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,5 Zhijie Yu,5 Gang Hu,6 Ronald Ching Wan Ma,7,8 Zhong-Ze Fang,6,9 Xilin Yang,3 Guozhi Jiang1

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 taurochenodeoxycholic acid (TCDCA) after adjustment for confounders (Beta (SE)=-0.030 (0.049); p=0.465). The BC-GRS was also associated with the risk of GDM (OR (95% CI): 1.00 (1.00 to 1.00); p=0.000), but not mediated by TCDCA. Compared with individuals in the low tertile of BC-GRS, the OR for GDM was 2.25 (95% CI: 1.40 to 1.77); p=0.005. Women with high BC-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 taurochenodeoxycholic acid (TCDCA), 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 TCDCA.

⇒ 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 Asians2 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.4–5 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 pathophysiology of GDM.6 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 colleagues10 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 physiological 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 BAs20 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 22302 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 22302 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.,24–25 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

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>1×10^-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:

\[
\text{individual GRS value} = \text{population mean GRS} \times \frac{\text{rescaled SD of GRS}}{\text{population SD of GRS}}.
\]

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 alanotransferase (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), glyc obsessoxycholic acid (GDCA), glycodeoxycholic acid (GCDCA), taurocholic acid (TCA), taurochenodeoxycholic acid (TCDA), taurodeoxycholic acid (TDCA), hydoxycholeic acid (HDDA), 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, HDDA, TCDA, 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=2.66×10^-5). Other BA species, including GDCDA (beta (SE)=-0.140 (0.053) per SD; p=0.009), TCDA (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;
Genetics/Genomes/Proteomics/Metabolomics

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.

**Table 1 Clinical and biochemical characteristics of participants**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Non-GDM</th>
<th>GDM</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>217</td>
<td>230</td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>29.1±3.32</td>
<td>29.2±2.72</td>
<td>0.808</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>21.9±3.45</td>
<td>23.9±3.64</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>DBP (mm Hg)</td>
<td>67.8±7.58</td>
<td>70.6±7.96</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>SBP (mm Hg)</td>
<td>104±10.5</td>
<td>108±10.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>ALT (U/L)</td>
<td>16 (11,22)</td>
<td>19 (14,26)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>GCT glucose (mmol/L)</td>
<td>6.27 (5.39, 7.20)</td>
<td>9.00 (8.36, 9.96)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Gestational weeks</td>
<td>10.1±2.05</td>
<td>10.1±2.10</td>
<td>0.937</td>
</tr>
<tr>
<td>Han nationality</td>
<td>210 (96.8%)</td>
<td>225 (97.8%)</td>
<td>0.693</td>
</tr>
<tr>
<td>Parity ≥1</td>
<td>11 (5.1%)</td>
<td>13 (5.7%)</td>
<td>0.949</td>
</tr>
<tr>
<td>Family history of diabetes</td>
<td>13 (6.0%)</td>
<td>28 (12.2%)</td>
<td>0.168</td>
</tr>
<tr>
<td>Habitual smoker (yes)</td>
<td>13 (6.0%)</td>
<td>15 (6.5%)</td>
<td>0.867</td>
</tr>
<tr>
<td>Alcohol drinker (yes)</td>
<td>53 (24.4%)</td>
<td>69 (30.0%)</td>
<td>0.999</td>
</tr>
<tr>
<td>Education&gt;12 years</td>
<td>216 (48.9%)</td>
<td>226 (51.1%)</td>
<td>0.080</td>
</tr>
<tr>
<td>Weight gain up to GCT (kg)</td>
<td>8.72±3.28</td>
<td>8.32±3.60</td>
<td>0.229</td>
</tr>
<tr>
<td>Bile acid (nmol/mL)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DCA</td>
<td>0.26 (0.16, 0.45)</td>
<td>0.20 (0.10, 0.34)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>GCA</td>
<td>0.07 (0.04, 0.13)</td>
<td>0.05 (0.03, 0.09)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>GDCA</td>
<td>0.12 (0.06, 0.23)</td>
<td>0.08 (0.04, 0.14)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>GCDCDA</td>
<td>0.35 (0.16, 0.63)</td>
<td>0.21 (0.12, 0.38)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>GUDCA</td>
<td>0.03 (0.02, 0.05)</td>
<td>0.02 (0.01, 0.03)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>TCDCDA</td>
<td>0.10 (0.05, 0.19)</td>
<td>0.06 (0.04, 0.10)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>TDCA</td>
<td>0.04 (0.02, 0.07)</td>
<td>0.03 (0.02, 0.05)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CDCA</td>
<td>0.09 (0.05, 0.20)</td>
<td>0.08 (0.04, 0.14)</td>
<td>0.024</td>
</tr>
<tr>
<td>HDCA</td>
<td>0.03 (0.02, 0.04)</td>
<td>0.02 (0.01, 0.04)</td>
<td>0.040</td>
</tr>
<tr>
<td>TCA</td>
<td>0.05 (0.04, 0.09)</td>
<td>0.06 (0.05, 0.08)</td>
<td>0.056</td>
</tr>
<tr>
<td>CA</td>
<td>0.10 (0.08, 0.15)</td>
<td>0.10 (0.09, 0.13)</td>
<td>0.153</td>
</tr>
</tbody>
</table>

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; GCDCDA, glycodeoxycholic acid; GCT, glucose challenge test; GDA, glycoursodeoxycholic acid; GDM, gestational diabetes mellitus; GUDCA, glycoursodeoxycholic acid; HDCA, hyodeoxycholic acid; SBP, systolic blood pressure; TCA, taurocholic acid; TCDCDA, 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 GRSs 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, 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**

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

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.
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 25 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 TCDCA. 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 colleagues19 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 colleagues31 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.25 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 al35 reported 1.6-fold increases in DGA 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 al39 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, suggesting the contribution of genetic variants to the metabolism of BAs. Moreover, through combining non-targeted metabolomics with genetic analyses, Fall et al. 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, and genetic variants in NR1H4, which encoding the BA receptor FXR, was identified to determine fasting glucose. 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. Ding et al. 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, which is positively responsive to BAs as a cell-surface receptor and improve insulin sensitivity and hepatic metabolism. 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. 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 (e.g., 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, 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.

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 Stomatolgy, 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
REFERENCES


