BMJ Open Diabetes Research & Care

Genetic variants associated with beta-cell function and insulin sensitivity potentially influence bile acid metabolites and gestational diabetes mellitus in a Chinese population

Qiulun Zhou,[1] Ying Wang,[2] Yuqin Gu,[1] Jing Li,[3] Hui Wang,[3] Junhong Leng ,[4] Weiqin Li,[4] Zhijie Yu,[5] Gang Hu ,[6] Ronald Ching Wan Ma ,[7,8] Zhong-Ze Fang ,[9] Xilin Yang ,[3] Guozhi Jiang [1]

To cite: Zhou Q, Wang Y, Gu Y, *et al*. Genetic variants associated with beta-cell function and insulin sensitivity potentially influence bile acid metabolites and gestational diabetes mellitus in a Chinese population. *BMJ Open Diab Res Care* 2021;**9**:e002287. doi:10.1136/bmjdrc-2021-002287

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjdrc-2021-002287).

QZ and YW contributed equally.

Received 22 March 2021
Accepted 17 August 2021

For numbered affiliations see end of article.

Correspondence to
Dr Zhong-Ze Fang; fangzhongze@tmu.edu.cn, Professor Xilin Yang; yangxilin@tmu.edu.cn and Professor Guozhi Jiang; jianggzh5@mail.sysu.edu.cn

ABSTRACT

Introduction To investigate associations between genetic variants related to beta-cell (BC) dysfunction or insulin resistance (IR) in type 2 diabetes (T2D) and bile acids (BAs), as well as the risk of gestational diabetes mellitus (GDM).

Research design and methods We organized a case-control study of 230 women with GDM and 217 without GDM nested in a large prospective cohort of 22 302 Chinese women in Tianjin, China. Two weighted genetic risk scores (GRSs), namely BC-GRS and IR-GRS, were established by combining 39 and 23 single nucleotide polymorphisms known to be associated with BC dysfunction and IR, respectively. Regression and mediation analyses were performed to evaluate the relationship of GRSs with BAs and GDM.

Results We found that the BC-GRS was inversely associated with taurodeoxycholic acid (TDCA) after adjustment for confounders (Beta (SE)=−0.177 (0.048); $p=2.66\times10^{-4}$). The BC-GRS was also associated with the risk of GDM (OR (95% CI): 1.40 (1.10 to 1.77); p=0.005), but not mediated by TDCA. Compared with individuals in the low tertile of BC-GRS, the OR for GDM was 2.25 (95% CI 1.26 to 4.01) in the high tertile. An interaction effect of IR-GRS with taurochenodeoxycholic acid (TCDCA) on the risk of GDM was evidenced (p=0.005). Women with high IR-GRS and low concentration of TCDCA had a markedly higher OR of 14.39 (95% CI 1.59 to 130.16; p=0.018), compared with those with low IR-GRS and high TCDCA.

Conclusions Genetic variants related to BC dysfunction and IR in T2D potentially influence BAs at early pregnancy and the development of GDM. The identification of both modifiable and non-modifiable risk factors may facilitate the identification of high-risk individuals to prevent GDM.

SIGNIFICANCE OF THIS STUDY

WHAT IS ALREADY KNOWN ABOUT THIS SUBJECT?

⇒ Gestational diabetes mellitus (GDM) and type 2 diabetes (T2D) may share risk factors.
⇒ Bile acids were reported to be associated with both GDM and T2D in different populations.

WHAT ARE THE NEW FINDINGS?

⇒ Beta-cell genetic risk score (BC-GRS) derived from 39 known risk variants for BC dysfunction in T2D were negatively associated with the concentration of taurodeoxycholic acid (TDCA), while the insulin resistance genetic risk score (IR-GRS) derived from 23 variants related to IR exhibited interaction effects with taurochenodeoxycholic acid (TCDCA).
⇒ The BC-GRS was associated with the risk of GDM, but not mediated by TDCA.
⇒ Women with a high tertile of IR-GRS and low concentration of TCDCA (≤0.2 nmol/mL) were at higher risk of GDM.

HOW MIGHT THESE RESULTS CHANGE THE FOCUS OF RESEARCH OR CLINICAL PRACTICE?

⇒ Our findings provide novel insights towards the underlying pathophysiology of GDM, highlighting some overlap with the pathogenesis of T2D.
⇒ This work highlights the importance of both genetic and modifiable risk factors, notably bile acids, which may facilitate the identification of high-risk individuals for optimal control of risk factors to prevent GDM.
⇒ Our findings provide novel insights on the pathogenesis of GDM, and the identification of both modifiable and non-modifiable risk factors may facilitate the identification of high-risk individuals to prevent GDM.

INTRODUCTION

Gestational diabetes mellitus (GDM), which is highly prevalent in the Asian population, has become one of the leading causes of mortality and morbidity for both mothers and children worldwide.[1] It is estimated that the pooled prevalence of GDM was 11.5% in Asians[2] and reached 14.8% in mainland China according to a recent review and meta-analysis.[3] Compared with women with a

normoglycemia during pregnancy, those with a history of GDM are prone to have a higher risk of developing type 2 diabetes (T2D) and cardiovascular diseases.[4–7] Given the continuous increase in the number of people with GDM, it is particularly important to identify those individuals at high risk of GDM for early intervention. However, since GDM is a complex multifactorial disease and the pathogenesis of the disease remains unclear, it is a key research priority to investigate the determinants of GDM.

Similar to the development of T2D, increased insulin resistance (IR) and defects in insulin secretion are the underlying pathophysiologies of GDM.[4] A growing body of literature has revealed that T2D and GDM share common risk factors including obesity, smoking, and genetic variants.[8 9] For example, Ding and colleagues[10] investigated the associations of 112 single nucleotide polymorphisms (SNPs) confirmed by genome-wide association studies (GWASs) for T2D with GDM and identified 11 variants significantly associated with GDM in White populations. In view of the fact that different genetic loci often imply different pathophysiological mechanisms, it is important to gain insight into the relationship between T2D-related variants and GDM according to their physiologic functions.

Bile acids (BAs) are cholesterol catabolites, which was known to affect the metabolism of glycemia, lipid, and energy.[11 12] Serum BA has been identified as a potentially modifiable risk factor for T2D.[13 14] Cariou *et al*[15] reported that the concentrations of serum BAs, including cholic acids (CA), chenodeoxycholic acid (CDCA), and deoxycholic acid (DCA), were negatively associated with insulin sensitivity in individuals with T2D. Multiple studies found that serum BAs distributed differently among women with and without intrahepatic cholestasis of pregnancy, which could increase the risk of GDM.[16–18] Moreover, a nested case-control study showed that a higher level of total BA in the first-trimester could contribute towards the development of GDM.[19] In addition, recent GWASs identified a number of genetic variants associated with the concentration of BAs[20] and indicated that the metabolism of BAs shares some common genetic origin with T2D.[21] However, it remains unclear whether T2D-related genetic variants linked the relationship between serum BAs and the development of GDM.

In this study, we hypothesized that genetic variants related to beta-cell (BC) dysfunction and IR in T2D may alter the concentration of serum BAs or interact with BA metabolites to influence the development of GDM. In a case-control study nested in a prospective cohort with documentation of clinical, genetic, and metabolite profiles, we examined the relationship between BC-related or IR-related genetic variants and BA species, as well as the risk of GDM in Chinese pregnant women. A better understanding of the genetic basis of GDM and the potential role of BAs in the development of GDM may enable physicians to identify high-risk patients for early intervention.

METHODS

Research design and population

The study design, recruitment methods, and biochemical assays were described in detail previously.[22 23] In brief, a population-based cohort of pregnant women was established in the six central urban districts of Tianjin, China, from October 2010 to August 2012. A total of 22 302 Chinese pregnant women were enrolled in this prospective cohort at their first antenatal care through a universal screening and management system for GDM. Once enrolled, participants were followed longitudinally from their first antenatal care visit to the time of glucose challenge test (GCT) at 24–28 gestational weeks and through the postpartum period. Written informed consent was obtained from all participants at the time of enrollment for archival and research purposes.

Among the 22 302 participants, 2991 provided overnight fasting blood samples at their first antenatal care visit (median (IQR): 10 (9–11) gestational weeks). We excluded 227 women without GCT results or oral glucose tolerance test (OGTT) results if their GCT ≥7.8 mmol/L. Of the remaining 2764 women, 243 developed GDM (see definition below) and were used as cases in this study, and 243 without GDM matched on age (±1 year) and gestational weeks (±2 weeks) were used as controls for this age-matched nested case-control study. The comparison of the 2764 participants with the rest of the entire cohort was previously described.[22]

Furthermore, after removing 8 subjects with a low capacity of DNA extraction, genome-wide genotyping was performed for 478 subjects using the Illumina Infinium Global Screening Array, and genotype data were imputed using minimac 3 with the 1000 Genomes Project phase 3 V.5 as a reference panel. After standard quality control (QC) according to the procedures illustrated by Anderson *et al*,[24] 23 subjects were discarded due to gender problem, call rate <97%, extreme heterozygosity or singleton, or DNA sample contamination. Finally, 447 subjects (230 GDM women and 217 non-GDM women) with detailed information on risk factors and outcomes were included in this study (figure 1).

Data collection procedures

Details of clinical assessment methods and definitions of clinical outcomes have been described.[22 23] Briefly, standardized procedures were used to measure maternal height, weight, and blood pressure (BP) at the first antenatal care visit. Weight was measured to the nearest 0.1 kg and was remeasured at the time of GCT. Body weight at the first antenatal care visit was used as the prepregnancy body weight to estimate prepregnancy body mass index (BMI). The weight difference between the first antenatal visit and GCT time was recorded as gestational weight gain. Other data were collected through a series of structured questionnaires filled out by nursing staffs and/or pregnant women at their first antenatal care visit, the time of the GCT, and subsequent antenatal care visits, respectively. We retrieved

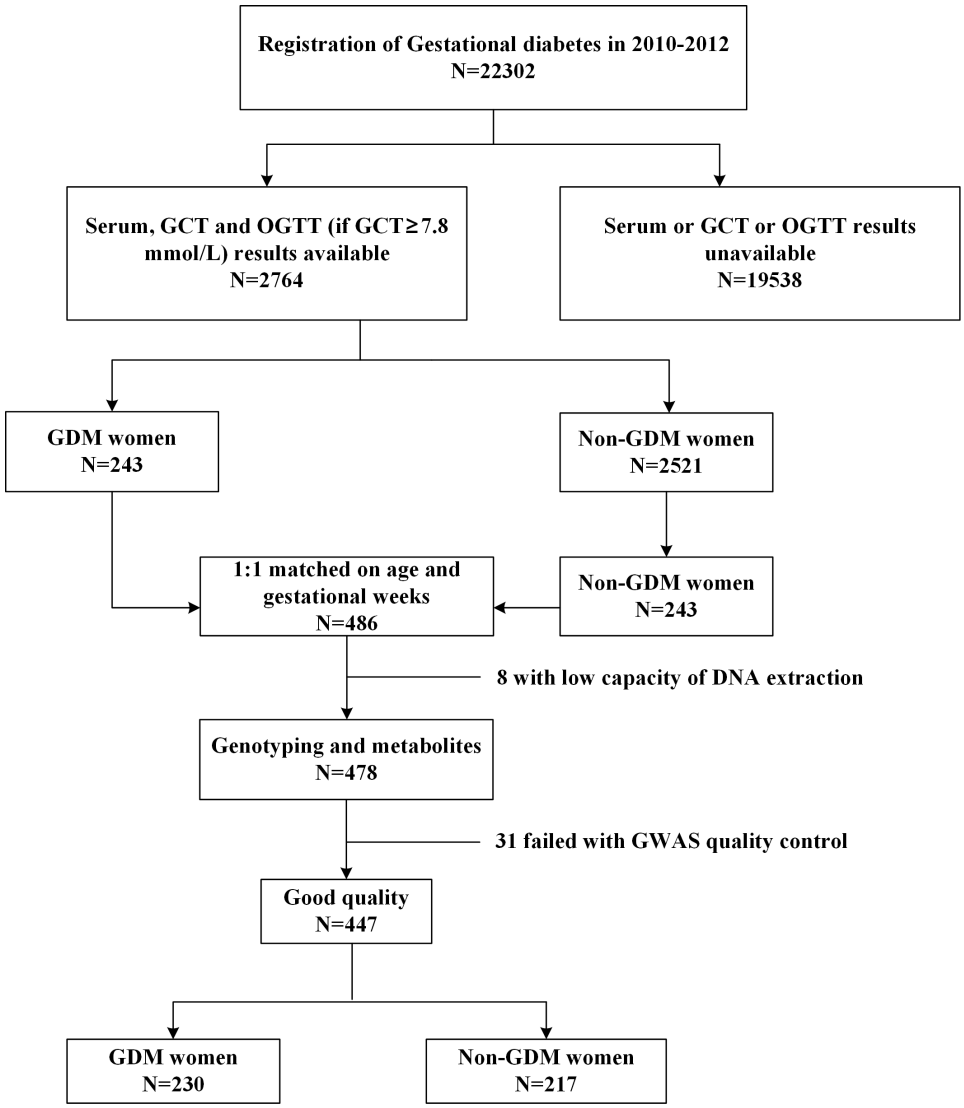


Figure 1 Flow diagram of sample selection in the cohort. GCT, glucose challenge test; GDM, gestational diabetes mellitus; GWAS, genome-wide association study; OGTT, oral glucose tolerance test.

pregnancy outcomes and other information from the centralized computer database of Tianjin Maternal and Child Health Information System, including the data of maternal age, family history of diabetes in first-degree relatives, parity, race, education, smoking habits before or during pregnancy, and alcohol consumption before or during pregnancy.

Definition of clinical outcomes

In this study, we used a two-step screening procedure to identify GDM cases. First, all pregnant women underwent a 50-g 1-hour GCT in non-fasting status at 24–28 weeks of gestation at primary care hospitals. Those women with plasma glucose ≥7.8 mmol/L were referred to the GDM clinic within Tianjin Women and Children's Health Center and then took a 75-g 2-hour OGTT in the morning after at least 8 hours of overnight fasting. All women with GDM fulfilled the 2013 WHO diagnostic criteria for GDM.[25]

Measurement of serum bile acids

LC-MS/MS analysis

Blood samples were collected at the first antenatal care visit (median at 10th gestational weeks). BAs were quantified using an liquid chromatography-mass spectrometry (LC-MS) based targeted metabolomics approach, and details of sample pretreatment and liquid chromatography-tandem mass spectrometry analysis were described in previous studies.[23 26] In brief, each blood sample was separated from the venous blood immediately and stored at –80°C and thawed in 4°C when used. QC samples were prepared by mixing all of the samples. After sample pretreatment, an Eksigentultral liquid chromatography 100 coupled with an AB 5600 TripleTOF system (AB SCIEX) was used to identify and quantify the BAs components. To separate the different BA components, a 2.1×100 mm XBridge Peptide BEH C18 column (waters) with a 4×2.0 mm guard column (phenomenex) was equipped. Under a column temperature of 40°C,

a controlled gradient of mobile phase A, composed of 0.1% (v/v) formic acid and 10 mM acetic acid amine in water, and mobile phase B, which was comprised by 0.1% formic acid and 80% (v/v) methanol and 20% (v/v) acetonitrile, was used for separation at a flow rate of 0.4 mL/min. The injection volume of the sample was 5 µL. During the analysis of the sample sequence, one QC sample was run after every 30 injections.

Data processing

On the basis of the m/z value and sample retention time, the Peak view V.1.2 and Multi-Quantum V.2.1 software were used to acquire the raw data.

Genetic risk score

We created two genetic risk scores (GRSs) based on 52 and 30 independent genetic variants known to be associated with BC function and IR from previous published GWAS, respectively.[27 28] We extracted these SNPs from genome-wide genotyping data by the Illumina Infinium Global Screening Array and explored the associations between each GRS and BAs, as well as GDM in this case-control study. Standard QC (minor allele frequency (MAF) >0.01; call rate >97%; $p>1\times10^{-4}$ in Hardy-Weinberg equilibrium) was performed and genotype data were imputed using minimac 3 with the 1000 Genomes Project phase 3 V.5 as a reference panel. Independent common SNPs (linkage disequilibrium coefficient r^2 <0.5; MAF >0.01) available in our dataset with good imputation quality (r^2 >0.8) were selected to construct the GRSs. A proxy SNP with r^2 >0.6 (according to the 1000 Genome CHB panel) was selected when the index SNP was imputed with poor quality. Finally, we obtained 39 and 23 independent common SNPs to develop the BC-GRS and IR-GRS, respectively (online supplemental tables 1 and 2).

The weighted GRS was developed by summing the score of reported risk allele for each SNP based on an additive genetic model, weighted by the effect size of the BC- or IR-related SNP as reported in the literature, then rescaled to a score to express the SD using the following formula:

$$\frac{\text{individual GRS value}-\text{population mean GRS}}{\text{population SD of GRS.}}.$$

The GRSs were standardized according to population means and SDs in the entire cohort. The final analysis is based on (1) the genetic risk per SD of the standardized GRSs and (2) tertile analysis of the GRSs for BAs and GDM (tertile 1: low; tertile 2: intermediate, and tertile 3: high).

Statistical analysis

In this study, we used logistic regression to explore the associations between GRSs and GDM outcome and employed linear regression to assess the associations between GRSs and serum BAs. Because of the skewed distribution, log-transformed BA was used as the dependent variable to fit the linear models. Mediation analysis was performed to examine the mediator effects of BA on the relationship between GRS and GDM using structural equation modeling.[29] Interaction effects between GRS and BA on the risk of GDM were also examined by adding a product term to the model. We adjusted for age, gestational weeks, baseline BMI, family history of diabetes, drinking history, smoking history, education attainment, ethnicity group, weight gain during pregnancy, alanine aminotransferase (ALT), systolic blood pressure (SBP), diastolic blood pressure (DBP), and parity to reduce the confounding effects induced by these variables. All data were expressed as percentages, means and SDs, or medians and IQRs as appropriate. Two-tailed p<0.05 was considered to indicate statistical significance for the comparison of baseline variables. Since a total of 11 BAs and 2 GRSs were included in the study, a Bonferroni-corrected significance level of p<0.0023 (0.05/22) was used for the association analysis of GRS with BAs. Analyses were performed using R (V.4.0.2; http://www.R-project.org).

RESULTS

Cohort description

The mean age of the cohort was 29.2±3.02 (SD) years, and the gestational age was 10.1±2.07 (SD) weeks at their first antenatal care visit. Compared with women without GDM, those who developed GDM had higher BMI, BP, alanine, and aminotransferase (ALT) when being tested at their first visit. No significant difference was observed for age, gestational weeks, ethnicity, parity, family history of diabetes, smoking and drinking habitats, and education levels between the two groups. A total of 11 BAs, namely glycocholic acid (GCA), glycodeoxycholic acid (GDCA), glycochenodeoxycholic acid (GCDCA), taurocholic acid (TCA), taurochenodeoxycholic acid (TCDCA), taurodeoxycholic acid (TDCA), hyodeoxycholic acid (HDCA), CA, glycoursodeoxycholic acid (GUDCA), CDCA, and DCA, were detectable in >90% of the serum samples and were used in this analysis (table 1), and others were listed in online supplemental table 3. Although CA and TCA were comparable between GDM and non-GDM groups, CDCA, DCA, GCA, GCDCA, GDCA, GUDCA, HDCA, TCDCA, and TDCA were lower in GDM group compared with non-GDM group (table 1).

Association of GRS with BA

We examined the associations of GRSs and 11 BAs with adjustment for clinical confounders. Five BA species were found to be potentially associated with GRSs derived from BC-related genetic variants in the entire cohort. The BC-GRS was negatively associated with log-transformed TDCA at the Bonferroni-corrected significance level (beta (SE)=−0.176 (0.048) per SD; $p=2.66\times10^{-4}$). Other BA species, including GDCA (beta (SE)=−0.140 (0.053) per SD; p=0.009), TCDCA (beta (SE)=−0.108 (0.046) per SD; p=0.020), GCA (beta (SE)=−0.113 (0.052) per SD; p=0.029), and TCA (beta (SE)=−0.072 (0.035) per SD;

Table 1 Clinical and biochemical characteristics of participants

Characteristics	Non-GDM	GDM	P
N	217	230	
Age (years)	29.1±3.32	29.2±2.72	0.808
BMI (kg/m^2)	21.9±3.45	23.9±3.64	<0.001
DBP (mm Hg)	67.8±7.58	70.6±7.96	<0.001
SBP (mm Hg)	104±10.5	108±10.6	<0.001
ALT (U/L)	16 (11,22)	19 (14,26)	<0.001
GCT glucose (mmol/L)	6.27 (5.39, 7.20)	9.00 (8.36, 9.96)	<0.001
Gestational weeks	10.1±2.05	10.1±2.10	0.937
Han nationality	210 (96.8%)	225 (97.8%)	0.693
Parity ≥1	11 (5.1%)	13 (5.7%)	0.949
Family history of diabetes	13 (6.0%)	28 (12.2%)	0.168
Habitual smoker (yes)	13 (6.0%)	15 (6.5%)	0.867
Alcohol drinker (yes)	53 (24.4%)	69 (30.0%)	0.999
Education>12 years	216 (48.9%)	226 (51.1%)	0.080
Weight gain up to GCT (kg)	8.72±3.28	8.32±3.60	0.229
Bile acid (nmol/mL)			
DCA	0.26 (0.16, 0.45)	0.20 (0.10, 0.34)	<0.001
GCA	0.07 (0.04, 0.13)	0.05 (0.03, 0.09)	<0.001
GDCA	0.12 (0.06, 0.23)	0.08 (0.04, 0.14)	<0.001
GCDCA	0.35 (0.16, 0.63)	0.21 (0.12, 0.38)	<0.001
GUDCA	0.03 (0.02, 0.05)	0.02 (0.01, 0.03)	<0.001
TCDCA	0.10 (0.05, 0.19)	0.06 (0.04, 0.10)	<0.001
TDCA	0.04 (0.02, 0.07)	0.03 (0.02, 0.05)	<0.001
CDCA	0.09 (0.05, 0.20)	0.08 (0.04, 0.14)	0.024
HDCA	0.03 (0.02, 0.04)	0.02 (0.01, 0.04)	0.040
TCA	0.05 (0.04, 0.09)	0.06 (0.05, 0.08)	0.056
CA	0.10 (0.08, 0.15)	0.10 (0.09, 0.13)	0.153

Non-GDM group was matched on age ±1 year of the GDM group. Values are described as n, n (%), or median (Q1, Q3) or means±SD.
ALT, alanine aminotransferase; BMI, body mass index; CA, cholic acid; CDCA, chenodeoxycholic acid; DBP, diastolic blood pressure; DCA, deoxycholic acid; GCA, glycocholic acid; GCDCA, glycochenodeoxycholic acid; GCT, glucose challenge test; GDCA, glycodeoxycholic acid; GDM, gestational diabetes mellitus; GUDCA, glycoursodeoxycholic acid; HDCA, hyodeoxycholic acid; SBP, systolic blood pressure; TCA, taurocholic acid; TCDCA, taurochenodeoxycholic acid; TDCA, taurodeoxycholic acid.

p=0.039), showed suggestive associations with BC-GRS, but not significant after Bonferroni correction for multiple comparisons (figure 2 and online supplemental table 4). The IR-GRS was not significantly associated with any BA species after Bonferroni correction.

Association of GRS with GDM

On multivariate logistic regression with adjustment for clinical confounders, the weighted BC-GRS was significantly associated with a higher risk of GDM (OR (95% CI): 1.40 (1.10 to 1.77) per SD; p=0.005). The top tertile of the BC-GRS showed significant association with GDM compared with the reference group with an OR of 2.25 (95% CI 1.26 to 4.01). Moreover, the associations for the risk of GDM persisted and remained significant when further adjusted for TDCA which were significantly associated with the BC-GRS (table 2). No significant associations were found between the IR-GRS and GDM with or without adjustment for BAs (table 2).

Mediation effect of TDCA on the relationship between BC-GRS and GDM

With consideration of the significant associations of the BC-GRS with both TDCA and GDM, we examined the potential mediation effect of TDCA on the relationship between the BC-GRS and GDM (online supplemental figure 1). In the mediation analysis using structural equation modeling, there was a significant association between the BC-GRS and TDCA (beta (SE)=−0.010 (0.005), p=0.032), as well as the association between TDCA and GDM (beta (SE)=−0.754 (0.300), p=0.012). Meanwhile, the BC-GRS was directly associated with GDM

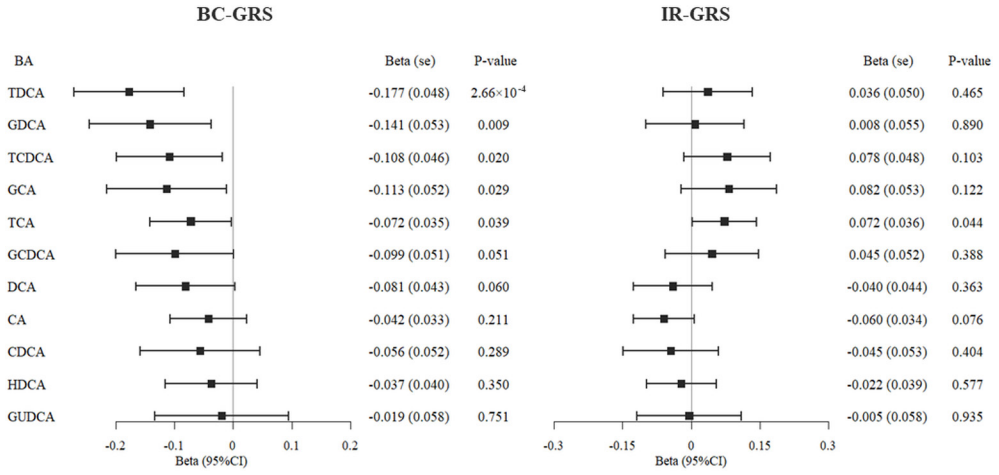


Figure 2 Associations of BC-GRS and IR-GRS with bile acids. Coefficients (beta) were adjusted for clinical risk factors including age, gestational weeks, baseline BMI, family history of diabetes, drinking history, smoking history, education attainment, ethnicity group, weight gain during pregnancy, ALT, SBP, DBP, and parity. ALT, alanine aminotransferase; BA, bile acid; BC-GRS, beta-cell genetic risk score; BMI, body mass index; CA, cholic acid; CDCA, chenodeoxycholic acid; DBP, diastolic blood pressure; DCA, deoxycholic acid; GCA, glycocholic acid; GCDCA, glycochenodeoxycholic acid; GDCA, glycodeoxycholic acid; GUDCA, glycoursodeoxycholic acid; HDCA, hyodeoxycholic acid; IR-GRS, insulin resistance genetic risk score; SBP, systolic blood pressure; TCA, taurocholic acid; TCDCA, taurochenodeoxycholic acid; TDCA, taurodeoxycholic acid.

after adjustment for TDCA (beta (SE)=0.074 (0.026), p=0.004). However, the indirect effect of the BC-GRS on GDM through TDCA was not statistically significant (beta (SE)=0.008 (0.005), p=0.103).

Interaction between GRS and BA on risk of GDM

We examined the interaction effects between GRSs and BA species on the risk of GDM. By adding the cross-product term to regression models, we observed a significant interaction of IR-GRS with TCDCA on risk of GDM (p=0.005). As there was a non-linear relationship between TCDCA and GDM (p=0.008 for non-linearity) and TCDCA≤0.2 nmol/mL was observed to be significantly associated with increased risk of GDM from our previous study,[23] we stratified the TCDCA into categorical variables at the specific cut-off point of 0.2 nmol/mL and examined interaction effects between

Table 2 Associations of BC-GRS and IR-GRS with GDM

	Model 1		Model 2		Model 3	
	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
BC-GRS (#SNP=39)						
GRS categorized as continuous	1.28 (1.06 to 1.54)	**0.011**	1.40 (1.10 to 1.77)	**0.005**	1.37 (1.07 to 1.75)	**0.013**
GRS categorized as tertiles						
Tertile 1	1	/	1	/	1	/
Tertile 2	1.13 (0.72 to 1.77)	0.607	1.17 (0.67 to 2.05)	0.575	1.24 (0.69 to 2.23)	0.464
Tertile 3	1.65 (1.05 to 2.60)	**0.030**	2.25 (1.26 to 4.01)	**0.005**	2.12 (1.14 to 3.93)	**0.017**
IR-GRS (#SNP=23)						
GRS categorized as continuous	1.18 (0.98 to 1.42)	0.080	1.24 (0.98 to 1.56)	0.069	1.24 (0.97 to 1.58)	0.079
GRS categorized as tertiles						
Tertile 1	1	/	1	/	1	/
Tertile 2	1.07 (0.68 to 1.68)	0.772	1.01 (0.58 to 1.75)	0.982	0.96 (0.53 to 1.74)	0.902
Tertile 3	1.43 (0.91 to 2.25)	0.122	1.41 (0.81 to 2.45)	0.221	1.44 (0.81 to 2.56)	0.219

Model 1 was the model without adjustment. Model 2 was adjusted for traditional confounders including age, gestational weeks, baseline BMI, family history of diabetes, drinking history, smoking history, education attainment, ethnicity group, weight gain during pregnancy, ALT, SBP, DBP, and parity. Model 3 was adjusted for traditional confounders in addition to TDCA. Bold values indicate the P values were statistically significant (P<0.05).
ALT, alanine aminotransferase; BC-GRS, beta-cell genetic risk score; BMI, body mass index; DBP, diastolic blood pressure; GDM, gestational diabetes mellitus; GRS, genetic risk score; IR-GRS, insulin resistance genetic risk score; SBP, systolic blood pressure; SNP, single nucleotide polymorphism; TDCA, taurodeoxycholic acid.

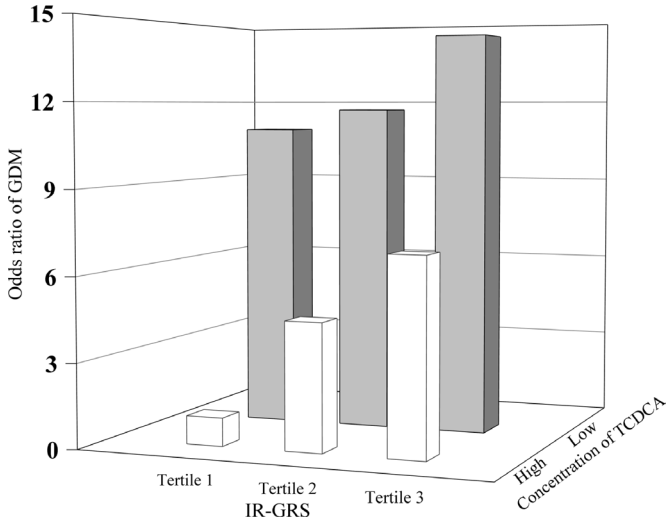


Figure 3 Interaction between IR-GRS and TCDCA. The group with low IR-GRS (Tertile 1) and high concentration of TCDCA (>0.2 nmol/L) was used as reference. ORs were adjusted for clinical risk factors including age, gestational weeks, baseline BMI, family history of diabetes, drinking history, smoking history, education attainment, ethnicity group, weight gain during pregnancy, ALT, SBP, DBP, and parity. ALT, alanine aminotransferase; BMI, body mass index; DBP, diastolic blood pressure; GDM, gestational diabetes mellitus; IR-GRS, insulin resistance genetic risk score; SBP, systolic blood pressure; TCDCA, taurochenodeoxycholic acid.

the categorized IR-GRS and TCDCA. The risk of developing GDM increased with an increasing number of risk alleles grouped by tertiles (p=0.025 for a trend in groups with a high concentration of TCDCA). Compared with women with low genetic risk (tertile 1) for GDM and high concentration of TCDCA (the reference group), women with high genetic risk (tertile 3) and low concentration of TCDCA had an OR of 14.39 (95% CI 1.59 to 130.16, p=0.018) (figure 3 and online supplemental table 5). No significant interaction effects were detected between the BC-GRS and BAs.

DISCUSSION

GDM, one of the most common complications of pregnancy, is closely linked to T2D. In view of the fact that increased IR and impaired insulin secretion are the main pathophysiological features of T2D, we constructed the BC-GRS and IR-GRS using genetic variants associated with BC dysfunction and IR in T2D, respectively, and analyzed the relationship between the GRSs and BA metabolites, as well as GDM in Chinese women. We revealed that the BC-GRS derived from 39 known risk variants for BC dysfunction in T2D were negatively associated with the concentration of TDCA, while the IR-GRS derived from 23 variants related to IR exhibited interaction effects with TCDCA. Furthermore, we found that the BC-GRS was also associated with GDM, but the effect was not mediated by TDCA. Compared with the IR-GRS, the magnitude of the association of BC-GRS with GDM was stronger, indicating that T2D SNPs related to defects in insulin secretion play a central role in the development of GDM in Chinese. These findings highlighted the importance of both genetic and modifiable risk factors, notably BAs, which may facilitate the identification of high-risk individuals for optimal control of risk factors to prevent GDM.

With respect to the relationship between serum BA and hyperglycemia, research in this area remains limited and conclusions are inconsistent. For example, Hou and colleagues[19] highlighted that the concentration of total serum BA was significantly higher in women who developed GDM when compared with healthy pregnant women. In the Joslin Diabetes Study, investigators proposed that patients with T2D had higher concentrations of fasting taurine-conjugated BA compared with normal glucose-tolerant persons.[30] On the contrary, Dudzik and colleagues[31] found that taurine-conjugated BAs were negatively associated with GDM in the European population. Similarly, our recent study has also shown inverse associations between the concentrations of BA species and risk of GDM in Chinese women.[23] Most of the published studies focused on the relationship between total BA and hyperglycemia, whereas total BA pools and their composition varied widely among different species. The total serum BA is composed of concentrations of individual primary BAs (ie, CA and CDCA), secondary BAs (ie, DCA), and their individual or total glycine-conjugated (ie, GCA, GCDCA, GDCA, and GUDCA) and taurine-conjugated forms (ie, TCA, TCDCA, and TDCA) as well as ratios of some of BAs such as CA/CDCA.[32] BAs from different species differ chemically and their effects on hyperglycemia are varied. Other potential reasons for this discrepancy may include the difference in study populations, the time of BAs being measured during pregnancy, and the marked heterogeneity of GDM.[33] In the present study, our genetic analysis provided significant evidence of a negative correlation between BA species and GDM at early pregnancy in Chinese women. The at-risk variants related to BC dysfunction were inversely associated with the concentration of TDCA, which was one of taurine-conjugated BAs, while TDCA was further inversely associated with the development of GDM, highlighting a genetic link between BAs and GDM. Further large-scale studies integrating genetic data and BA metabolites in various populations are needed to validate our findings.

Increased fasting serum BAs were associated with IR, impaired islet BC function, and increased glucagon levels in patients with T2D.[34] Cariou *et al*[15] reported 1.6-fold increases in DCA in patients with T2D and insulin resistance index (HOMA-IR) was positively related with CA, CDCA, and DCA after adjustment for potential confounders. In addition, Hou *et al*[19] showed that serum total BA level was positively correlated with HOMA-IR and pancreatic BC insulin secretion (HOMA-beta) and increased risk of GDM in Chinese women. In our study, the BC-GRS was negatively associated with plasma BA concentration and was independently

associated with the risk of GDM. However, the association between BC-GRS and GDM was not mediated by BA. One potential explanation is that these genetic variants exhibit pleiotropic features having associations with both BC function and BA concentration. Recent GWASs have identified several genetic variants associated with BAs in White populations,[20] suggesting the contribution of genetic variants to the metabolism of BAs. Moreover, through combining non-targeted metabolomics with genetic analyses, Fall *et al*[21] found that the metabolism of BAs shares some common genetic origin with T2D. It is reported that fasting serum BAs contributed to the effects on glycemia possibly by manipulating BA receptors farnesoid X receptor (FXR) and G-protein coupled receptor (TGR5), in enteroendocrine cells and pancreatic BCs,[12] and genetic variants in *NR1H4,* which encoding the BA receptor FXR, was identified to determine fasting glucose.[35] The above evidence highlighted the presence of heritable factors that can modify BA, as well as glucose metabolism. Taking into consideration genetic pleiotropy in complex traits and diseases, it is likely that the concentration of BAs was also influenced by these genetic variants related to BC dysfunction. Large-scale studies are needed to investigate the causal effects of BA on the risk of GDM.

This analysis is consistent with and extends previous work. Previous studies have found that the magnitude of the association of BC-GRS with the risk of GDM was stronger than that of IR-GRS in Caucasian women.[36] Ding *et al*[10] identified three risk variants related to T2D that were also associated with an increased risk of GDM and assumed that those SNPs impaired the expression of glucagon-like peptide 1 (GLP-1) in enteroendocrine cells to decrease the secretion of insulin. The secretion of GLP-1 could be regulated by TGR5,[37] which is positively responsive to BAs as a cell-surface receptor and improve insulin sensitivity and hepatic metabolism.[38 39] It has been reported that the leading cause of GDM was linked to dysfunction of islet BCs to meet the increased insulin requirements of gestation.[9 40] From a genetic perspective, we demonstrated the contribution of impaired BC function towards the development of GDM, highlighting the potentially pivotal role of BC dysfunction in the pathogenesis of GDM. Interestingly, these genetic variants related to BC dysfunction affected the concentrations of BAs but were independent of BAs to predict GDM. Despite no direct association between IR-GRS and GDM, we detected a significant interaction effect between IR-GRS and BA of TCDCA, which showed that pregnant women with low levels of TCDCA and high genetic risk (high-risk group) were more likely to develop GDM compared with women with higher TCDCA and lower genetic risk. As genetic variants are known to be non-modifiable risk factors for diabetes, it could be used to evaluate the diabetes risk in any stage. However, it is worthwhile to note that there may be interactions between genetic variants and modifiable factors (eg, BAs), and the genetic risk of GDM may vary by these modifiable factors. These findings potentially offer novel information to improve our understanding of the etiology of GDM and help identify women who are at risk of GDM during their early pregnancy.

We acknowledge that there are several limitations in our study. First, a two-step procedure was used to identify incident GDM in this study, which may lead to misclassification of GDM and underestimation of the effect size. Second, although our findings provide evidence of a genetic link between BAs and GDM, we cannot establish a causal relationship between them. Mendelian randomization is one of the approaches to investigate whether BAs are causally linked to GDM,[41] but a large sample size is needed. Third, we did not validate our findings in an independent cohort. However, our findings of associations of BC-GRS and IR-GRS with GDM were consistent with previous studies from the White population. Further replication studies are needed to confirm the results in other populations. Fourth, since some lifestyle and dietary factors which may influence the concentration of BAs were not available for adjustment, we cannot exclude the possibility of residual bias from unmeasured confounders, despite detailed clinical and biochemical information was available in our study. In addition, since there are several fundamental assumptions for tests of mediation, including no misspecification due to unmeasured variables that cause variables in the mediation analysis and no misspecification due to imperfect measurement, the estimates of mediation effects could be biased in our study. However, since these assumptions are often difficult to test and may be untestable in most situations, further work with additional information from prior research, including randomized experimental studies, and larger sample size are needed to consolidate our conclusion.[42 43]

CONCLUSION

In conclusion, as a further study based on our previous findings that serum BAs at the early pregnancy predicted GDM, we discovered a genetic link between BAs and GDM in our Chinese pregnant women. We found that the T2D SNPs related to BC dysfunction independently predicted GDM, and genetic variants related to IR exhibited interaction effects with BAs on the risk of GDM. Women with a high BC-GRS or a high IR-GRS and low concentrations of TCDCA had an increased risk of GDM. These findings may advance our understanding of the genetic basis of GDM and the potential role of BAs in the development of hyperglycemia during pregnancy. The contributions of both modifiable and non-modifiable risk factors may facilitate the identification of high-risk individuals to prevent GDM.

Author affiliations
[1]School of Public Health (Shenzhen), Sun Yat-Sen University, Shenzhen, Guangdong, China
[2]The Second School of Clinical Medicine, Key Laboratory of 3D Printing Technology in Stomatology, Guangdong Medical University, Dongguan, Guangdong, China
[3]Department of Epidemiology and Biostatistics, School of Public Health, Tianjin Medical University, Tianjin, China
[4]Project Office, Tianjin Women and Children's Health Center, Tianjin, China
[5]Population Cancer Research Program and Department of Pediatrics, Dalhousie University, Halifax, Nova Scotia, Canada

[6]Chronic Disease Epidemiology Laboratory, Pennington Biomedical Research Center, Baton Rouge, Louisiana, USA
[7]Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Hong Kong, China
[8]Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Hong Kong, China
[9]Department of Toxicology and Sanitary Chemistry, School of Public Health, Tianjin Medical University, Tianjin, China

Contributors Study concept and design: QZ, YW, Z-ZF, XY, GJ. Acquisition of data: QZ, YW, YG, JLi, HW, JLe, WL, ZY, GH, RCWM, Z-ZF, XY, GJ. Analysis and interpretation of data: QZ, YW, YG, GJ. Writing—original draft: QZ, YW, GJ. Writing—review and editing: QZ, YW, YG, JLi, HW, JLe, WL, ZY, GH, RCWM, Z-ZF, XY, GJ.

Funding This research was supported by the National Key Research and Development Program of China (Grant No:2019YFA0802300), the Fundamental Research Funds for the Central Universities, Sun Yat-sen University (Grant No: 2021QNTD41), the National Natural Science Foundation of China (Grant No: 81870549), and Sailing Plan of Guangdong Province (Grant No:4YF16001G).

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval Ethics approval (ID: 2009–02) was obtained from the Ethics Committee of Tianjin Women and Children's Health Center (TWCHC).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. Data cannot be shared publically due to individual level genetic data which was not consented for sharing on a public platform. Data are available for analysis by qualified researchers who write to contact us requesting the data, who meet the criteria for access to our confidential data. Readers and colleagues who are interested to obtain further information about the study can contact Dr. Xilin Yang at yangxilin@tmu.edu.cn.

ORCID iDs
Junhong Leng http://orcid.org/0000-0002-6742-0830
Gang Hu http://orcid.org/0000-0002-6172-8017
Ronald Ching Wan Ma http://orcid.org/0000-0002-1227-803X
Zhong-Ze Fang http://orcid.org/0000-0003-1113-9262
Xilin Yang http://orcid.org/0000-0002-9462-7992
Guozhi Jiang http://orcid.org/0000-0002-2795-714X

REFERENCES

1 Kim Y, Lee JL, Jang IS, *et al*. Knowledge and health beliefs of gestational diabetes mellitus associated with breastfeeding intention among pregnant women in Bangladesh. *Asian Nurs Res* 2020;14:144–9.
2 Lee KW, Ching SM, Ramachandran V, *et al*. Prevalence and risk factors of gestational diabetes mellitus in Asia: a systematic review and meta-analysis. *BMC Pregnancy Childbirth* 2018;18:494.
3 Gao C, Sun X, Lu L, *et al*. Prevalence of gestational diabetes mellitus in mainland China: a systematic review and meta-analysis. *J Diabetes Investig* 2019;10:154–62.
4 McIntyre HD, Catalano P, Zhang C, *et al*. Gestational diabetes mellitus. *Nat Rev Dis Primers* 2019;5:47.
5 Daly B, Toulis KA, Thomas N, *et al*. Increased risk of ischemic heart disease, hypertension, and type 2 diabetes in women with previous gestational diabetes mellitus, a target group in general practice for preventive interventions: a population-based cohort study. *PLoS Med* 2018;15:e1002488.
6 Fan Y, Li W, Liu H, *et al*. Effects of obesity and a history of gestational diabetes on the risk of postpartum diabetes and hyperglycemia in Chinese women: obesity, GDM and diabetes risk. *Diabetes Res Clin Pract* 2019;156:107828.
7 Vounzoulaki E, Khunti K, Abner SC, *et al*. Progression to type 2 diabetes in women with a known history of gestational diabetes: systematic review and meta-analysis. *BMJ* 2020;369:m1361.
8 Kawai VK, Levinson RT, Adefurin A, *et al*. A genetic risk score that includes common type 2 diabetes risk variants is associated with gestational diabetes. *Clin Endocrinol* 2017;87:149–55.
9 Sullivan SD, Jablonski KA, Florez JC, *et al*. Genetic risk of progression to type 2 diabetes and response to intensive lifestyle or metformin in prediabetic women with and without a history of gestational diabetes mellitus. *Diabetes Care* 2014;37:909–11.
10 Ding M, Chavarro J, Olsen S, *et al*. Genetic variants of gestational diabetes mellitus: a study of 112 SNPs among 8722 women in two independent populations. *Diabetologia* 2018;61:1758–68.
11 Ahmad TR, Haeusler RA. Bile acids in glucose metabolism and insulin signalling - mechanisms and research needs. *Nat Rev Endocrinol* 2019;15:701–12.
12 Rajani C, Jia W. Bile acids and their effects on diabetes. *Front Med* 2018;12:608–23.
13 Shapiro H, Kolodziejczyk AA, Halstuch D, *et al*. Bile acids in glucose metabolism in health and disease. *J Exp Med* 2018;215:383–96.
14 Wu Y, Zhou A, Tang L, *et al*. Bile acids: key regulators and novel treatment targets for type 2 diabetes. *J Diabetes Res* 2020;2020:1–11.
15 Cariou B, Chetiveaux M, Zaïr Y, *et al*. Fasting plasma chenodeoxycholic acid and cholic acid concentrations are inversely correlated with insulin sensitivity in adults. *Nutr Metab* 2011;8:48.
16 Martineau M, Raker C, Powrie R, *et al*. Intrahepatic cholestasis of pregnancy is associated with an increased risk of gestational diabetes. *Eur J Obstet Gynecol Reprod Biol* 2014;176:80–5.
17 Majewska A, Godek B, Bomba-Opon D, *et al*. Association between intrahepatic cholestasis in pregnancy and gestational diabetes mellitus. A retrospective analysis. *Ginekol Pol* 2019;90:458–63.
18 Williamson C, Geenes V. Intrahepatic cholestasis of pregnancy. *Obstet Gynecol* 2014;124:120–33.
19 Hou W, Meng X, Zhao W, *et al*. Elevated first-trimester total bile acid is associated with the risk of subsequent gestational diabetes. *Sci Rep* 2016;6:34070.
20 Ferkingstad E, Oddsson A, Gretarsdottir S, *et al*. Genome-Wide association meta-analysis yields 20 loci associated with gallstone disease. *Nat Commun* 2018;9:5101.
21 Fall T, Salihovic S, Brandmaier S, *et al*. Non-Targeted metabolomics combined with genetic analyses identifies bile acid synthesis and phospholipid metabolism as being associated with incident type 2 diabetes. *Diabetologia* 2016;59:2114–24.
22 Leng J, Shao P, Zhang C, *et al*. Prevalence of gestational diabetes mellitus and its risk factors in Chinese pregnant women: a prospective population-based study in Tianjin, China. *PLoS One* 2015;10:e0121029.
23 Li J, Huo X, Cao Y-F, *et al*. Bile acid metabolites in early pregnancy and risk of gestational diabetes in Chinese women: a nested case-control study. *EBioMedicine* 2018;35:317–24.
24 Anderson CA, Pettersson FH, Clarke GM, *et al*. Data quality control in genetic case-control association studies. *Nat Protoc* 2010;5:1564–73.
25 International Association of Diabetes and Pregnancy Study Groups Consensus Panel, Metzger BE, Gabbe SG, *et al*. International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. *Diabetes Care* 2010;33:676–82.
26 Fang Z-Z, Zhang D, Cao Y-F, *et al*. Irinotecan (CPT-11)-induced elevation of bile acids potentiates suppression of IL-10 expression. *Toxicol Appl Pharmacol* 2016;291:21–7.
27 Mahajan A, Taliun D, Thurner M, *et al*. Fine-Mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet* 2018;50:1505–13.
28 Goodarzi MO, Palmer ND, Cui J, *et al*. Classification of type 2 diabetes genetic variants and a novel genetic risk score association with insulin clearance. *J Clin Endocrinol Metab* 2020;105:1251–60. doi:10.1210/clinem/dgz198
29 Rosseel Y. Lavaan: an R package for structural equation modeling and more. version 0.5–12 (beta). *J Stat Softw* 2012;48:1–36.

30 Wewalka M, Patti M-E, Barbato C, *et al*. Fasting serum taurine-conjugated bile acids are elevated in type 2 diabetes and do not change with intensification of insulin. *J Clin Endocrinol Metab* 2014;99:1442–51.
31 Dudzik D, Zorawski M, Skotnicki M, *et al*. Metabolic fingerprint of gestational diabetes mellitus. *J Proteomics* 2014;103:57–71.
32 Manzotti C, Casazza G, Stimac T, *et al*. Total serum bile acids or serum bile acid profile, or both, for the diagnosis of intrahepatic cholestasis of pregnancy. *Cochrane Database Syst Rev* 2019;7:CD012546.
33 Huvinen E, Eriksson JG, Stach-Lempinen B, *et al*. Heterogeneity of gestational diabetes (GDM) and challenges in developing a GDM risk score. *Acta Diabetol* 2018;55:1251–9.
34 Wang XH, Xu F, Cheng M. Fasting serum total bile acid levels are associated with insulin sensitivity, islet β-cell function and glucagon levels in response to glucose challenge in patients with type 2 diabetes. *Endocr J* 2020:EJ20-0201.
35 Heni M, Wagner R, Ketterer C, *et al*. Genetic variation in NR1H4 encoding the bile acid receptor FXR determines fasting glucose and free fatty acid levels in humans. *J Clin Endocrinol Metab* 2013;98:E1224–9.
36 Shen Y, Jia Y, Li Y, *et al*. Genetic determinants of gestational diabetes mellitus: a case-control study in two independent populations. *Acta Diabetol* 2020;57:843–52.
37 Sonne DP, Hansen M, Knop FK. Bile acid sequestrants in type 2 diabetes: potential effects on GLP1 secretion. *Eur J Endocrinol* 2014;171:R47–65.
38 Pathak P, Xie C, Nichols RG, *et al*. Intestine farnesoid X receptor agonist and the gut microbiota activate G-protein bile acid receptor-1 signaling to improve metabolism. *Hepatology* 2018;68:1574–88.
39 Kawamata Y, Fujii R, Hosoya M, *et al*. A G protein-coupled receptor responsive to bile acids. *J Biol Chem* 2003;278:9435–40.
40 Lipotoxicity NCJ. Beta cell dysfunction, and gestational diabetes. *Cell Metab* 2014;19:553–4.
41 Emdin CA, Khera AV, Kathiresan S. Mendelian randomization. *JAMA* 2017;318:1925–6.
42 MacKinnon DP, Fairchild AJ, Fritz MS. Mediation analysis. *Annu Rev Psychol* 2007;58:593–614.
43 VanderWeele TJ. Mediation analysis: a practitioner's guide. *Annu Rev Public Health* 2016;37:17–32.

Supplemental Table S1 Association of 39 SNPs related to beta-cell dysfunction with GDM

						Model 1 (non-adjustment)		Model 2 (adjustment)	
SNP	**Chr**	**Position**	**Nearest gene**	**MAF**	**risk allele**	**OR**	***P***	**OR**	***P***
rs340874	1	214159256	PROX1	0.398	C	0.94 (0.71-1.24)	0.669	0.96 (0.68-1.34)	0.800
rs6780171	3	185503456	IGF2BP2	0.266	A	0.97 (0.72-1.31)	0.840	1.14 (0.78-1.66)	0.500
rs4481184	3	185505787	IGF2BP2	0.268	T	0.95 (0.70-1.28)	0.722	1.09 (0.75-1.59)	0.648
rs11717959	3	185541213	IGF2BP2	0.201	G	1.01 (0.73-1.42)	0.931	0.88 (0.58-1.33)	0.536
rs1801212	4	6302519	WFS1	0.011	A	1.02 (0.29-3.58)	0.972	0.57 (0.12-2.70)	0.481
rs4457053	5	76424949	ZBED3	0.055	G	0.90(0.51-1.58)	0.714	0.98 (0.49-2.00)	0.966
rs9379084	6	7231843	RREB1	0.126	G	2.41 (1.57-3.70)	1×10^{-4}	1.99 (1.19-3.33)	0.009
rs9505097	6	7255650	RREB1	0.059	C	1.54 (0.87-2.72)	0.135	2.11 (1.02-4.33)	0.043
rs7756992	6	20679709	CDKAL1	0.476	G	1.42 (1.08-1.85)	0.011	1.40 (1.01-1.95)	0.046
rs10228066	7	15063569	DGKB	0.365	T	0.89 (0.68-1.17)	0.392	0.86 (0.61-1.20)	0.364
rs1708302	7	28198677	JAZF1	0.016	C	1.56 (0.54-4.44)	0.409	1.68 (0.53-5.34)	0.379
rs878521	7	44255643	GCK	0.419	A	1.23 (0.94-1.60)	0.126	1.30 (0.95-1.79)	0.106
rs791595	7	127862802	LEP	0.097	A	0.98 (0.64-1.50)	0.917	1.03 (0.62-1.70)	0.911
rs13262861	8	41508577	ANK1	0.131	C	0.89 (0.60-1.32)	0.565	1.17 (0.72-1.89)	0.523
rs4736819	8	41509915	ANK1	0.395	T	0.99 (0.76-1.29)	0.950	0.99 (0.72-1.36)	0.955
rs3802177	8	118185025	SLC30A8	0.407	G	1.08 (0.83-1.40)	0.589	1.19 (0.87-1.63)	0.286
rs10974438	9	4291928	GLIS3	0.357	C	1.07 (0.82-1.41)	0.609	1.08 (0.77-1.49)	0.665
rs10811660	9	22134068	CDKN2A/B	0.486	G	1.06 (0.82-1.37)	0.656	1.05 (0.76-1.44)	0.771
rs10757283	9	22134172	CDKN2A/B	0.355	T	0.90 (0.69-1.18)	0.432	0.91 (0.66-1.27)	0.584
rs505922	9	136149229	ABO	0.469	C	1.00 (0.76-1.32)	0.983	0.94 (0.67-1.31)	0.705
rs12378717	9	139286060	GPSM1	0.029	G	1.21 (0.54-2.67)	0.645	1.40 (0.53-3.66)	0.497
rs11257655	10	12307894	CDC123/CAMK1D	0.435	T	0.76 (0.59-0.98)	0.032	0.81 (0.59-1.10)	0.180

rs10882101	10	94462427	HHEX/IDE	0.286	T	1.11 (0.83-1.48)	0.497	0.98 (0.68-1.40)	0.893
rs1112718	10	94479107	HHEX/IDE	0.177	A	1.21 (0.87-1.70)	0.261	1.22 (0.82-1.83)	0.332
rs7903146	10	114758349	TCF7L2	0.048	T	0.98 (0.52-1.82)	0.939	1.22 (0.56-2.64)	0.612
rs34855922	10	114871594	TCF7L2	0.012	A	0.85 (0.26-2.82)	0.789	0.52 (0.09-2.94)	0.463
rs234853	11	2850828	KCNQ1	0.353	G	0.80 (0.61-1.06)	0.118	0.83 (0.59-1.16)	0.274
rs2237895	11	2857194	KCNQ1	0.352	C	1.27 (0.96-1.66)	0.091	1.15 (0.82-1.61)	0.414
rs2237897	11	2858546	KCNQ1	0.311	C	1.28 (0.97-1.70)	0.085	1.40 (0.98-1.99)	0.062
rs445084	11	2908754	KCNQ1	0.114	G	1.02 (0.68-1.53)	0.930	0.89 (0.54-1.46)	0.636
rs102275	11	61557803	TMEM258	0.302	T	1.02 (0.77-1.36)	0.882	0.99 (0.70-1.40)	0.964
rs77464186	11	72460398	CENTD2/ARAP1	0.087	A	0.83 (0.52-1.35)	0.455	0.69 (0.37-1.27)	0.234
rs10830963	11	92708710	MTNR1B	0.452	G	1.44 (1.12-1.86)	0.005	1.72 (1.26-2.34)	0.001
rs57235767	11	93013531	MTNR1B	0.218	C	1.08 (0.79-1.48)	0.621	1.30 (0.88-1.91)	0.185
rs10842994	12	27965150	KLHDC5	0.167	C	0.94 (0.66-1.32)	0.703	0.88 (0.57-1.34)	0.545
rs1359790	13	80717156	SPRY2	0.274	G	1.12 (0.83-1.50)	0.462	1.05 (0.72-1.52)	0.812
rs8038040	15	62394264	C2CD4A/B	0.392	G	1.17 (0.89-1.52)	0.255	0.95 (0.68-1.33)	0.775
rs1005752	15	77818128	HMG20A	0.354	A	0.89 (0.68-1.18)	0.422	0.92 (0.66-1.28)	0.617
rs12910825	15	91511260	PRC1	0.018	G	1.74 (0.62-4.86)	0.293	1.73 (0.50-6.02)	0.391

Model 2 was adjusted for factors age, gestational weeks, baseline BMI, family history of diabetes, drinking history, smoking history, education attainment, ethnicity group, weight gain during pregnancy, ALT, SBP, DBP and parity.

Supplemental Table S2 Association of 23 SNPs related to insulin resistance with GDM

SNP	Chr	Position	Nearest gene	MAF	risk allele	Model 1 (non-adjustment)		Model 2 (adjustment)	
						OR	*P*	OR	*P*
rs3768321	1	40035928	MACF1	0.121	T	1.02 (0.68-1.53)	0.934	1.30 (0.78-2.16)	0.319
rs1260326	2	27730940	GCKR	0.459	C	1.24 (0.95-1.62)	0.114	1.26 (0.92-1.74)	0.154
rs2249105	2	65287896	CEP68	0.354	A	1.25 (0.95-1.64)	0.107	1.28 (0.92-1.78)	0.142
rs2052261	2	65355270	CEP68	0.359	G	0.87 (0.67-1.14)	0.319	0.88 (0.63-1.21)	0.428
rs10195252	2	165513091	GRB14/COBLL1	0.119	T	0.98 (0.65-1.48)	0.933	1.00 (0.60-1.66)	0.992
rs2972144	2	227101411	IRS1	0.081	G	1.46 (0.90-2.37)	0.122	1.21 (0.66-2.23)	0.531
rs11709077	3	12336507	PPARG	0.041	G	1.87 (0.95-3.66)	0.070	1.77 (0.82-3.79)	0.143
rs11926707	3	46925539	KIF9	0.318	C	0.96 (0.73-1.27)	0.784	0.77 (0.54-1.10)	0.158
rs9860730	3	64701146	ADAMTS9	0.257	A	1.17 (0.87-1.59)	0.303	1.23 (0.86-1.77)	0.260
rs28819812	4	157652753	PDGFC	0.318	C	0.96 (0.72-1.28)	0.779	1.01 (0.71-1.43)	0.962
rs702634	5	53271420	ARL15	0.123	A	1.11 (0.75-1.62)	0.609	1.15 (0.73-1.82)	0.553
rs465002	5	55808475	ANKRD55	0.446	T	1.21 (0.94-1.57)	0.141	1.47 (1.07-2.01)	0.017
rs9687832	5	55861595	ANKRD55	0.107	A	1.15 (0.75-1.75)	0.523	1.01 (0.62-1.67)	0.957
rs2280141	10	124193181	PLEKHA1	0.373	T	0.94 (0.73-1.22)	0.655	1.09 (0.80-1.49)	0.574
rs4148856	12	123450765	MPHOSPH9	0.153	C	1.04 (0.73-1.50)	0.815	1.11 (0.72-1.69)	0.638
rs2738809	16	75516534	BCAR1	0.467	G	0.85 (0.66-1.09)	0.203	0.72 (0.52-0.99)	0.042
rs2925979	16	81534790	CMIP	0.373	T	1.11 (0.85-1.44)	0.450	1.28 (0.93-1.76)	0.134
rs12454712	18	60845884	BCL2A	0.474	T	1.05 (0.81-1.37)	0.702	1.04 (0.75-1.44)	0.821
rs8107974	19	19388500	TM6SF2	0.075	T	0.92 (0.56-1.50)	0.731	1.40 (0.75-2.61)	0.293
rs889138	19	33890838	PEPD	0.467	C	1.05 (0.81-1.36)	0.711	0.96 (0.70-1.33)	0.820
rs10406431	19	46157019	GIPR	0.295	A	0.89 (0.67-1.18)	0.421	0.89 (0.63-1.26)	0.500
rs2238689	19	46178661	GIPR	0.302	C	1.04 (0.79-1.39)	0.771	1.13 (0.80-1.59)	0.491

rs112915006	22	50604696	PIM3	0.152	G	1.15 (0.81-1.64)	0.443	1.43 (0.92-2.23)	0.112

Model 2 was adjusted for factors: age, gestational weeks, baseline BMI, family history of diabetes, drinking history, smoking history, education attainment, ethnicity group, weight gain during pregnancy, ALT, SBP, DBP and parity.

Supplemental Table S3 Distribution of other bile acid species measured in the cohort

	Bile acid (nmol/mL)			Missing
	Non-GDM group	GDM group	Total	
THDCA	0.009 (0.008,0.010)	NA	0.009 (0.008,0.010)	351 (78.5%)
TUDCA	0.009 (0.006,0.013)	NA	0.009 (0.006,0.013)	337 (75.4%)
LCA	0.049 (0.039,0.062)	NA	0.050 (0.040,0.060)	311 (69.6%)
GLCA	0.011 (0.007,0.023)	NA	0.010 (0.010,0.020)	278 (62.2%)
TLCA	0.002 (0.001,0.005)	0.003 (0.002,0.004)	0.002 (0.001,0.005)	270 (60.4%)
DHCA	0.350 (0.250,0.480)	0.170 (0.140,0.220)	0.250 (0.160,0.400)	142 (31.8%)
UDCA	0.024 (0.014,0.037)	0.025 (0.019,0.038)	0.024 (0.016,0.038)	104 (23.3%)

Non-GDM group were matched on age ±1 year of the GDM group. Values are described as median (Q1, Q3) or n, n (%).

Abbreviations: THDCA, taurohyodeoxycholic acid; TUDCA, tauroursodeoxycholic acid; LCA, lithocholic acid; GLCA, glycolithocholic acid; TLCA, taurolithocholic acid; DHCA, dehydrocholic acid; UDCA, ursodeoxycholic acid.

Supplemental Table S4 Associations of GRS and bile acids in Chinese women

	BC-GRS (#SNP=39)				**IR-GRS (#SNP=23)**			
	Model 1 (non-adjustment)		Model 2 (adjustment)		Model 1 (non-adjustment)		Model 2 (adjustment)	
	Beta (SE)	p-value	Beta (SE)	p-value	Beta (SE)	p-value	Beta (SE)	p-value
TDCA	-0.130(0.044)	0.003	-0.177(0.048)	**2.66×10^{-4}**	0.001(0.044)	0.984	0.036(0.050)	0.465
GDCA	-0.110(0.048)	0.024	-0.141(0.053)	0.009	-0.043(0.049)	0.376	0.008(0.055)	0.890
TCDCA	-0.100(0.041)	0.014	-0.108(0.046)	0.020	0.035(0.041)	0.390	0.078(0.048)	0.103
GCA	-0.104(0.044)	0.019	-0.113(0.052)	0.029	0.042(0.045)	0.349	0.082(0.053)	0.122
TCA	-0.073(0.029)	0.014	-0.072(0.035)	0.039	0.054(0.030)	0.070	0.072(0.036)	0.044
GCDCA	-0.106(0.045)	0.018	-0.099(0.051)	0.051	0.032(0.045)	0.481	0.045(0.052)	0.388
DCA	-0.045(0.037)	0.229	-0.081(0.043)	0.060	-0.060(0.037)	0.104	-0.040(0.044)	0.363
CA	-0.017(0.027)	0.522	-0.042(0.033)	0.211	-0.046(0.027)	0.089	-0.060(0.034)	0.076
CDCA	-0.033(0.045)	0.460	-0.056(0.052)	0.289	-0.021(0.045)	0.648	-0.045(0.053)	0.404
HDCA	-0.015(0.034)	0.654	-0.037(0.040)	0.350	-0.020(0.033)	0.546	-0.022(0.039)	0.577
GUDCA	-0.034(0.049)	0.489	-0.019(0.058)	0.751	0.004(0.049)	0.930	-0.005(0.058)	0.935

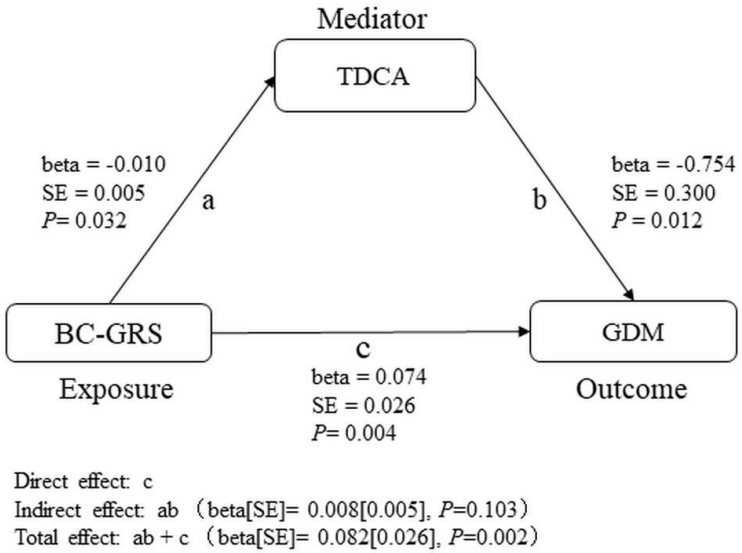
Abbreviations: BC, beta cell; IR, insulin resistance; TDCA, taurodeoxycholic acid; GDCA, glycodeoxycholic acid; TCDCA, taurochenodeoxycholic acid; GCA, glycocholic acid; TCA, taurocholic acid; GCDCA, glycochenodeoxycholic acid; DCA, deoxycholic acid; CA, cholic acid; CDCA, chenodeoxycholic acid; HDCA, hyodeoxycholic acid; GUDCA, glycoursodeoxycholic acid.

Model 2 was adjusted for conventional risk factors including age, gestational weeks, baseline BMI, family history of diabetes, drinking history, smoking history, education attainment, ethnicity group, weight gain during pregnancy, ALT, SBP, DBP and parity.

Supplemental Table S5 Interaction between IR-GRS and TCDCA

Concentration	OR (95% CI)			*P* for trend
	Tertile 1	Tertile 2	Tertile 3	
High (> 0.2 nmol/L)	1	4.51 (0.41,49.26)	6.92 (0.68,70.77)	0.025
Low (≤ 0.2 nmol/L)	10.95 (1.21,99.51)	11.70 (1.29,106.44)	14.39 (1.59,130.16)	0.366

ORs were adjusted for conventional risk factors including age, gestational weeks, baseline BMI, family history of diabetes, drinking history, smoking history, education attainment, ethnicity group, weight gain during pregnancy, ALT, SBP, DBP and parity. Abbreviations: IR, insulin resistance. TCDCA, taurochenodeoxycholic acid.



Supplemental Figure 1 Mediation analysis of the role of TDCA in mediating the relationship between BC-GRS and GDM.
a, the association between BC-GRS and TDCA; b, the association between TDCA and GDM; c, the direct association between BC-GRS and GDM after adjustment for TDCA; ab, the indirect effect (or mediation effect) of BC-GRS on GDM through TDCA; ab + c, the total effect of BC-GRS on GDM. All associations were adjusted for age, gestational weeks, baseline BMI, family history of diabetes, drinking history, smoking history, education attainment, ethnicity group, weight gain during pregnancy, ALT, SBP, DBP, and parity.