parameters, including slope (representing the gradient of the two lines), scale (indicating effect size

magnitude), and correlation (defining allowed deviations), are detailed in **Supplementary Table 23**.

To account for LD structure on classification, we further conducted colocalization analyses of gestational phenotype variants and variants associated with the same phenotypes in Taiwan Biobank. Adjustments to the *linemodels* classification were made based on colocalization results. The prior probabilities for a genetic locus being associated with gestational phenotypes (p_1) and phenotypes in Taiwan Biobank (p_2), or both (p_{12}) were set at 1×10^{-4} , 1×10^{-4}, and 5×10^{-6} , respectively [66]. Adjustments to classifications were based on the following criteria:

(1) If a variant classified into the gestation cluster in *linemodels* had a colocalization posterior probability (PPH4) < 0.2, its classification remained unchanged. Otherwise, it was classified as uncertain.

(2) If a variant in a general cluster had a PPH4 > 0.8, its classification remained unchanged. Otherwise, it was classified as uncertain.

(3) Variants in the *linemodel* uncertain cluster were reclassified as gestation-specific if their PPH4 < 0.2, or as general if their PPH4 > 0.8.

(4) In all other cases, the classification results from linemodels were retained.

Pathway enrichment analysis

We conducted pathway enrichment analysis using Metascape (v3.5.20240101)[67]. This analysis calculates the hit rate for our gene list and compares it with the background hit rate based on the number of genes in each category, with the enrichment factor defined as the ratio of these two rates. The *P*-value represents the probability of observing multiple pathways by chance, modeled using a cumulative hypergeometric distribution. Lower *P*-values indicate that the observed enrichment is less likely due to randomness.

The analysis was performed using default parameters: the minimum overlap was set to 3, the *P*-value cutoff was 0.01, and the minimum enrichment factor was 1.5. The ontology categories included Gene Ontology (GO) Biological Processes, Reactome Gene Sets, KEGG Pathways, WikiPathways, Canonical Pathways, and PANTHER Pathways.

Detection of G by T interaction effects

We applied TrajGWAS[42] to identify genetic variants associated with changes in gestational phenotypes. For 645 associated signals, we applied the following mixed effects model with an interaction term of gene (G) and gestational day (T):

$$y_{ij} = x_{ij}\beta + g_i\beta_g + G \times T \times \beta_{G\times T} + \varepsilon_{ij} \quad (3)$$

y_{ij} represents the phenotypic value of the jth measurement for the ith pregnant woman. x_{ij} represents fixed-effects covariates, including the first ten principal components, maternal age and gestational days at phenotypic testing, β is its corresponding effect value. g_i denotes the genotype dosage for the ith pregnant woman, corresponding to a genetic effect of β_g. $G \times T$ represents the interaction between maternal genotype dosage and gestational day, corresponding to an effect value of $\beta_{G\times T}$. ε_{ij} is a random-effect term reflecting the variance of the genetic variants affecting the within-subject (WS) variance of the biomarker. The formula for the variance $\sigma^2_{\varepsilon_{ij}}$ of ε_{ij} is given below:

$$\sigma^2_{\varepsilon_{ij}} = \exp(\tau w_{ij} + \tau_g g_i + \tau_{G\times T} G \times T + \omega_i) \quad (4)$$

w_{ij} is a covariate including the first 10 principal components, maternal age, and gestational days at phenotypic measurement, and τ is the corresponding effect value. τ_g denotes the genetic effect affecting the variance of WS. $\tau_{G\times T}$ indicates the effect value for the interaction of maternal genotype dosage and gestational time. ω_i denotes the random intercept.

We analyzed 24 repeated measure CBC phenotypes (**Supplementary Table 27)**. Following the approach by Ko et al., we first performed a score test for SNPs to determine the direction of the genetic effect (β_g or τ_g) and calculated *P*-value for mean and within-subject variability. SNPs with a *P*-value of β_g or τ_g less than 5×10^{-8} underwent a Wald test to estimate effect sizes for genotype dosage (β_g or τ_g) and gene-time interaction ($\beta_{G\times T}$ or $\tau_{G\times T}$). Significant SNPs were subjected to meta-analysis using METAL[56]. Manhattan and QQ plots were generated using the Julia package MendelPlots.jl (https://github.com/OpenMendel/MendelPlots.jl)[68] , considering SNPs with minor allele frequency>0.01. Independent significant loci were identified using a Python script, where the lead variant with the smallest *P* value in a 1Mb window was selected, and variants within ±500kb of the lead variant were grouped into the same locus.

To examine how genetic effects varied across the five periods- first, second, and third trimester, delivery day, and within 42 days postpartum- we conducted localized association analyses for identified loci. The methodology followed previous GWAS protocols, utilizing PLINK[54] (version 2.0) with the parameters --chr, --from-bp, and --to-bp to target specific loci. A variant was classified as exhibiting temporal change if its 95% confidence interval (CI) for genetic effects were non-overlapping in at least two of the five periods.

Mendelian randomization network of pregnancy, Taiwanese, and BBJ traits

To investigate whether pregnancy status influences the progression of future diseases, bidirectional Mendelian randomization (MR) analyses were conducted using data from the cohort of this study and the Japanese Biobank (BBJ), employing the TwoSampleMR package[69,70]. GWAS summary statistics from both cohorts were utilized, and independent SNPs identified through GCTA-COJO ($P < 5\times10^{-8}$) were used as genetic

instruments in the MR analyses. The proportion of phenotypic variance (PVE) explained by the instruments was calculated using the formula[71] (5). The F-statistic was calculated using the formula (6) to evaluate the strength of the instruments, with F values greater than 10 indicating sufficient strength.

$$PVE = \frac{2\hat{\beta}^2 MAF(1-MAF)}{2\hat{\beta}^2 MAF(1-MAF) + \left(se\left(\hat{\beta}\right)\right)^2 2NMAF(1-MAF)} \quad (5)$$

$$F = \frac{N - nsnp - 1}{nsnp} \times \frac{PVE}{1 - PVE} \quad (6)$$

Assessments of instrumental variables are provided in **Supplementary Table 39-42**. Sensitivity analyses were preformed using alternative MR methods including: 1) MR Egger; 2) Simple mode; 3) Weighted median; and 4) Weighted mode. Among these methods, the inverse variance weight (IVW) method, which provides the greatest statistical power for estimating causal associations, was used as the primary reference[72].

To address horizontal pleiotropy in causal relationships, we applied MR-PRESSO[43]. Outliers identified by MR-PRESSO were tested for differences between pre-correction and post-correction results. If significant differences were detected the outliers were excluded, the two-sample MR was repeated.

Robust causal estimates were defined as those satisfying the following criteria: significant at an FDR-adjusted IVW q value [73] ($q < 0.01$); No evidence of heterogeneity ($P < 0.05$); and no evidence of horizontal pleiotropy ($P < 0.05$) after MR-PRESSO.

To account for the potential influence of mid-life health status on causal relationships, we conducted multivariable MR analyses[74,75] incorporating 23 phenotypes from the Taiwan Biobank[22] as secondary exposures. In these analyses, gestational phenotypes and Taiwan Biobank phenotypes were used as exposures, while the 220 phenotypes

from the BBJ were used as outcomes. The TwoSampleMR package was used for these analyses, with a significance threshold of $q < 0.01$.

Data availability

The full GWAS summary data from this study have been deposited in the GWAS catalog database (https://www.ebi.ac.uk/gwas/) under accession numbers (GCSTXXXXXX), with approval from the China's National Health Commission (permission number: 2024SQXB002074). Raw sequencing data have been deposited in the Genome Sequence Archive (GSA) for Human at the National Genomics Data Center82,83, under the BioProject accession number GSA-Human: HRA006833 with approval from China’s National Health Commission (permission number: 2024BAT01079). Data can be accessed via applications, following the GSA guide (https://ngdc.cncb.ac.cn/gsa-human/document). The access authority can be obtained for academic research use only.

Contributions

Design, management, or subject recruitment of the individual studies: S.L., F.W., J.Z., S.H., and data curation: J.Z., Y.W., Z.Y., H.Z., L.H., S.C., and genotyping of the individual studies: Y.W., Y.G., Y.L., Z.Y., and statistical methods, analysis, bioinformatics, or interpretation of the results in the individual studies: S.L., H.Z., Y.G., Z.Y., Y.L., Y.W., X.G, Y.C., and providing professional guidance and

interpretation of data: S.L., F.W., J.Z., S.H., G.C., X.Q., Q.Z., and drafting of manuscript: S.L., H.Z., and critical revision of manuscript: all authors.

Acknowledgement

The study was supported by National Natural Science Foundation of China (32470642, 32470679, 31900487, 82203291, 82173525), Shenzhen Science and Technology Program (20220818100717002, ZDSYS20230626091203007), Guangdong Basic and Applied Basic Research Foundation (2022B1515120080, 2020A1515110859), the Shenzhen Health Elite Talent Training Project and the High-performance Computing Public Platform (Shenzhen Campus) of SUN YAT-SEN UNIVERSITY.

Reference

1. Stock, S.J. & Aiken, C.E. Barriers to progress in pregnancy research: How can we break through? *Science* **380**, 150-153 (2023).
2. Qiao, J. *et al.* A Lancet Commission on 70 years of women's reproductive, maternal, newborn, child, and adolescent health in China. *Lancet* **397**, 2497-2536 (2021).
3. Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950-2019: a comprehensive demographic analysis for the Global Burden of Disease Study 2019. *Lancet* **396**, 1160-1203 (2020).
4. Chen, X. *et al.* The path to healthy ageing in China: a Peking University-Lancet Commission. *Lancet* **400**, 1967-2006 (2022).
5. Yu, D. *et al.* A multi-tissue metabolome atlas of primate pregnancy. *Cell* **187**, 764-781.e14 (2024).
6. Hunter, P.J. *et al.* Biological and pathological mechanisms leading to the birth of a small vulnerable newborn. *Lancet* **401**, 1720-1732 (2023).
7. Liu, S. *et al.* Genomic Analyses from Non-invasive Prenatal Testing Reveal Genetic Associations, Patterns of Viral Infections, and Chinese Population History. *Cell* **175**, 347-359.e14 (2018).
8. Zhen, J. *et al.* Genome-wide association and Mendelian randomisation analysis among 30,699 Chinese pregnant women identifies novel genetic and molecular risk factors for gestational diabetes and glycaemic traits. *Diabetologia* **67**, 703-713 (2024).
9. Yang, Z. *et al.* Genetic Basis of Altered Platelet Counts and Gestational Thrombocytopenia in Pregnancy. *Blood* (2023).
10. Xiao, H. *et al.* Genetic analyses of 104 phenotypes in 20,900 Chinese pregnant women reveal pregnancy-specific discoveries. *Cell Genom* **4**, 100633 (2024).

11. Guo, J. *et al.* Phenome-wide association study in 25,639 pregnant Chinese women reveals loci associated with maternal comorbidities and child health. *Cell Genom* **4**, 100632 (2024).
12. Sakaue, S. *et al.* A cross-population atlas of genetic associations for 220 human phenotypes. *Nat Genet* **53**, 1415-1424 (2021).
13. Graham, S.E. *et al.* The power of genetic diversity in genome-wide association studies of lipids. *Nature* **600**, 675-679 (2021).
14. Elliott, A. *et al.* Distinct and shared genetic architectures of gestational diabetes mellitus and type 2 diabetes. *Nat Genet* **56**, 377-382 (2024).
15. Juliusdottir, T. *et al.* Distinction between the effects of parental and fetal genomes on fetal growth. *Nat Genet* **53**, 1135-1142 (2021).
16. Huang, S. *et al.* The Born in Guangzhou Cohort Study enables generational genetic discoveries. *Nature* **626**, 565-573 (2024).
17. Mitha, A. *et al.* Neurological development in children born moderately or late preterm: national cohort study. *Bmj* **384**, e075630 (2024).
18. Crump, C. *et al.* Adverse pregnancy outcomes and long term risk of ischemic heart disease in mothers: national cohort and co-sibling study. *Bmj* **380**, e072112 (2023).
19. Liu, S. *et al.* Utilizing Non-Invasive Prenatal Test Sequencing Data Resource for Human Genetic Investigation. *bioRxiv*, 2023.12. 11.570976 (2023).
20. Liu, S. *et al.* Utilizing non-invasive prenatal test sequencing data for human genetic investigation. *Cell Genom* **4**, 100669 (2024).
21. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat Genet* **44**, 369-75, s1-3 (2012).
22. Chen, C.Y. *et al.* Analysis across Taiwan Biobank, Biobank Japan, and UK Biobank identifies hundreds of novel loci for 36 quantitative traits. *Cell Genom* **3**, 100436 (2023).

23. Sollis, E. *et al.* The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Res* **51**, D977-d985 (2023).
24. Kelleher, A.M., DeMayo, F.J. & Spencer, T.E. Uterine Glands: Developmental Biology and Functional Roles in Pregnancy. *Endocr Rev* **40**, 1424-1445 (2019).
25. Mills, M.C. *et al.* Identification of 371 genetic variants for age at first sex and birth linked to externalising behaviour. *Nat Hum Behav* **5**, 1717-1730 (2021).
26. Gibson, D.A. *et al.* Profiling the expression and function of oestrogen receptor isoform ER46 in human endometrial tissues and uterine natural killer cells. *Hum Reprod* **35**, 641-651 (2020).
27. Schlaikjær Hartwig, T. *et al.* Cell-free fetal DNA for genetic evaluation in Copenhagen Pregnancy Loss Study (COPL): a prospective cohort study. *Lancet* **401**, 762-771 (2023).
28. Chitty, L.S. Use of cell-free DNA to screen for Down's syndrome. *N Engl J Med* **372**, 1666-7 (2015).
29. Wang, W. *et al.* Homozygous variants in PANX1 cause human oocyte death and female infertility. *Eur J Hum Genet* **29**, 1396-1404 (2021).
30. Sang, Q. *et al.* A pannexin 1 channelopathy causes human oocyte death. *Sci Transl Med* **11**(2019).
31. Brinkmann, V. *et al.* Neutrophil extracellular traps kill bacteria. *Science* **303**, 1532-5 (2004).
32. Bouvier, S. *et al.* NETosis Markers in Pregnancy: Effects Differ According to Histone Subtypes. *Thromb Haemost* **121**, 877-890 (2021).
33. Guillotin, F. *et al.* Vital NETosis vs. suicidal NETosis during normal pregnancy and preeclampsia. *Front Cell Dev Biol* **10**, 1099038 (2022).
34. Yu, G. *et al.* APRIL and TALL-I and receptors BCMA and TACI: system for regulating humoral immunity. *Nat Immunol* **1**, 252-6 (2000).

35. Duan, T. *et al.* USP3 plays a critical role in the induction of innate immune tolerance. *EMBO Rep* **24**, e57828 (2023).
36. Bulik-Sullivan, B.K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* **47**, 291-5 (2015).
37. lab, N. UK Biobank-Neale lab. (2018).
38. Dalmia, A. *et al.* A genetic epidemiological study in British adults and older adults shows a high heritability of the combined indicator of vitamin B(12) status (cB(12)) and connects B(12) status with utilization of mitochondrial substrates and energy metabolism. *J Nutr Biochem* **70**, 156-163 (2019).
39. Pirinen, M. linemodels: clustering effects based on linear relationships. *Bioinformatics* **39**(2023).
40. Li, A., Yang, S., Zhang, J. & Qiao, R. Establishment of reference intervals for complete blood count parameters during normal pregnancy in Beijing. *J Clin Lab Anal* **31**(2017).
41. Zhu, J., Li, Z., Deng, Y., Lan, L. & Yang, J. Comprehensive reference intervals for white blood cell counts during pregnancy. *BMC Pregnancy Childbirth* **24**, 35 (2024).
42. Ko, S. *et al.* GWAS of longitudinal trajectories at biobank scale. *Am J Hum Genet* **109**, 433-445 (2022).
43. Verbanck, M., Chen, C.Y., Neale, B. & Do, R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet* **50**, 693-698 (2018).
44. Bonaca, M.P., Hamburg, N.M. & Creager, M.A. Contemporary Medical Management of Peripheral Artery Disease. *Circ Res* **128**, 1868-1884 (2021).
45. Ranaei Pirmardan, E. *et al.* Pre-hyperglycemia immune cell trafficking underlies subclinical diabetic cataractogenesis. *J Biomed Sci* **30**, 6 (2023).

46. Jonasson, L., Tompa, A. & Wikby, A. Expansion of peripheral CD8+ T cells in patients with coronary artery disease: relation to cytomegalovirus infection. *J Intern Med* **254**, 472-8 (2003).

47. Borgia, M.C. *et al.* Further evidence against the implication of active cytomegalovirus infection in vascular atherosclerotic diseases. *Atherosclerosis* **157**, 457-62 (2001).

48. Parikh, N.I. *et al.* Adverse Pregnancy Outcomes and Cardiovascular Disease Risk: Unique Opportunities for Cardiovascular Disease Prevention in Women: A Scientific Statement From the American Heart Association. *Circulation* **143**, e902-e916 (2021).

49. Auger, N., Tang, T., Healy-Profitós, J. & Paradis, G. Gestational diabetes and the long-term risk of cataract surgery: A longitudinal cohort study. *J Diabetes Complications* **31**, 1565-1570 (2017).

50. Norton, M.E. *et al.* Cell-free DNA analysis for noninvasive examination of trisomy. *N Engl J Med* **372**, 1589-97 (2015).

51. Patel, R.R., Steer, P., Doyle, P., Little, M.P. & Elliott, P. Does gestation vary by ethnic group? A London-based study of over 122,000 pregnancies with spontaneous onset of labour. *Int J Epidemiol* **33**, 107-13 (2004).

52. Cheung, S.W., Patel, A. & Leung, T.Y. Accurate description of DNA-based noninvasive prenatal screening. *N Engl J Med* **372**, 1675-7 (2015).

53. Rubinacci, S., Ribeiro, D.M., Hofmeister, R.J. & Delaneau, O. Efficient phasing and imputation of low-coverage sequencing data using large reference panels. *Nature Genetics* **53**, 120-126 (2021).

54. Chang, C.C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).

55. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol* **17**, 122 (2016).

56. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-2191 (2010).

57. Krzywinski, M. *et al.* Circos: an information aesthetic for comparative genomics. *Genome Res* **19**, 1639-45 (2009).

58. Kanai, M. *et al.* Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. *Nat Genet* **50**, 390-400 (2018).

59. Yang, J., Lee, S.H., Goddard, M.E. & Visscher, P.M. GCTA: A Tool for Genome-wide Complex Trait Analysis. *The American Journal of Human Genetics* **88**, 76-82 (2011).

60. Zhu, H. *et al.* Novel insights into the genetic architecture of pregnancy glycemic traits from 14,744 Chinese maternities. *Cell Genom* **4**, 100631 (2024).

61. Machiela, M.J. & Chanock, S.J. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* **31**, 3555-7 (2015).

62. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-504 (2003).

63. Chin, C.-H. *et al.* cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Systems Biology* **8**, S11 (2014).

64. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* **47**, 1236-41 (2015).

65. Yang, J. *et al.* Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nat Genet* **47**, 1114-20 (2015).

66. Wallace, C. Eliciting priors and relaxing the single causal variant assumption in colocalisation analyses. *PLoS Genet* **16**, e1008720 (2020).

67. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* **10**, 1523 (2019).
68. Zhou, H. *et al.* OPENMENDEL: a cooperative programming project for statistical genetics. *Hum Genet* **139**, 61-71 (2020).
69. Hemani, G. *et al.* The MR-Base platform supports systematic causal inference across the human phenome. *eLife* **7**(2018).
70. Li, J., Hemani, G., Tilling, K. & Davey Smith, G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLOS Genetics* **13**(2017).
71. Teslovich, T.M. *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707-13 (2010).
72. Slob, E.A.W. & Burgess, S. A comparison of robust Mendelian randomization methods using summary data. *Genetic Epidemiology* **44**, 313-329 (2020).
73. John D. Storey, A.J.B., Alan Dabney, David Robinson. qvalue: Q-value estimation for false discovery rate control. R package version 2.36.0 (2024).
74. Sanderson, E. Multivariable Mendelian Randomization and Mediation. *Cold Spring Harb Perspect Med* **11**(2021).
75. Vabistsevits, M. Setting up Multivariable Mendelian Randomization analysis in R. (22 March, 2021).

Figure Legends

Figure 1. Schematic overview of the study

We curated a comprehensive dataset encompassing 122 gestational phenotypes from 120,579 Chinese pregnant women. This study involved an extensive analysis of the genetic architecture of these phenotypes, identification of biological pathways enriched with associated genetic variants, and investigation of genetic variants influencing longitudinal trajectories of these phenotypes. Furthermore, we explored potential causal relationships between gestational phenotypes and later-age disorders.

Figure 2. Genetic associations across nine molecular test categories of gestational phenotypes

The top-left inset displays a histogram depicting the number of loci associated with each gestational phenotypes, categorized as novel or previously known loci. The main panel features a Fuji plot illustrating the GWAS results across all phenotypes within nine molecular test categories. Independent variants identified through meta-analysis ($P < 4.1 \times 10^{-10}$) are shown. The number of phenotypes associated with each pleiotropic locus is depicted using filled colored boxes, with functional and newly discovered loci labeled directly on the plot. Additional details on functional loci are provided in **Supplementary Table 13**. Pleiotropic and trait-specific associations are distinguished by circles of different sizes. The order of the 88 gestational phenotypes in both the histogram and Fuji plot corresponds to **Supplementary Table 2**.

Figure 3. Characteristics of genetic loci exhibiting gestation-specific genetic effects

Panel **A** displays a classification of the genetic effects of SNPs ($P < 4.1 \times 10^{-10}$, Bonferroni correction threshold) on gestational phenotypes using a Bayesian classification algorithm combined with colocalization analysis. Panel **B** features a butterfly bar plot highlighting the top 20 pathways significantly enriched for genes

with gestation-specific variants, identified through Metascape enrichment analysis. On the left, the plot depicts the enrichment *P* values for gestation-specific variants and the proportion of genes enriched in each pathway relative to the total number of pregnancy-specific genes. On the right, the corresponding results for general variants are shown. Pathways with significantly different gene enrichment, as determined by a chi-square test, are marked with an asterisk (*).

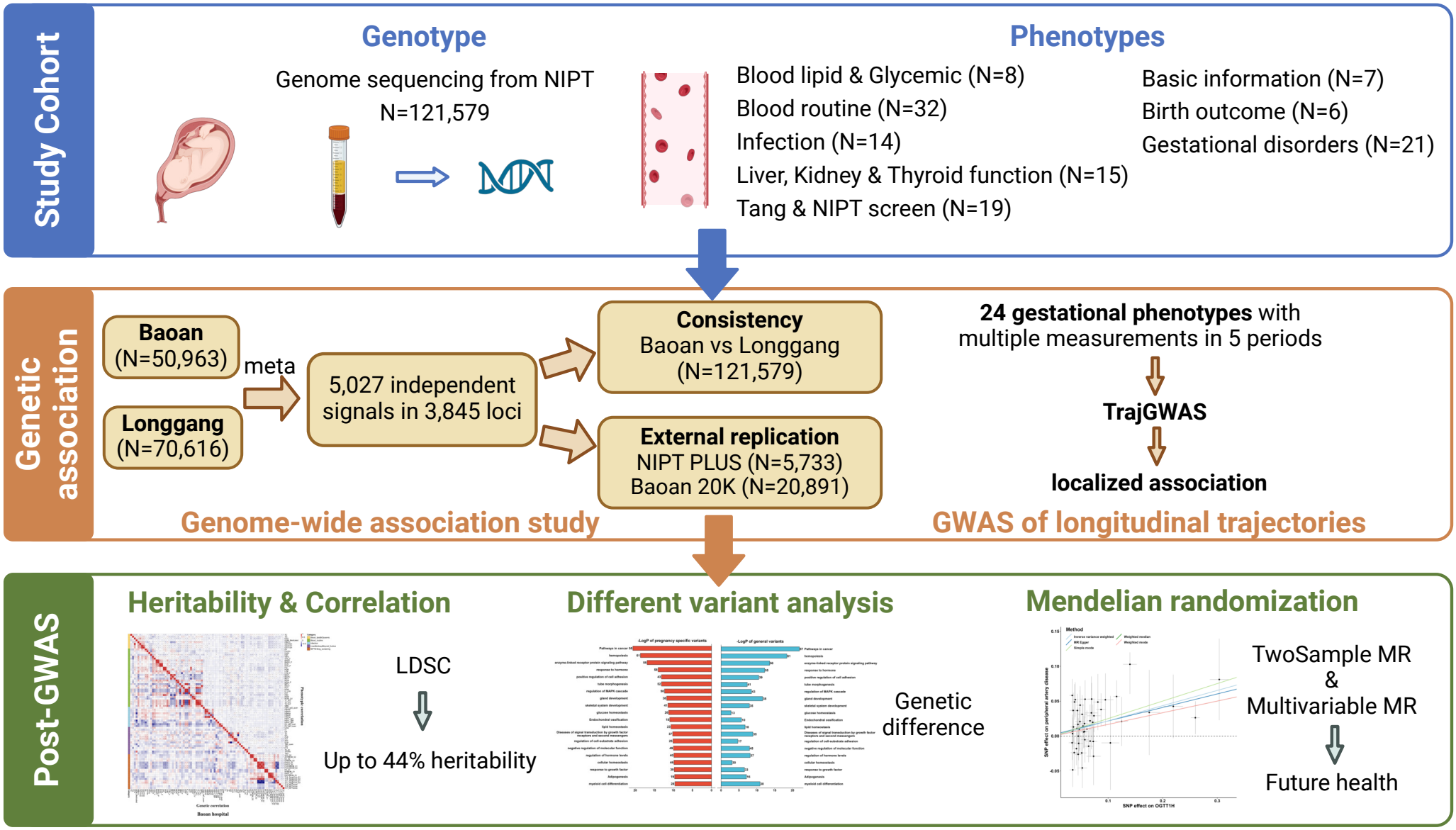
Figure 4. Gene-by-Time interactions for repeated measurements of 24 blood cell phenotypes during gestation

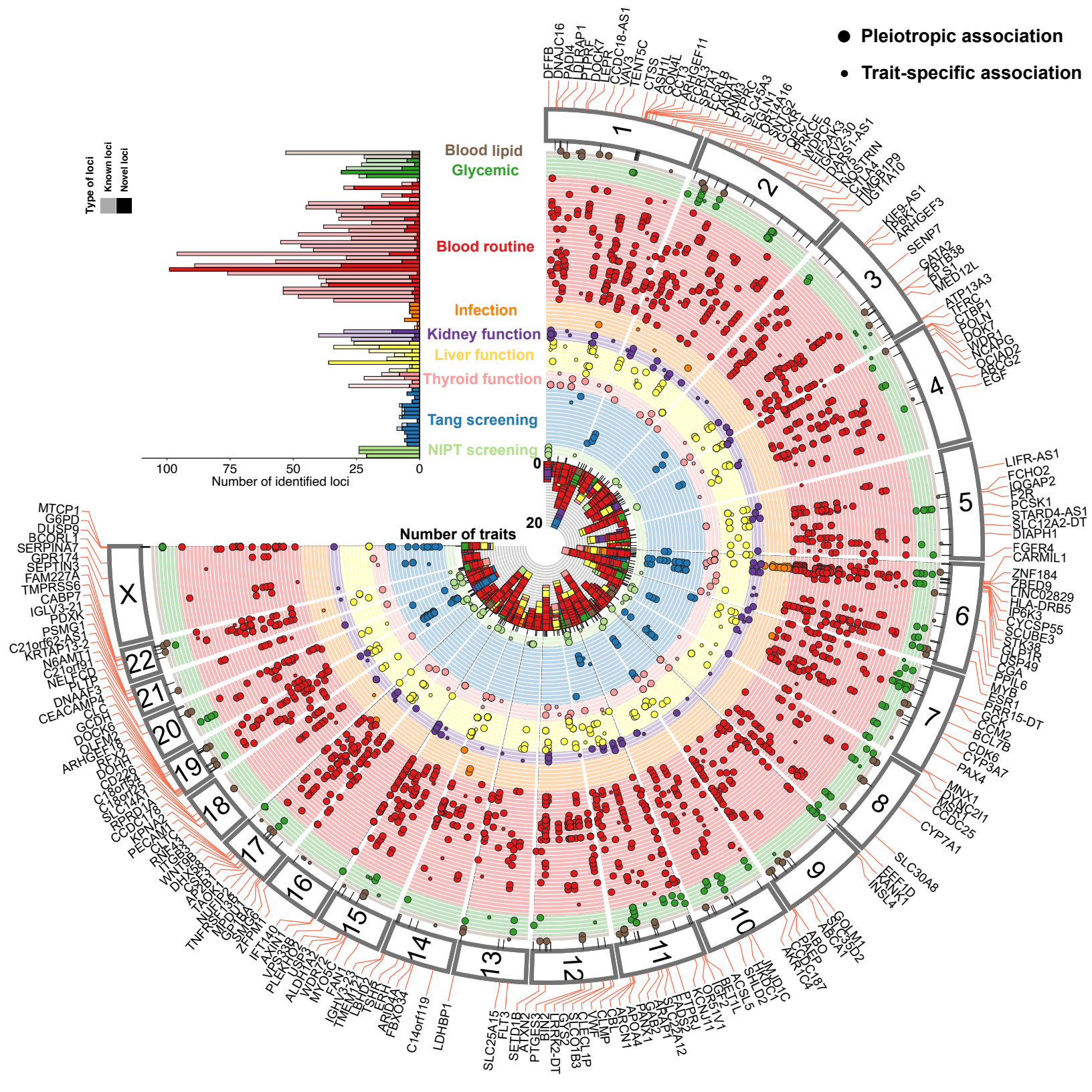
Panel **A** presents forest plots illustrating the genetic effects across five time periods, focusing on the first 25 variants with most significant G by T interaction effects and demonstrated substantial effect changes (non-overlapping genetic effect 95% confidence interval in at least two of the five periods). Each gestational period is represented by a different color, with beta values and 95% confidence intervals indicated. Annotations highlighting variant with a *P*-value less than 5×10^{-8}. The eQTL status of each variant is noted next to the variant label. Detailed information is availbale in Supplementary Table 31. Panel **B** depicts the changes in white blood cell (WBC) counts across the first, second, and third trimesters, at delivery, and during the postpartum period, using data from two hospitals. A generalized additive model was applied to smooth the curve, with the ribbon around the curve representing the 95% confidence interval. The first trimester spans from conception to 14 weeks (98 days), the second trimester from weeks 14 to 28 (196 days), and the third trimester extends from week 28 to delivery. Panel **C** displays a forest plot of the genetic effect of the rs2270401-A allele at 17q21.1 (eQTL for *GSDMA*) on WBC counts, based on a meta-analysis across the first, second, and third trimesters, at delivery, and postpartum between the two hospitals. Points represent the genetic effect (beta), with error bars indicating the 95% confidence interval. Panel **D** presents the results of enrichment analyses for variants exhibiting significant and non-significant G by T interactions.

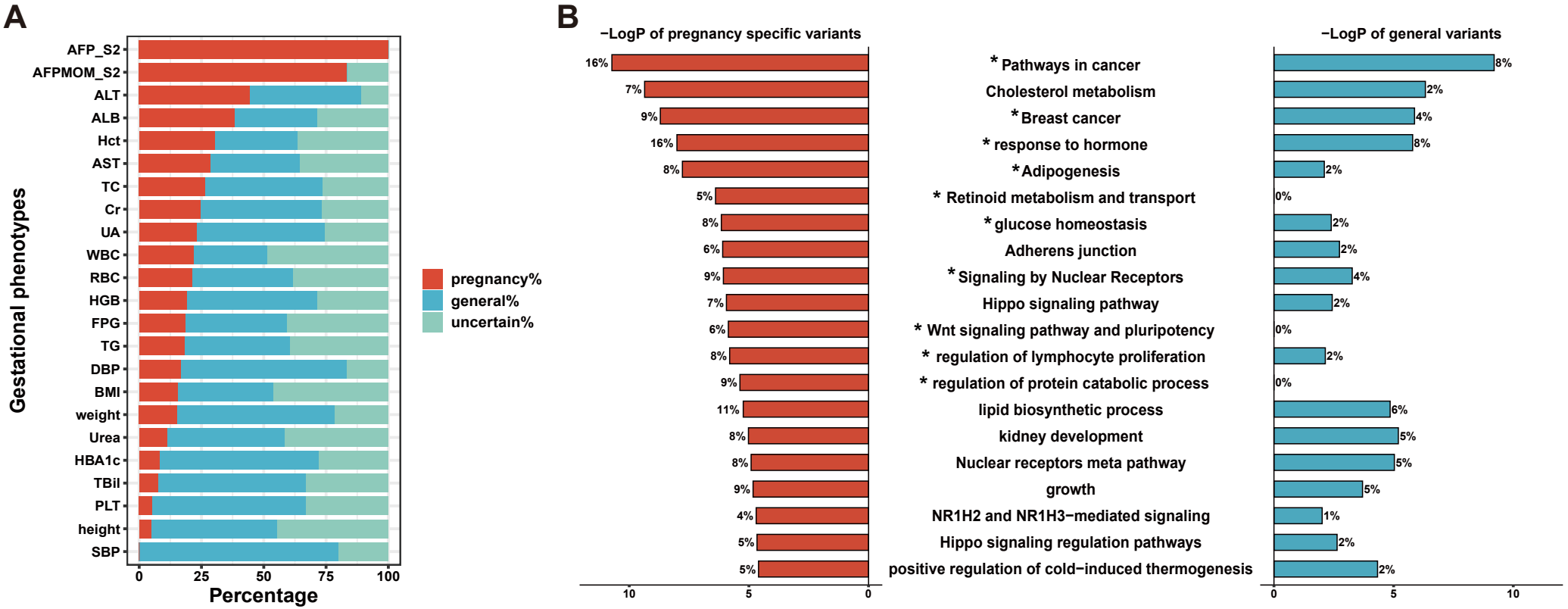
BASO: basophil absolute value; BASO_P: basophil percentage; HGB: hemoglobin level; MCV: mean corpuscular volume; MCH: mean corpuscular hemoglobin; RDW_CV: red blood cell volume distribution width variation; PLT: platelet count; MPV: mean platelet volume; PCT: platelet crit; PDW: platelet distribution width; P_LCR: platelet-large cell ratio.

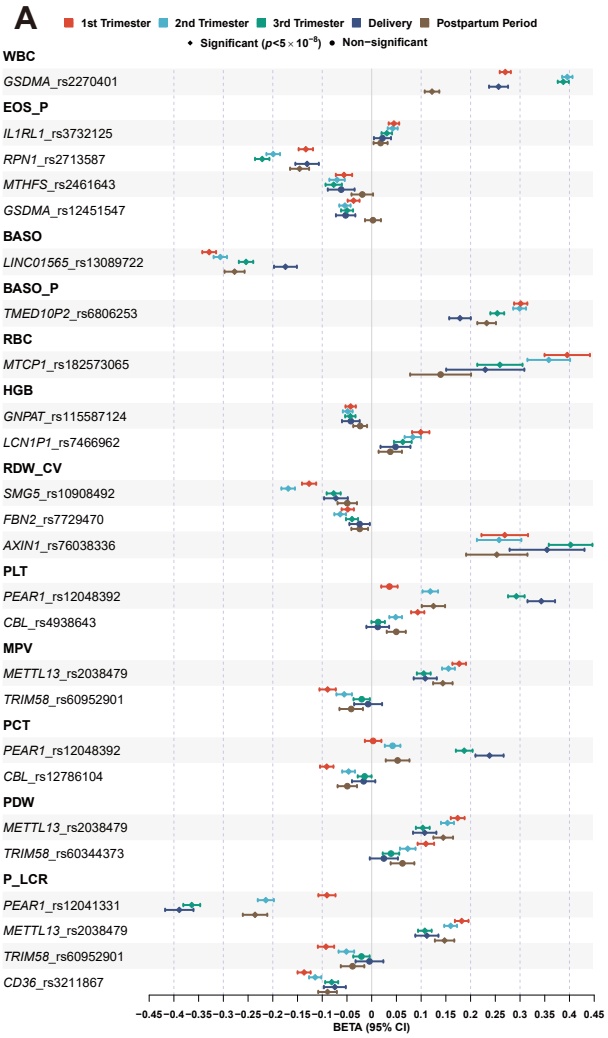
Figure 5. Causal inferences between gestational phenotypes and older-age phenotypes.

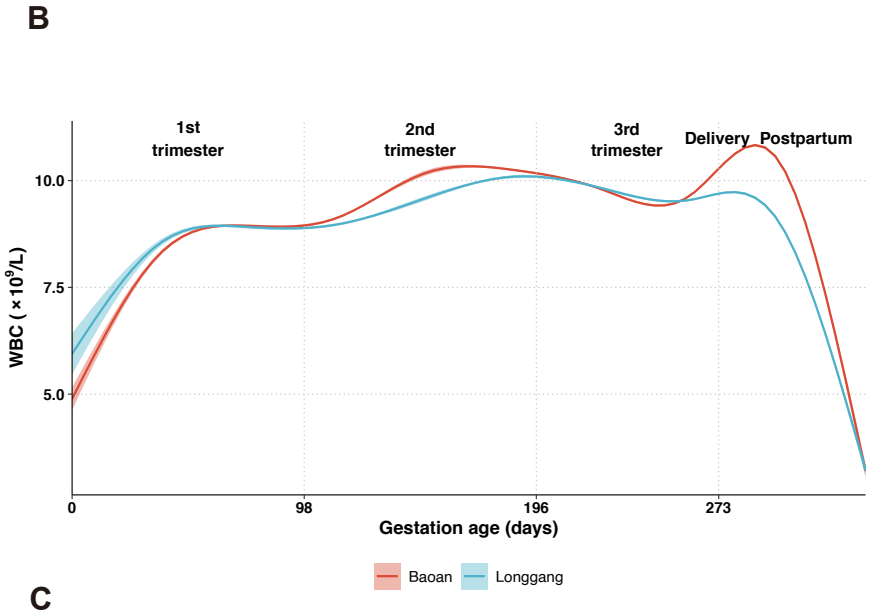
Panel **A** displays bidirectional Mendelian Randomization (MR) estimates, illustrating causal links between all gestational phenotypes (blue nodes) and older-age phenotypes from Biobank Japan (BBJ) cohort (red nodes). Arrows indicate the direction of association, with red denoting an increasing effect and blue representing a decreasing effect. Panel **B** features a heatmap of MR results, showing nine molecular test categories of gestational phenotypes as the exposures (right) and BBJ phenotypes as the outcomes (bottom). Associates with *P*-values below the FDR threshold are annotated with an asterisk (*, $q < 0.01$). Panel **C** presents an MR scatterplot where maternal triglyceride (TG) levels serve as the exposure and the use of antithrombotic agents in the BBJ cohort as the outcome. Panel **D** shows an MR scatterplot with maternal oral glucose tolerance test 1-hour plasma glucose (OGTT1H) levels as the exposure, and peripheral artery disease in the BBJ cohort as the outcome.

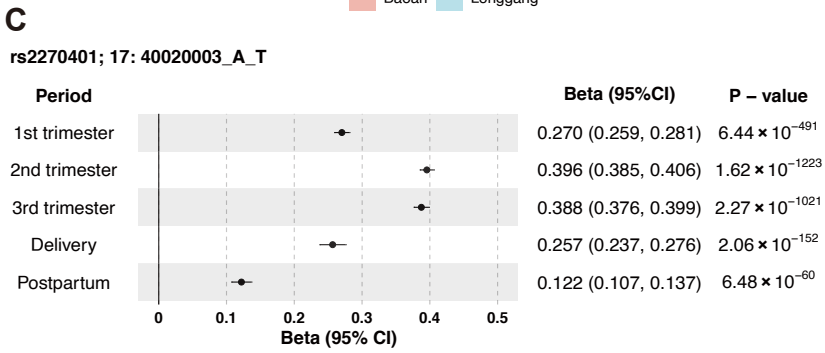

Study Cohort
Genotype
Genome sequencing from NIPT
N=121,579
Phenotypes
Blood lipid & Glycemic (N=8)
Blood routine (N=32)
Infection (N=14)
Liver, Kidney & Thyroid function (N=15)
Tang & NIPT screen (N=19)
Basic information (N=7)
Birth outcome (N=6)
Gestational disorders (N=21)
Genetic association
Baoan (N=50,963)
Longgang (N=70,616)
meta
5,027 independent signals in 3,845 loci
Consistency
Baoan vs Longgang
(N=121,579)
External replication
NIPT PLUS (N=5,733)
Baoan 20K (N=20,891)
Genome-wide association study
24 gestational phenotypes with multiple measurements in 5 periods
TrajGWAS
localized association
GWAS of longitudinal trajectories
Post-GWAS
Heritability & Correlation
LDSC
Up to 44% heritability
Genetic correlation
Baoan hospital
Different variant analysis
Genetic difference
Mendelian randomization
TwoSample MR
&
Multivariable MR
Future health
SNP effect on peripheral artery disease
SNP effect on OGTT1H

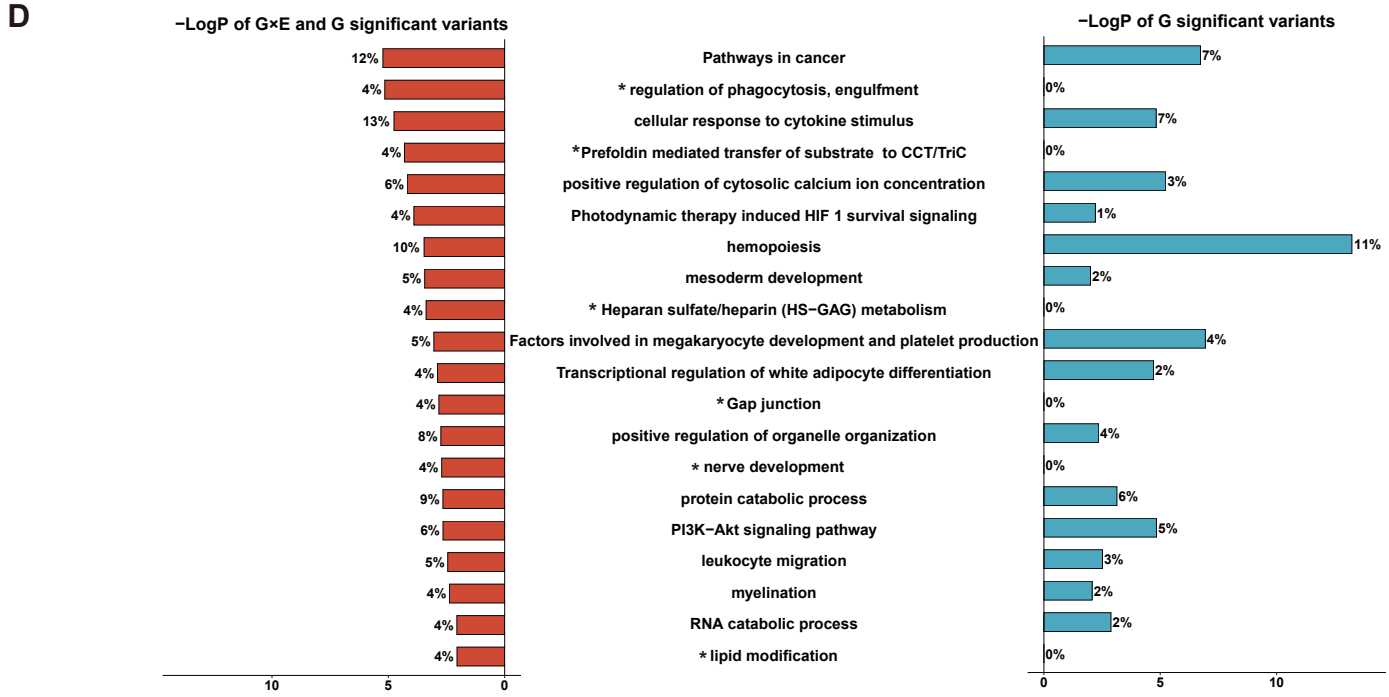

Pleiotropic association
Trait-specific association
Type of loci
Known loci
Novel loci
Blood lipid
Glycemic
Blood routine
Infection
Kidney function
Liver function
Thyroid function
Tang screening
NIPT screening
100
75
50
25
0
Number of identified loci
Number of traits
0
20
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X

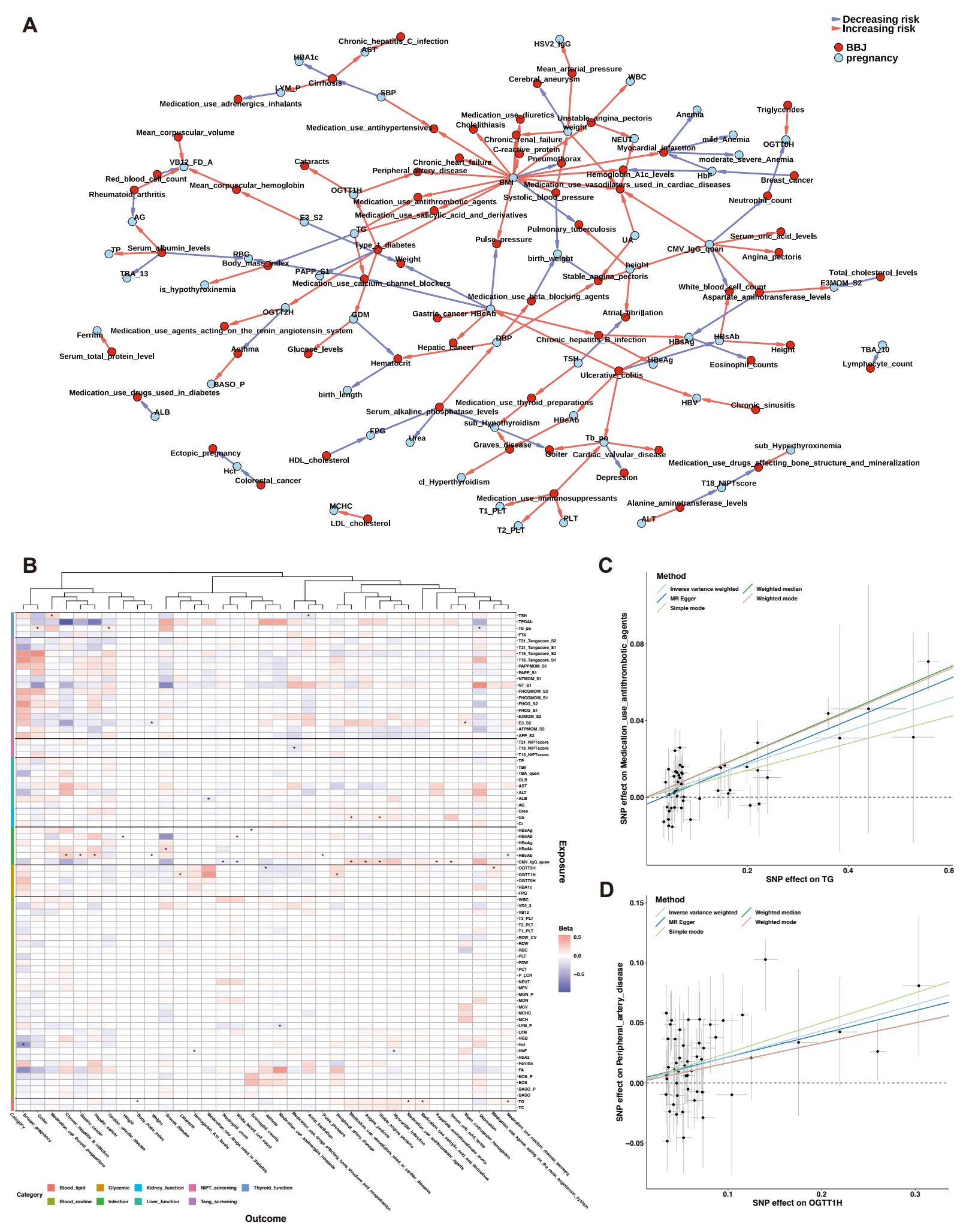

A
Gestational phenotypes
AFP_S2
AFPMOM_S2
ALT
ALB
Hct
AST
TC
Cr
UA
WBC
RBC
HGB
FPG
TG
DBP
BMI
weight
Urea
HBA1c
TBil
PLT
height
SBP
0
25
50
75
100
Percentage
pregnancy%
general%
uncertain%
B
−LogP of pregnancy specific variants
16%
7%
9%
16%
8%
5%
8%
6%
9%
7%
6%
8%
9%
11%
8%
8%
9%
4%
5%
5%
10
5
0
* Pathways in cancer
Cholesterol metabolism
* Breast cancer
* response to hormone
* Adipogenesis
* Retinoid metabolism and transport
* glucose homeostasis
Adherens junction
* Signaling by Nuclear Receptors
Hippo signaling pathway
* Wnt signaling pathway and pluripotency
* regulation of lymphocyte proliferation
* regulation of protein catabolic process
lipid biosynthetic process
kidney development
Nuclear receptors meta pathway
growth
NR1H2 and NR1H3−mediated signaling
Hippo signaling regulation pathways
positive regulation of cold−induced thermogenesis
−LogP of general variants
8%
2%
4%
8%
2%
0%
2%
2%
4%
2%
0%
2%
0%
6%
5%
5%
5%
1%
2%
2%
0
5
10


A
1st Trimester
2nd Trimester
3rd Trimester
Delivery
Postpartum Period
Significant (p<5×10^-8)
Non-significant
WBC
GSDMA_rs2270401
EOS_P
IL1RL1_rs3732125
RPN1_rs2713587
MTHFS_rs2461643
GSDMA_rs12451547
BASO
LINC01565_rs13089722
BASO_P
TMED10P2_rs6806253
RBC
MTCP1_rs182573065
HGB
GNPAT_rs115587124
LCN1P1_rs7466962
RDW_CV
SMG5_rs10908492
FBN2_rs7729470
AXIN1_rs76038336
PLT
PEAR1_rs12048392
CBL_rs4938643
MPV
METTL13_rs2038479
TRIM58_rs60952901
PCT
PEAR1_rs12048392
CBL_rs12786104
PDW
METTL13_rs2038479
TRIM58_rs60344373
P_LCR
PEAR1_rs12041331
METTL13_rs2038479
TRIM58_rs60952901
CD36_rs3211867
−0.45 −0.4 −0.35 −0.3 −0.25 −0.2 −0.15 −0.1 −0.05 0 0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45
BETA (95% CI)


B
1st trimester
2nd trimester
3rd trimester
Delivery
Postpartum
WBC (× 10^9/L)
10.0
7.5
5.0
0
98
196
273
Gestation age (days)
Baoan
Longgang


C
rs2270401; 17: 40020003_A_T
Period
Beta (95%CI)
P – value
1st trimester
0.270 (0.259, 0.281)
6.44 × 10^-491
2nd trimester
0.396 (0.385, 0.406)
1.62 × 10^-1223
3rd trimester
0.388 (0.376, 0.399)
2.27 × 10^-1021
Delivery
0.257 (0.237, 0.276)
2.06 × 10^-152
Postpartum
0.122 (0.107, 0.137)
6.48 × 10^-60
0 0.1 0.2 0.3 0.4 0.5
Beta (95% CI)


D
−LogP of G×E and G significant variants
−LogP of G significant variants
12% Pathways in cancer 7%
4% * regulation of phagocytosis, engulfment 0%
13% cellular response to cytokine stimulus 7%
4% *Prefoldin mediated transfer of substrate to CCT/TriC 0%
6% positive regulation of cytosolic calcium ion concentration 3%
4% Photodynamic therapy induced HIF 1 survival signaling 1%
10% hemopoiesis 11%
5% mesoderm development 2%
4% * Heparan sulfate/heparin (HS−GAG) metabolism 0%
5% Factors involved in megakaryocyte development and platelet production 4%
4% Transcriptional regulation of white adipocyte differentiation 2%
4% * Gap junction 0%
8% positive regulation of organelle organization 4%
4% * nerve development 0%
9% protein catabolic process 6%
6% PI3K−Akt signaling pathway 5%
5% leukocyte migration 3%
4% myelination 2%
4% RNA catabolic process 2%
4% * lipid modification 0%
10 5 0
0 5 10


A
Decreasing risk
Increasing risk
BBJ
pregnancy
Chronic_hepatitis_C_infection
AST
HBA1c
Cirrhosis
LYM_P
SBP
Medication_use_adrenergics_inhalants
HSV2_IgG
Mean_arterial_pressure
Cerebral_aneurysm
WBC
Medication_use_diuretics
Unstable_angina_pectoris
Cholelithiasis
Medication_use_antihypertensives
Anemia
Triglycerides
Mean_corpuscular_volume
Chronic_renal_failure
weight
NEUT
mild_Anemia
OGTT0H
C-reactive_protein
Myocardial_infarction
moderate_severe_Anemia
VB12_FD_A
Cataracts
Pneumothorax
Chronic_heart_failure
Peripheral_artery_disease
Hemoglobin_A1c_levels
HbF
Breast_cancer
Red_blood_cell_count
Mean_corpuscular_hemoglobin
BMI
Medication_use_vasodilators_used_in_cardiac_diseases
Rheumatoid_arthritis
OGTT1H
Medication_use_antithrombotic_agents
Systolic_blood_pressure
Neutrophil_count
AG
F3_S2
Medication_use_salicylic_acid_and_derivatives
TG
Pulmonary_tuberculosis
Serum_uric_acid_levels
TP
Serum_albumin_levels
Type_2_diabetes
Pulse_pressure
UA
CMV_IgG_quan
RBC
Weight
birth_weight
Angina_pectoris
Body_mass_index
height
TBA_13
PAPP_S1
Stable_angina_pectoris
Total_cholesterol_levels
is_hypothyroxinemia
Medication_use_calcium_channel_blockers
White_blood_cell_count
E3MOM_S2
Medication_use_beta_blocking_agents
Aspartate_aminotransferase_levels
OGTT2H
GDM
Gastric_cancer
HBcAb
Atrial_fibrillation
Ferritin
Medication_use_agents_acting_on_the_renin_angiotensin_system
DBP
Chronic_hepatitis_B_infection
HBsAg
HBsAb
Height
Serum_total_protein_level
Asthma
Glucose_levels
Hepatic_cancer
TSH
HBeAg
Eosinophil_counts
TBA_10
Hematocrit
Lymphocyte_count
BASO_P
Ulcerative_colitis
Medication_use_drugs_used_in_diabetes
birth_length
Medication_use_thyroid_preparations
HBV
Chronic_sinusitis
ALB
Serum_alkaline_phosphatase_levels
HBeAb
sub_Hypothyroidism
FPG
Urea
Tb_po
Graves_disease
Goiter
Cardiac_valvular_disease
sub_Hyperthyroxinemia
Ectopic_pregnancy
HDL_cholesterol
Medication_use_drugs_affecting_bone_structure_and_mineralization
Hct
Colorectal_cancer
cl_Hyperthyroidism
Depression
T18_NIPTscore
Medication_use_immunosuppressants
Alanine_aminotransferase_levels
MCHC
T1_PLT
PLT
ALT
LDL_cholesterol
T2_PLT
B
Exposure
Beta
0.5
0.0
−0.5
Category
Blood_lipid
Glycemic
Kidney_function
NIPT_screening
Thyroid_function
Blood_routine
Infection
Liver_function
Tang_screening
Outcome
C
Method
Inverse variance weighted
Weighted median
MR Egger
Weighted mode
Simple mode
SNP effect on Medication_use_antithrombotic_agents
SNP effect on TG
0.08
0.04
0.00
0.2
0.4
0.6
D
Method
Inverse variance weighted
Weighted median
MR Egger
Weighted mode
Simple mode
SNP effect on Peripheral_artery_disease
SNP effect on OGTT1H
0.15
0.10
0.05
0.00
−0.05
0.1
0.2
0.3