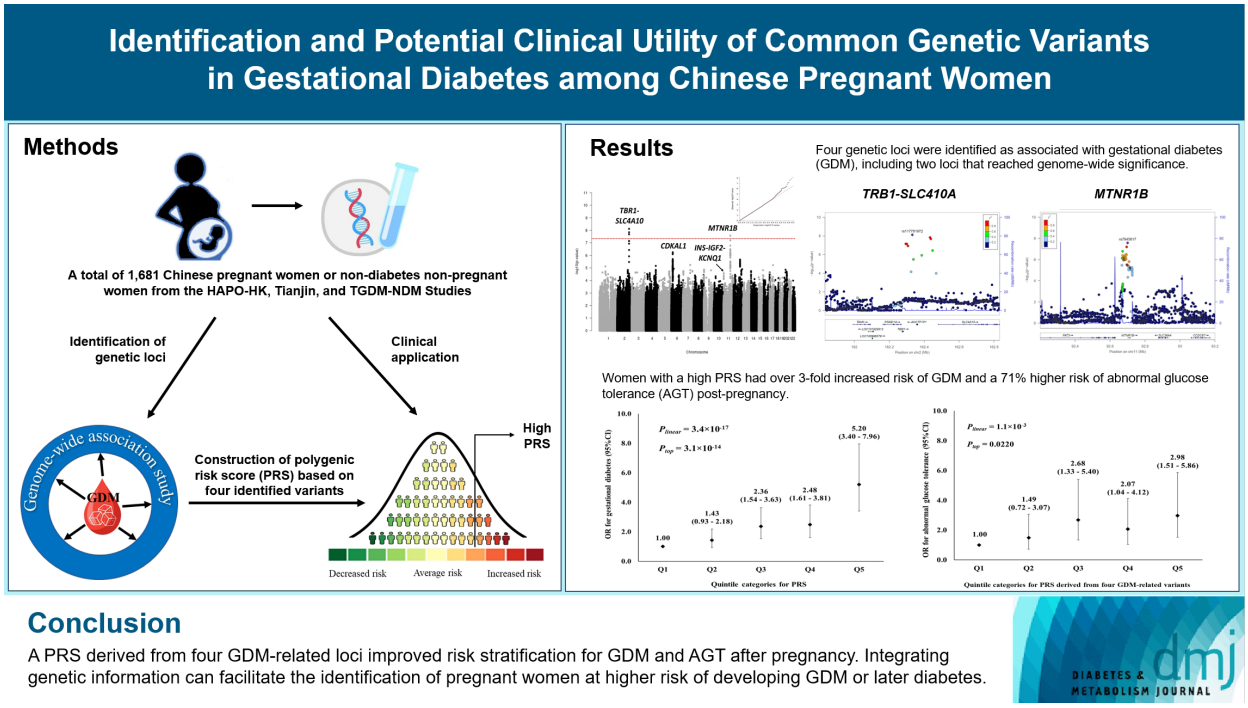
dmj
DIABETES & METABOLISM JOURNAL

Identification and Potential Clinical Utility of Common Genetic Variants in Gestational Diabetes among Chinese Pregnant Women

Claudia Ha-ting Tam, Ying Wang[4], Chi Chiu Wang, Lai Yuk Yuen, Cadmon King-poo Lim, Junhong Leng, Ling Wu, Alex Chi-wai Ng, Yong Hou, Kit Ying Tsoi, Hui Wang, Risa Ozaki, Albert Martin Li, Qingqing Wang, Juliana Chung-ngor Chan, Yan Chou Ye, Wing Hung Tam, Xilin Yang, Ronald Ching-wan Ma

Diabetes Metab J 2025;49:128-143 | *https://doi.org/10.4093/dmj.2024.0139*



Highlights

- **We explored the genetic factors of gestational diabetes (GDM) and their applications.**
- **We confirmed three loci and identified a novel locus for GDM.**
- **A polygenic risk score (PRS) from these loci effectively stratifies GDM risk.**
- **The PRS also stratifies the risk of abnormal glucose tolerance post-pregnancy.**
- **Genetic information may help identify high-risk women to prevent GDM.**

How to cite this article:
Tam CH, Wang Y, Wang CC, Yuen LY, Lim CK, Leng J et al. Identification and Potential Clinical Utility of Common Genetic Variants in Gestational Diabetes among Chinese Pregnant Women. Diabetes Metab J 2025;49:128-143. https://doi.org/10.4093/dmj.2024.0139

Original Article
Genetics

Diabetes Metab J 2025;49:128-143
https://doi.org/10.4093/dmj.2024.0139
pISSN 2233-6079 · eISSN 2233-6087

dmj
DIABETES & METABOLISM JOURNAL

Identification and Potential Clinical Utility of Common Genetic Variants in Gestational Diabetes among Chinese Pregnant Women

Claudia Ha-ting Tam[1,2,3], Ying Wang[4], Chi Chiu Wang[5,6,7,8], Lai Yuk Yuen[5], Cadmon King-poo Lim[1,2,3], Junhong Leng[9], Ling Wu[5], Alex Chi-wai Ng[1], Yong Hou[1], Kit Ying Tsoi[1], Hui Wang[10], Risa Ozaki[1,2], Albert Martin Li[11], Qingqing Wang[12], Juliana Chung-ngor Chan[1,2,3,13], Yan Chou Ye[14], Wing Hung Tam[5], Xilin Yang[10], Ronald Ching-wan Ma[1,2,3,13]

[1]Department of Medicine and Therapeutics, Prince of Wales Hospital, The Chinese University of Hong Kong, Hong Kong,
[2]Hong Kong Institute of Diabetes and Obesity, [3]CUHK-SJTU Joint Research Center in Diabetes Genomics and Precision Medicine, The Chinese University of Hong Kong, Hong Kong,
[4]Scientific Research Platform of the Second School of Clinical Medicine, Guangdong Medical University, Dongguan,
[5]Department of Obstetrics and Gynecology, Prince of Wales Hospital, The Chinese University of Hong Kong, Hong Kong,
[6]Development and Reproduction Laboratory, Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Hong Kong,
[7]School of Biomedical Sciences, The Chinese University of Hong Kong, Hong Kong, Hong Kong,
[8]Chinese University of Hong Kong-Sichuan University Joint Laboratory in Reproductive Medicine, The Chinese University of Hong Kong, Hong Kong,
[9]Department of Children's Health, Tianjin Women and Children's Health Center, Tianjin,
[10]Department of Epidemiology and Biostatistics, School of Public Health, Tianjin Medical University, Tianjin,
[11]Department of Pediatrics, Prince of Wales Hospital, The Chinese University of Hong Kong, Hong Kong,
[12]Department of Obstetrics and Gynecology, The Third Affiliated Hospital of Sun Yat-Sen University, Guangzhou,
[13]Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Hong Kong,
[14]Department of Obstetrics and Gynecology, The Seventh Affiliated Hospital of Sun Yat-Sen University, Shenzhen, China

Background: The genetic basis for hyperglycaemia in pregnancy remain unclear. This study aimed to uncover the genetic determinants of gestational diabetes mellitus (GDM) and investigate their applications.
Methods: We performed a meta-analysis of genome-wide association studies (GWAS) for GDM in Chinese women (464 cases and 1,217 controls), followed by *de novo* replications in an independent Chinese cohort (564 cases and 572 controls) and *in silico* replication in European (12,332 cases and 131,109 controls) and multi-ethnic populations (5,485 cases and 347,856 controls). A polygenic risk score (PRS) was derived based on the identified variants.
Results: Using the genome-wide scan and candidate gene approaches, we identified four susceptibility loci for GDM. These included three previously reported loci for GDM and type 2 diabetes mellitus (T2DM) at *MTNR1B* (rs7945617, odds ratio [OR], 1.64; 95% confidence interval [CI], 1.38 to 1.96), *CDKAL1* (rs7754840, OR, 1.33; 95% CI, 1.13 to 1.58), and *INS-IGF2-KCNQ1* (rs2237897, OR, 1.48; 95% CI, 1.23 to 1.79), as well as a novel genome-wide significant locus near *TBR1-SLC4A10* (rs117781972, OR, 2.05; 95% CI, 1.61 to 2.62; $P_{meta}=7.6\times10^{-9}$), which has not been previously reported in GWAS for T2DM or glycaemic traits. Moreover, we found that women with a high PRS (top quintile) had over threefold (95% CI, 2.30 to 4.09; $P_{meta}=3.1\times10^{-14}$) and 71% (95% CI, 1.08 to 2.71; $P=0.0220$) higher risk for GDM and abnormal glucose tolerance post-pregnancy, respectively, compared to other individuals.
Conclusion: Our results indicate that the genetic architecture of glucose metabolism exhibits both similarities and differences between the pregnant and non-pregnant states. Integrating genetic information can facilitate identification of pregnant women at a higher risk of developing GDM or later diabetes.

Keywords: Diabetes, gestational; Genetic risk score; Genome-wide association study; Glucose intolerance; Pregnant women

Corresponding authors: Xilin Yang https://orcid.org/0000-0002-9462-7992
Department of Epidemiology and Biostatistics, School of Public Health, Tianjin Medical University, 22 Qixiangtai Road, Tianjin, China
E-mail: yangxilin@tmu.edu.cn

Ronald Ching-wan Ma https://orcid.org/0000-0002-1227-803X
Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Prince of Wales Hospital, 30-32 Ngan Shing Street, Shatin, Hong Kong, China
E-mail: rcwma@cuhk.edu.hk

Received: Mar. 20, 2024; Accepted: Jun. 17, 2024

INTRODUCTION

Gestational diabetes mellitus (GDM) occurs when insulin production or utilization is impaired during pregnancy, resulting in hyperglycemia. It typically resolves after delivery. GDM is a prevalent condition, affecting 14.0% of live births worldwide in 2021 [1]. Unmanaged GDM can lead to adverse perinatal outcomes [2] and increased long-term metabolic risks for both mothers and children. A recent meta-analysis assessing over 1.3 million individuals reported that women with GDM have a nearly 10-fold increased risk of developing type 2 diabetes mellitus (T2DM) compared with healthy controls [3]. Moreover, offspring exposed to hyperglycemia *in utero* are more likely to be obese and insulin resistant in childhood and early adulthood than offspring of mothers with normoglycemia [4-6]. These observations indicate that (1) GDM shares a common pathology with T2DM; and (2) GDM can provide insights into the predisposition of T2DM, with pregnancy serving as a stressor that unmask hyperglycemia during pregnancy among those who are at risk of future diabetes because of impaired β-cell function.

Given the increased lifetime risk of T2DM associated with GDM, many genetic studies have focused on the assumption that the genetic architecture of both conditions shared similarities. A number of T2DM susceptibility loci have been examined for their association with GDM in modest sample sizes [7]. In a recent meta-analysis of 23 studies, eight T2DM loci (insulin like growth factor 2 mRNA binding protein 2 [*IGF2BP2*], CDK5 regulatory subunit-associated protein 1-like 1 [*CDKAL1*], GLIS family zinc finger 3 [*GLIS3*], cyclin dependent kinase inhibitor 2A/2B [*CDKN2A/2B*], hematopoietically expressed homeobox [*HHEX*]/insulin degrading enzyme [*IDE*] transcription factor 7 like 2 [*TCF7L2*], melatonin receptor 1B [*MTNR1B*], and HNF1 homeobox A [*HNF1A*]) showed association with GDM after correcting for multiple comparisons [7]. Notably, these loci demonstrated a consistent direction of association with T2DM, supporting a shared genetic background between the two conditions. More recently, the GENetics of Diabetes In Pregnancy (GenDIP) Consortium conducted a multi-ancestry genome-wide association studies (GWAS) meta-analysis for GDM, involving over 300K women [8]. Among the five genome-wide significant loci identified, four were previously reported to be associated with T2DM: *MTNR1B, TCF7L2, CDKAL1*, and *CDKN2A/B*. Interestingly, the novel locus at hexokinase domain containing 1 (*HKDC1*) did not show a strong association with T2DM, highlighting the presence of genetic determinants specific to glucose regulation in pregnancy. The FinnGen Study further explored genetic features distinct from T2DM and identified eight loci with effects on GDM that were three times stronger compared to T2DM [9].

There is growing interest in utilizing a polygenic risk score (PRS) to assess the individual's risks of developing GDM and T2DM postpartum. Prior efforts to construct PRSs for GDM typically involved variants associated with T2DM, and suffered from a small number of variants and participants. As a result, these scores had limited success in enhancing predictive capability compared to models that included only clinical variables [10]. For instance, in a study of Caucasian women involving 458 GDM cases and 1,538 controls, a PRS consisting of 34 variants related to T2DM and/or fasting glucose was associated with increased risk of GDM. However, its utility in identifying GDM cases was limited, with only a modest improvement in GDM prediction when added to clinical factors (increase in c-statistic=0.03) [11].

This study aimed to (1) identify novel GDM susceptibility loci by performing a meta-analysis of three GWASs in Chinese women, followed by replication studies in independent Chinese, European, and multi-ethnic cohorts; and (2) explore the clinical utility of the identified variants by deriving a PRS for GDM, and evaluating its predictive value for GDM and abnormal glucose tolerance (AGT) at 7-year postpartum.

METHODS

Study design and participants

The overall design of the current study is shown in Supplementary Fig. 1. Details of the design, ascertainment methods, inclusion criteria, and phenotyping procedures for each cohort are outlined in the "cohort descriptions" (Supplementary Methods). Supplementary Tables 1 and 2 summarize the clinical characteristics of participants in the discovery and replication studies. At the time of assessment, all participants provided written informed consent for DNA collection and data analysis for research purposes. Institutional review boards of the respective institutions approved each study.

For the genome-wide scan, a meta-analysis was conducted on 1,681 (464 cases and 1,217 controls) women of Han Chinese ancestry from three independent cohorts, including: (1) The Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong (HAPO-HK) Study, which consisted of 960 pregnant

women (149 GDM cases vs. 811 non-GDM controls) attending a follow-up visit at the Hong Kong center [6]; (2) The Tianjin GDM Study, which involved 455 pregnant women (229 GDM cases vs. 226 non-GDM controls) participating in a two-step GDM screening program conducted by Tianjin's Women and Children Health Center [12]; and (3) The Treated GDM Cases vs. Non-diabetes Controls (TGDM-NDM) Study, which included 86 pregnant women with GDM receiving antenatal treatment at the Hong Kong center [13], and 180 non-diabetes, non-pregnant women recruited from hospital staff and a community-based health screening program [14]. The flow diagrams of cohort selection in the HAPO-HK and Tianjin Studies are shown in Supplementary Figs. 2 and 3.

The first-stage *de novo* replication involved 1,136 Southern Han Chinese pregnant women (564 GDM cases and 572 non-GDM controls) from the Guangzhou Study [15]. In the second-stage *in silico* replication, we accessed data from two published GDM GWAS, including the FinnGen Study (12,332 cases and 131,109 controls) [9] and a multi-ancestry meta-analysis of GDM contributed by the GenDIP Consortium (5,485 GDM cases and 347,856 non-GDM controls) [8]. The HAPO-HK Study, our discovery cohort, participated in the GenDIP meta-analysis. Additionally, individual-level data from the Thai (HAPO-Thai; 260 GDM cases and 933 non-GDM controls) and Hispanic (HAPO-Hispanic; 207 GDM cases and 596 non-GDM controls) populations of the HAPO Study were utilized to validate the results of the PRS analysis [2].

Clinical outcomes

GDM was diagnosed using criteria established in 2010 by the International Association of the Diabetes and Pregnancy Study Groups (IADPSG) [16]. These criteria included fasting plasma glucose levels ≥5.1 mmol/L, or 1-hour glucose levels ≥10.0 mmol/L, or 2-hour glucose levels ≥8.5 mmol/L, as determined from a 75-g oral glucose tolerance test conducted during the third trimester of gestation.

AGT was defined as the presence of impaired glucose tolerance, impaired fasting glucose, or diabetes, as per the diagnostic criteria of the American Diabetes Association [17].

Statistical analyses

Within cohort, logistic regression was used to examine the association between genetic markers (e.g., a single nucleotide polymorphism [SNP] or a PRS) and the risks of GDM and AGT after pregnancy, with the adjustments for the first four principal components (PCs), age and/or body mass index (BMI). Meta-analysis was conducted with use of the inverse-variance weighted method under a fixed-effects model. The area under the curve (AUC) of the receiver operating characteristic and continuous net reclassification improvement (NRI) index were used to evaluate the incremental predictive value of PRS in GDM and AGT after pregnancy, over the clinical risk factors and PCs.

Detailed methods for (1) clinical and laboratory measurements; (2) genotyping, quality controls, and imputation; (3) construction of PRS; and (4) statistical analysis can be found in Supplementary Methods.

Data availability

The data that support the findings of this study are not openly available due to reasons of sensitivity and are available from the corresponding author upon reasonable request. The HAPO-Thai and -Hispanic data were obtained from database of Genotypes and Phenotypes (dbGAP; phs000096.v4.p1).

RESULTS

Discovery genome-wide scan analysis

We performed a meta-analysis of three GWASs for GDM in Chinese women. The results of the genome-wide scan are illustrated in Fig. 1, displaying Manhattan plots and corresponding quantile-quantile (Q-Q) plots. Among the 6,322,337 biallelic and autosomal SNPs available in all discovery cohorts, we identified two loci that reached genome-wide significance. These included a previously reported locus for GDM and T2DM susceptibility at *MTNR1B* (rs7945617, odds ratio [OR], 1.64; 95% confidence interval [CI], 1.38 to 1.96; $P_{meta}=2.6\times10^{-8}$), and a novel locus near T-box brain transcription factor 1 (*TBR1*)-solute carrier family 4 member 10 (*SLC4A10*) (rs117781972, OR, 2.05; 95% CI, 1.61 to 2.62; $P_{meta}=7.6\times10^{-9}$) (Fig. 2, Supplementary Tables 3 and 4). The results remained largely unchanged after applying genomic control correction (Supplementary Tables 3 and 4). Conditioning on *TBR1-SLC4A10* rs117781972 and *MTNR1B* rs7945617 did not reveal any residual associations (data not shown).

The lead SNP from the two genome-wide significant loci underwent (1) *de novo* replication in the Guangzhou Study and (2) *in silico* replication using data from the FinnGen Study and the GenDIP meta-analysis. The *TBR1-SLC4A10* rs117781972 variant did not replicate in the Guangzhou and FinnGen Studies ($P>0.05$) and it only showed nominal significance in the

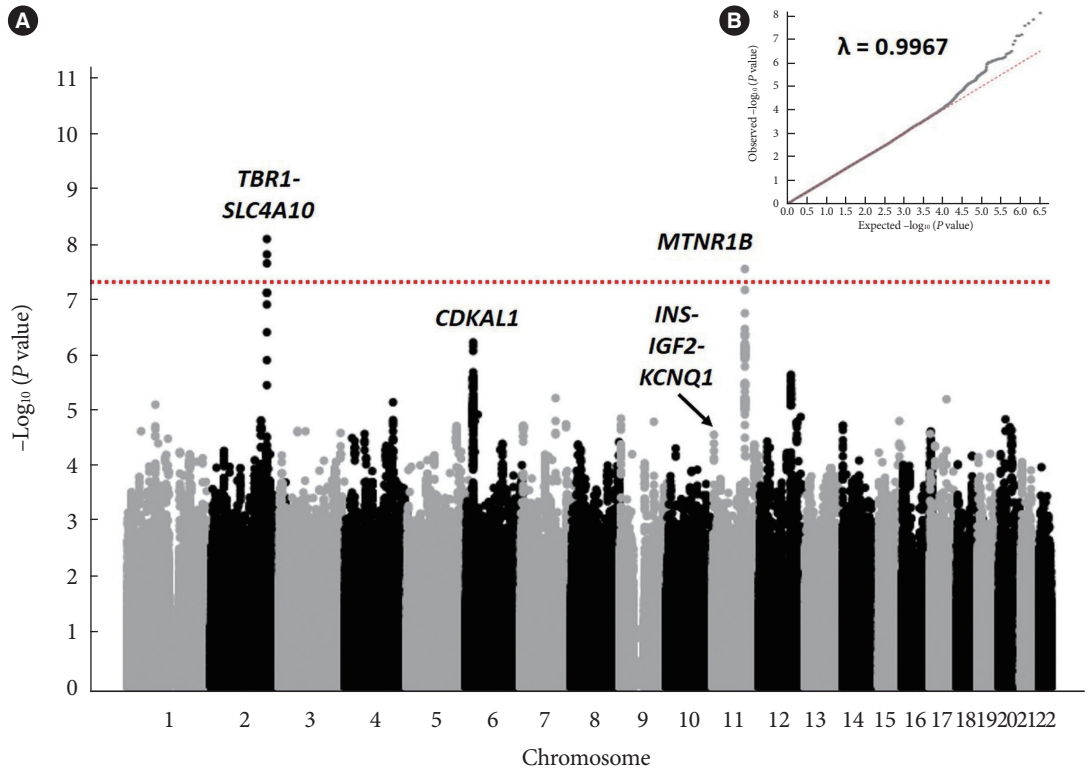


Fig. 1. Results for meta-analysis of genome-wide association study for gestational diabetes. (A) Manhattan plot. The y-axis represents the $-\log_{10}P$ value (adjusted for principal components and age), and the x-axis represents the 6,322,337 analyzed biallelic single nucleotide polymorphisms. The dashed red horizontal line corresponds to the genome-wide significance threshold for $P<5\times10^{-8}$. There are 4 points with $P<5\times10^{-8}$, and the arrow and labels localize the susceptibility loci to gestational diabetes mellitus (GDM) discovered in the present study. (B) Quantile-quantile (Q-Q) plot. The dotted line corresponds to the null hypothesis. *TBR1*, T-box brain transcription factor 1; *SLC4A10*, solute carrier family 4 member 10; *CDKAL1*, CDK5 regulatory subunit-associated protein 1-like 1; *MTNR1B*, melatonin receptor 1B; *INS-IGF2*, insulin-insulin-like growth factor 2; *KCNQ1*, potassium voltage-gated channel subfamily Q member 1.

GenDIP meta-analysis (OR, 1.31; 95% CI, 1.07 to 1.60; P_{meta}= 0.0164) (Fig. 2A and Supplementary Table 3). In contrast, the *MTNR1B* rs7945617 showed a *P* value ≤0.05 in all replication cohorts and demonstrated a smaller but concordant direction of effect in all replication cohorts (1.21< OR <1.34), compared to the discovery cohorts (Fig. 2B and Supplementary Table 4).

Candidate gene analysis

We utilized the meta-analysis results from this study to investigate the associations of established GDM-related loci in Chinese women (Table 1). In this analysis, we specifically chose 14 unique variants from nine loci that have shown associations with GDM at genome-wide significant levels in previous studies [8,9,18,19]. Of these, we confirmed the associations for four variants at the *CDKAL1* locus (rs7754840 and rs9348441) and the *MTNR1B* locus (rs10830962 and rs10830963) after correcting for multiple comparisons (*P* value threshold: 0.05/14= 3.6×10^{-3}). Our meta-analysis in Chinese women revealed a 1.33- to 1.58-fold ($1.5\times10^{-6}<P_{meta}<7.9\times10^{-4}$) increased risk of GDM per copy of the risk allele for these variants.

In our meta-analysis, we further investigated the impact of known T2DM-associated variants on the risk of GDM. We examined 338 variants that showed genome-wide significance in a multi-ethnic GWAS conducted by the Diabetes Meta-Analysis of Trans-Ethnic association studies (DIAMANTE) Consortium, which included 180,834 T2DM cases and 1,159,055 controls [20]. From these variants, we successfully obtained 286 independent variants at 216 loci in our meta-analysis (Supplementary Table 5). This analysis identified three variants, namely rs9348441 at the *CDKAL1* locus (OR, 1.40; 95% CI, 1.18 to

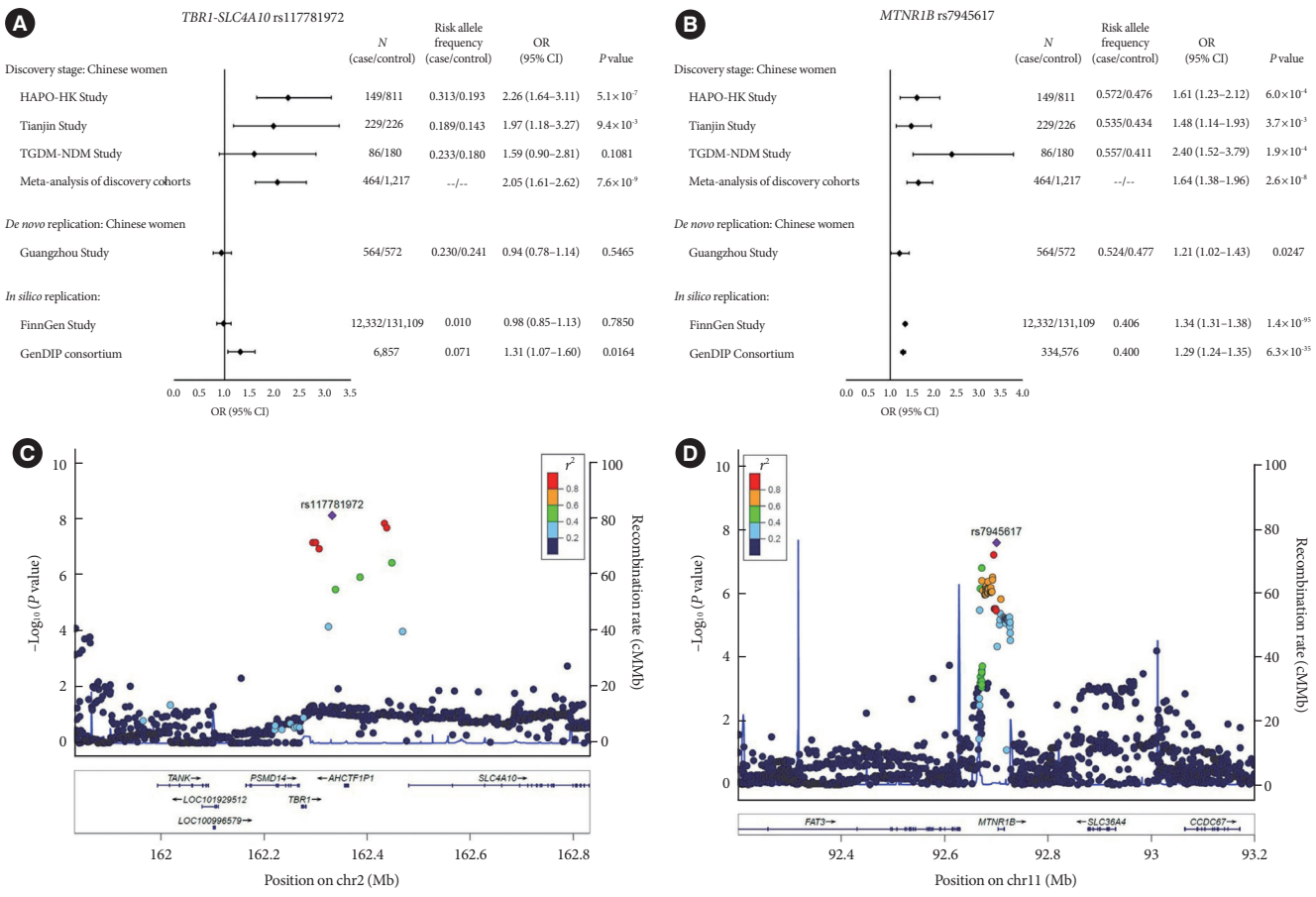


Fig. 2. Results for the two genome-wide significant loci for gestational diabetes. (A) Forest plot for the association between T-box brain transcription factor 1 (*TBR1*)-solute carrier family 4 member 10 (*SLC4A10*) rs117781972 and gestational diabetes mellitus (GDM) in all discovery and replication cohorts. Odds ratio (OR) and 95% confidence interval (CI) were reported according to the A-allele of rs117781972 (i.e., the GDM-associated risk allele). (B) Forest plot for the association between melatonin receptor 1B (*MTNR1B*) rs7945617 and GDM in all discovery and replication cohorts. ORs and 95% CIs were reported according to the C-allele of rs7945617 (i.e., the GDM-associated risk allele). A total of three studies (i.e., the Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong [HAPO-HK] Study, the Tianjin Study, and the Treated GDM Cases vs. Non-diabetes Controls [TGDM-NDM] Study) were included in the "meta-analysis of discovery cohorts." For the GENetics of Diabetes In Pregnancy (GenDIP) meta-analysis, the *P* values of the associations were obtained from the meta-regression implemented in Meta-Regression of Multi-AncEstry Genetic Association (MR-MEGA) and the combined OR and 95% CI was estimated by meta-analysis under a fixed effect model. (C) Regional plot of the *TBR1-SLC4A10* locus. (D) Regional plot of the *MTNR1B* locus. The purple diamonds represent the sentinel single nucleotide polymorphisms (SNPs) rs117781972 and rs7945617 identified from the meta-analysis of genome-wide association studies. Other SNPs are colored according to their level of linkage disequilibrium, which is measured by r^2, with the sentinel SNPs. The recombination rates estimated from the 1000 Genomes project Asian data are shown. The genes in the interval are indicated in the bottom panel. TANK, TRAF family member associated NFKB activator; PSMD14, proteasome 26S subunit, non-ATPase 14; AHCTF1P1, AT-hook containing transcription factor 1 pseudogene 1; FAT3, FAT atypical cadherin 3; CCDC67, coiled-coil domain containing 87.

1.65; $P_{meta}=9.4\times10^{-5}$), rs2237897 at the insulin-insulin-like growth factor 2 (*INS-IGF2*)-potassium voltage-gated channel subfamily Q member 1 (*KCNQ1*) locus (OR, 1.48; 95% CI, 1.23 to 1.79; $P_{meta}=3.8\times10^{-5}$), and rs10830963 at the *MTNR1B* locus (OR, 1.58; 95% CI, 1.31 to 1.91; $P_{meta}=1.5\times10^{-6}$), that remained significant after correcting for multiple comparisons (*P* value threshold: $0.05/286=1.7\times10^{-4}$).

Previous studies consistently found a strong association between GDM and the *CDKAL1* and *MTNR1B* loci, across diverse populations. To support the signals at the *INS-IGF2-KC-*

Table 1. Associations of established variants and gestational diabetes in Chinese women

Ref	Discovery population	Chr	SNP	Nearest genes(s)	Risk/non-risk allele	Cohort	Imputation quality, Rsq	Number		RAF		Association test		
								GDM case	Non-GDM control	GDM case	Non-GDM control	OR (95% CI)	$P_{additive}$	P_Q
[19]	Finnish	2	rs780094	*GCKR*	C/T	HAPO-HK Study	0.998	149	811	0.560	0.527	1.21 (0.93–1.57)	0.1493	-
						Tianjin Study	0.999	229	226	0.491	0.433	1.28 (0.98–1.68)	0.0736	-
						TGDM-NDM Study	0.997	86	180	0.536	0.564	0.97 (0.66–1.44)	0.8922	-
						Meta-analysis	-	464	1,217	-	-	1.19 (1.00–1.41)	0.0462	0.5160
[9]	Finnish	2	rs780093	*GCKR*	C/T	HAPO-HK Study	1.000	149	811	0.560	0.526	1.22 (0.94–1.58)	0.1397	-
						Tianjin Study	1.000	229	226	0.489	0.434	1.27 (0.97–1.66)	0.0883	-
						TGDM-NDM Study	1.000	86	180	0.535	0.564	0.97 (0.65–1.43)	0.8678	-
						Meta-analysis	-	464	1,217	-	-	1.18 (1.00–1.40)	0.0511	0.5193
[9]	Finnish	2	rs1402837	*SPC25*	T/C	HAPO-HK Study	0.999	149	811	0.406	0.386	1.03 (0.79–1.35)	0.8012	-
						Tianjin Study	1.000	229	226	0.389	0.369	1.11 (0.85–1.45)	0.4487	-
						TGDM-NDM Study	0.999	86	180	0.384	0.342	1.31 (0.86–1.99)	0.2120	-
						Meta-analysis	-	464	1,217	-	-	1.11 (0.93–1.32)	0.2453	0.6553
[9]	Finnish	5	rs1820176	*PCSK1*	T/C	HAPO-HK Study	0.978	149	811	0.663	0.645	1.06 (0.81–1.38)	0.6835	
						Tianjin Study	0.975	229	226	0.701	0.635	1.32 (1.01–1.74)	0.0456	-
						TGDM-NDM Study	0.982	86	180	0.699	0.704	0.95 (0.60–1.50)	0.8133	-
						Meta-analysis	-	464	1,217	-	-	1.14 (0.96–1.36)	0.1434	0.3541
[18]	Korean	6	rs7754840[a]	*CDKAL1*	C/G	HAPO-HK Study	0.989	149	811	0.451	0.361	1.47 (1.14–1.90)	3.4×10^{-3}	-
						Tianjin Study	1.000	229	226	0.465	0.381	1.41 (1.08–1.84)	0.0110	-
						TGDM-NDM Study	0.983	86	180	0.353	0.346	0.92 (0.61–1.38)	0.6836	-
						Meta-analysis	-	464	1,217	-	-	1.33 (1.13–1.58)	7.9×10^{-4}	0.1400
[8]	Multi-ethnicities	6	rs9348441[a]	*CDKAL1*	A/T	HAPO-HK Study	0.989	149	811	0.445	0.341	1.59 (1.22–2.06)	4.9×10^{-4}	-
						Tianjin Study	0.999	229	226	0.445	0.376	1.34 (1.02–1.75)	0.0340	-
						TGDM-NDM Study	0.985	86	180	0.403	0.341	1.16 (0.79–1.71)	0.4565	-
						Meta-analysis	-	464	1,217	-	-	1.40 (1.18–1.65)	9.4×10^{-5}	0.3808
[8]	Multi-ethnicities	9	rs10811662	*CDKN2A/ CDKN2B*	G/A	HAPO-HK Study	0.883	149	811	0.586	0.601	0.89 (0.68–1.17)	0.4150	-
						Tianjin Study	0.991	229	226	0.522	0.505	1.08 (0.83–1.39)	0.5703	-
						TGDM-NDM Study	0.869	86	180	0.660	0.564	1.84 (1.17–2.90)	0.0084	-
						Meta-analysis	-	464	1,217	-	-	1.08 (0.91–1.29)	0.3789	0.0278
[9]	Finnish	9	rs1333051	*CDKN2B*	A/T	HAPO-HK Study	0.873	149	811	0.885	0.872	1.07 (0.71–1.61)	0.7449	-
						Tianjin Study	0.996	229	226	0.830	0.843	0.93 (0.65–1.31)	0.6688	-
						TGDM-NDM Study	0.867	86	180	0.901	0.861	1.57 (0.77–3.19)	0.2161	-
						Meta-analysis	-	464	1,217	-	-	1.04 (0.81–1.33)	0.7458	0.4245

(*Continued to the next page*)

Table 1. Continued

Ref	Discovery population	Chr	SNP	Nearest genes(s)	Risk/non-risk allele	Cohort	Imputation quality, Rsq	Number		RAF		Association test		
								GDM case	Non-GDM control	GDM case	Non-GDM control	OR (95% CI)	$P_{additive}$	P_Q
[8]	Multi-ethnicities	10	rs9663238	*HKDC1*	G/A	HAPO-HK Study	0.990	149	811	0.294	0.274	1.11 (0.84–1.47)	0.4693	-
						Tianjin Study	0.996	229	226	0.260	0.292	0.84 (0.62–1.12)	0.2343	-
						TGDM-NDM Study	0.980	86	180	0.339	0.271	1.40 (0.87–2.24)	0.1646	-
						Meta-analysis	-	464	1,217	-	-	1.03 (0.85–1.24)	0.7825	0.1501
[9]	Finnish	10	rs34872471	*TCF7L2*	C/T	HAPO-HK Study	0.993	149	811	0.020	0.022	1.01 (0.41–2.50)	0.9767	-
						Tianjin Study	0.993	229	226	0.048	0.049	0.99 (0.53–1.86)	0.9782	-
						TGDM-NDM Study	0.990	86	180	0.052	0.030	1.10 (0.39–3.13)	0.8543	-
						Meta-analysis	-	464	1,217	-	-	1.02 (0.64–1.62)	0.9391	0.9854
[8]	Multi-ethnicities	10	rs7903146	*TCF7L2*	T/C	HAPO-HK Study	1.000	149	811	0.020	0.022	1.01 (0.41–2.48)	0.9823	-
						Tianjin Study	1.000	229	226	0.048	0.049	0.99 (0.53–1.86)	0.9804	-
						TGDM-NDM Study	1.000	86	180	0.052	0.031	1.10 (0.39–3.11)	0.8533	-
						Meta-analysis	-	464	1,217	-	-	1.02 (0.64–1.62)	0.9395	0.9853
[18]	Korean	11	rs10830962[a]	*MTNR1B*	G/C	HAPO-HK Study	0.879	149	811	0.522	0.439	1.52 (1.16–2.00)	2.4×10^{-3}	-
						Tianjin Study	0.995	229	226	0.505	0.421	1.36 (1.05–1.76)	0.0180	-
						TGDM-NDM Study	0.874	86	180	0.504	0.384	2.01 (1.28–3.16)	2.6×10^{-3}	-
						Meta-analysis	-	464	1,217	-	-	1.51 (1.27–1.79)	3.1×10^{-6}	0.3381
[8]	Multi-ethnicities	11	rs10830963[a]	*MTNR1B*	G/C	HAPO-HK Study	0.775	149	811	0.498	0.426	1.51 (1.13–2.01)	5.3×10^{-3}	-
						Tianjin Study	0.812	229	226	0.478	0.393	1.47 (1.11–1.96)	7.9×10^{-3}	-
						TGDM-NDM Study	0.779	86	180	0.486	0.357	2.22 (1.37–3.60)	1.2×10^{-3}	-
						Meta-analysis	-	464	1,217	-	-	1.58 (1.31–1.91)	1.5×10^{-6}	0.3257
[9]	Finnish	16	rs2926003	*CMIP*	C/T	HAPO-HK Study	0.982	149	811	0.666	0.718	0.75 (0.57–1.00)	0.0511	-
						Tianjin Study	0.824	229	226	0.659	0.673	0.93 (0.68–1.27)	0.6588	-
						TGDM-NDM Study	0.987	86	180	0.733	0.735	0.95 (0.58–1.54)	0.8255	-
						Meta-analysis	-	464	1,217	-	-	0.85 (0.70–1.03)	0.0919	0.5472

ORs were estimated according to the reported risk allele. $P_{additive}$ was obtained from individual cohort using logistic regression model with the adjustments of age and principal components. P_Q was obtained from heterogeneity test (Cochran's Q test).

Chr, chromosome; SNP, single nucleotide polymorphism; RAF, risk allele frequency; GDM, gestational diabetes mellitus; OR, odds ratio; CI, confidence interval; *GCKR*, glucokinase regulator; HAPO-HK, Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong; TGDM-NDM, Treated GDM Cases vs. Non-diabetes Controls; *SPC25*, SPC25 component of NDC80 kinetochore complex; *PCSK1*, proprotein convertase subtilisin/kexin type 1; *CDKAL1*, CDK5 regulatory subunit-associated protein 1-like 1; *CDKN2A*, cyclin dependent kinase inhibitor 2A; *HKDC1*, hexokinase domain containing 1; *TCF7L2*, transcription factor 7 like 2; *MTNR1B*, melatonin receptor 1B; *CMIP*, c-Maf inducing protein.

[a]SNP remained significant after considering multiple comparisons (P value threshold for 14 tests: $0.05/14=3.6\times10^{-3}$).

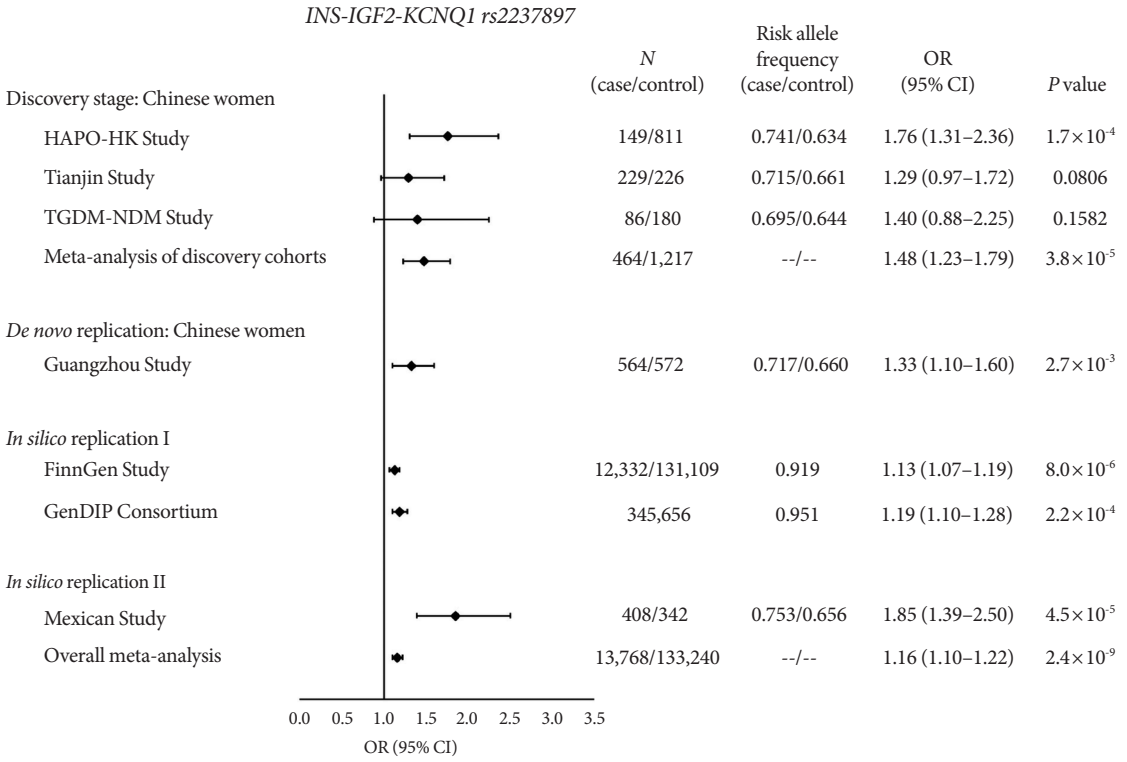


Fig. 3. Forest plot for the association between potassium voltage-gated channel subfamily Q member 1 (*KCNQ1*) rs2237897 and gestational diabetes mellitus in all discovery and replication cohorts. Odds ratio (OR) and 95% confidence interval (CI) were reported according to the C-allele of rs2237897 (i.e., the type 2 diabetes mellitus-associated risk allele). For the GENetics of Diabetes In Pregnancy (GenDIP) meta-analysis, the *P* value of the association was obtained from the meta-regression implemented in Meta-Regression of Multi-AncEstry Genetic Association (MR-MEGA) and the combined OR and 95% CI was estimated by meta-analysis under a fixed effect model. A total of three studies (i.e., the Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong [HAPO-HK] Study, the Tianjin Study, and the Tianjin Study and the Treated GDM Cases vs. Non-diabetes Controls [TGDM-NDM] Study) were included in the "meta-analysis of discovery cohorts." A total of six studies (i.e., the HAPO-HK Study, the Tianjin Study, the TGDM-NDM Study, the Guangzhou Study, the FinnGen Study, and the Mexican Study) were included in the "overall meta-analysis." We did not include the GenDIP samples in the overall meta-analysis because they overlapped with both the HAPO-HK and FinnGen Studies.

NQ1 locus, we conducted independent replication studies in the Guangzhou Study, the FinnGen Study and the GenDIP meta-analysis. Our analysis of *INS-IGF2-KCNQ1* rs2237897 yielded *P* values ranging from 8.0×10^{-6} to 2.7×10^{-3}, with consistent ORs of 1.13 to 1.33 in all these studies, aligning with the original findings (Fig. 3 and Supplementary Table 6). Moreover, a previous study conducted in the Mexican population (408 cases and 342 controls) reported a positive association between rs2237897 and GDM (OR, 1.85; 95% CI, 1.39 to 2.50; $P=4.5\times10^{-5}$) (Fig. 3 and Supplementary Table 6) [21]. We further performed a meta-analysis for rs2237897 using data from the HAPO-HK Study, the Tianjin Study, the TGDM-NDM Study, the Guangzhou Study, the FinnGen Study, and the Mexican Study, comprising a total of 13,768 GDM cases and 133,240 controls. In this meta-analysis, the rs2237897 variant exhibited a genome-wide significant association with GDM risk (OR, 1.16; 95% CI, 1.10 to 1.22; $P_{meta}=2.4\times10^{-9}$) (Fig. 3 and Supplementary Table 6).

Conditional analysis at the *CDKAL1* and *MTNR1B* loci

Supplementary Fig. 4 summarizes the GDM variants identified in this study. Through the genome-wide scanning and candidate gene approaches, we found two variants at the *CDKAL1* locus (rs7754840 and rs9348441) and three variants at the *MTNR1B* locus (rs7945617, rs10830962, and rs10830963). These variants showed high linkage disequilibrium (LD) within each locus in our Chinese samples (LD $r^2=0.81$ for *CDKAL1* and 0.78–0.93 for *MTNR1B*) (Supplementary Table 7). Conditional

analysis revealed that the effect sizes of variants within the two loci were attenuated when conditioned on each other (Supplementary Table 7). As a result, we selected four variants for further analysis: *CDKAL1* rs7754840, *MTNR1B* rs7945617, *TBR1-SLC4A10* rs117781972, and *INS-IGF2-KCNQ1* rs2237897.

Results for (1) sensitivity analysis and (2) associations of identified variants with glycemic and metabolic traits during pregnancy are available in the Supplementary Results.

Development and evaluation of PRS for GDM

We constructed a PRS using the four identified variants, weighted by their corresponding effect sizes for GDM risk, as

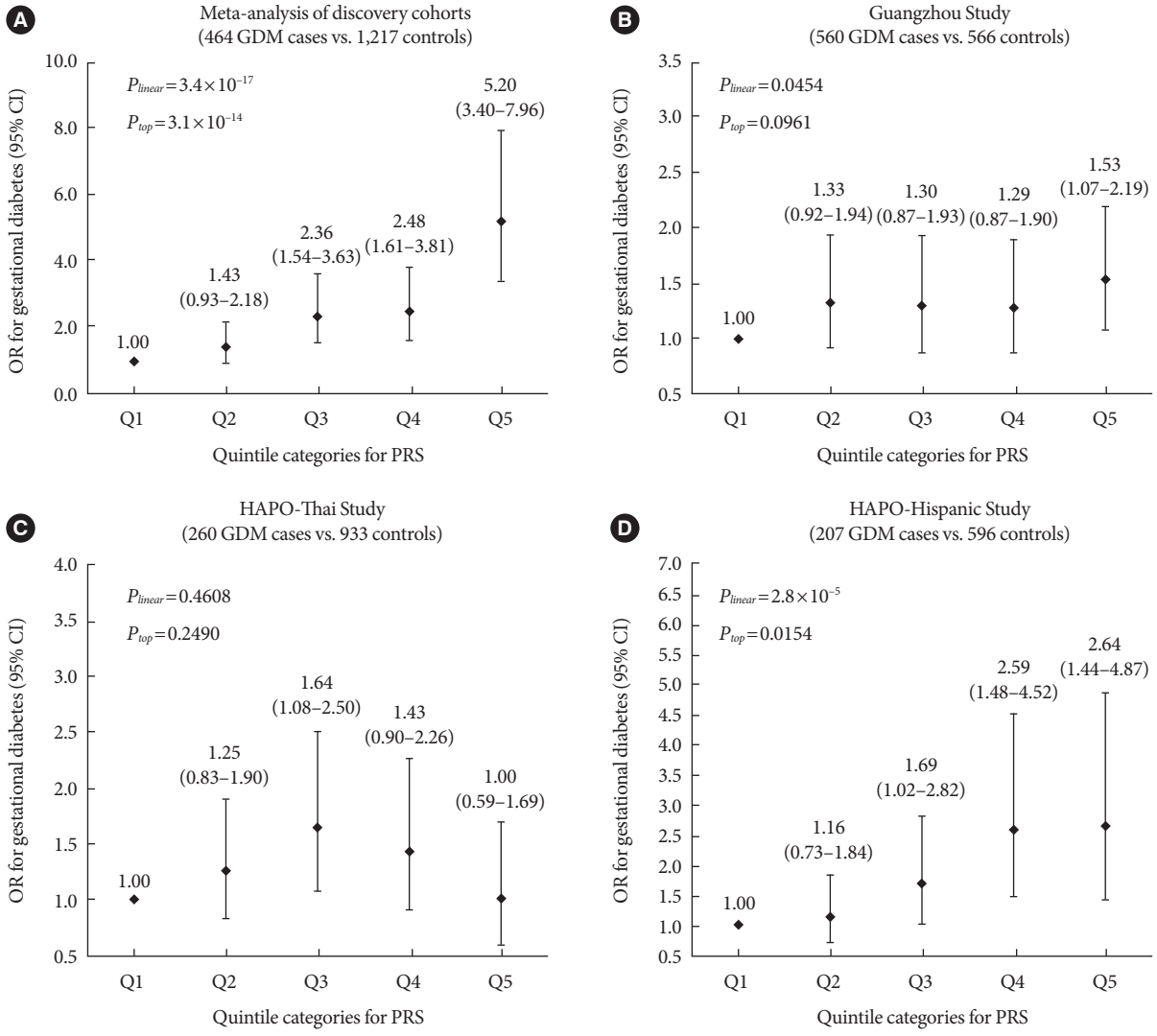


Fig. 4. Association between quintiles of polygenic risk score (PRS) and gestational diabetes. (A) Meta-analysis of three discovery cohorts of Chinese women (the Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong [HAPO-HK] Study, Tianjin Study, and Tianjin Study and the Treated GDM Cases vs. Non-diabetes Controls [TGDM-NDM] Study). (B) Guangzhou Study. (C) HAPO-Thai Study. (D) HAPO-Hispanic Study. P_{linear} is the *P* value testing for a linear trend across the quintile categories of PRS. P_{top} is the *P* value testing for the association of a high PRS with gestational diabetes mellitus (GDM) by comparing the top 20% with the remaining 80% of the PRS distribution. Odds ratio (OR) and 95% confidence interval (CI) of GDM were stratified by quintile categories of PRS. Within each individual cohort, *P* values were obtained from logistic regression with the adjustment of principal components, age and body mass index, except for the Guangzhou Study which did not adjust for any covariates. Results from the three discovery cohorts were then meta-analyzed using a fixed-effects model.

estimated in our meta-analysis. This PRS demonstrated robust association with increased GDM risk among the three discovery cohorts of Chinese women ($1.2\times10^{-13}<P<1.0\times10^{-3}$) (Supplementary Table 8). Our meta-analysis revealed a 1.84-fold (95% CI, 1.62 to 2.09; $P_{meta}=9.2\times10^{-21}$) increase in GDM risk per 1-standard deviation (SD) increase in the PRS. Each quintile increase in the PRS was associated with a 1.48-fold (95% CI, 1.35 to 1.62; $P_{meta}=3.4\times10^{-17}$) increased odds of GDM (Fig. 4A). Women in the top quintile had a three-fold (OR, 3.07; 95% CI, 2.30 to 4.09; $P_{meta}=3.1\times10^{-14}$) and more than five-fold (OR, 5.20; 95% CI, 3.40 to 7.96; $P_{meta}=3.0\times10^{-14}$) increased risk of GDM, compared with the remaining individuals and those in the bottom quintile, respectively (Fig. 4A and Supplementary Table 8). The association between PRS and GDM was further validated in the Guangzhou Study (OR, 1.16; 95% CI, 1.03 to 1.31; $P=0.0133$) and the HAPO-Hispanic Study (OR, 1.41; 95% CI, 1.18 to 1.67; $P=1.2\times10^{-4}$) (Fig. 4B, C and Supplementary Table 8). However, no significant association was seen in the HAPO-Thai Study ($P=0.2695$) (Fig. 4D and Supplementary Table 8).

Next, we assessed the predictive value of our PRS for GDM (Supplementary Table 9). In the model incorporating only the PRS, the AUC for predicting GDM varied between 0.63 and 0.69 in the three discovery cohorts; but it was below 0.57 in all the validation cohorts. When PRS was added to the base model consisting of PCs, age and BMI, there was a significant increase in the AUC by 5.0% to 7.8% in the HAPO-HK and Tianjin Studies ($1.4\times10^{-4}<P_{increase}<0.0134$). Computation of NRI index indicated enhanced reclassification in all discovery and HAPO-Hispanic Studies (25.8%< NRI <57.4%; $P_{NRI}<0.05$). Similar findings were observed when additional clinical risk factors were considered in the base model of both the HAPO-HK and HAPO-Hispanic Studies.

Evaluation of PRS for AGT after pregnancy

Given the well-established association between GDM and post-pregnancy hyperglycemia, we evaluated the role of our GDM-related PRS in predicting AGT risk at 7-year postpartum in the HAPO-HK Study (129 AGT cases and 829 controls). The clinical characteristics of these participants during the follow-up

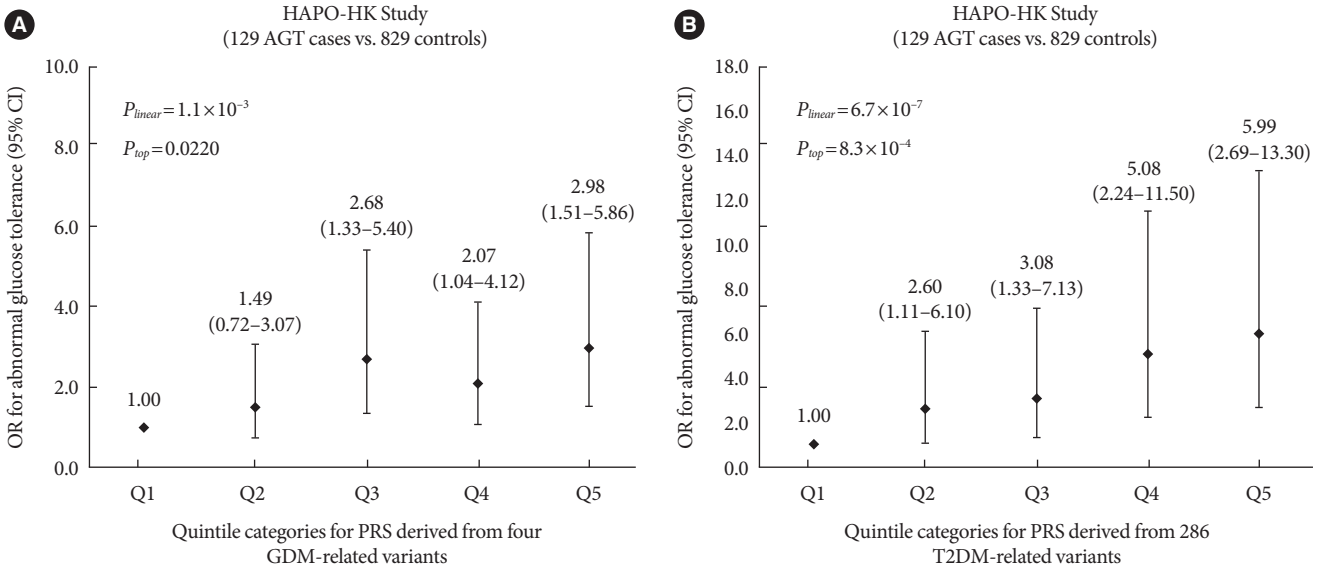


Fig. 5. Association between quintiles of polygenic risk score (PRS) and abnormal glucose tolerance (AGT) at 7-year postpartum in the Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong (HAPO-HK) Study. (A) PRS derived based on four gestational diabetes mellitus (GDM)-related variants. (B) PRS derived based on 286 type 2 diabetes mellitus (T2DM)-related variants. The T2DM-related PRS was derived based on 286 T2DM-related variants reported by the Diabetes Meta-Analysis of Trans-Ethnic association studies (DIAMANTE) consortium [20]. P_{linear} is the *P* value testing for a linear trend across the quintile categories of PRS. P_{top} is the *P* value testing for the association of a high PRS with AGT after pregnancy by comparing the top 20% with the remaining 80% of the PRS distribution. Odds ratio (OR) and 95% confidence interval (CI) of GDM were stratified by quintile categories of PRS. *P* values were obtained from logistic regression with the adjustment of principal components, age and body mass index.

study are summarized in Supplementary Table 10. With the adjustments for PCs, age and BMI, the PRS was associated with a 1.40-fold (95% CI, 1.15 to 1.71; $P=6.7\times10^{-4}$) increased risk of postpartum AGT per 1-SD increase in the PRS (Supplementary Table 11). Each quintile increase in this PRS raised the odds of AGT by approximately 27% (OR, 1.27; 95% CI, 1.10 to 1.46; $P=1.1\times10^{-3}$) (Fig. 5). Women with a high PRS (top quintile) had a 71% (OR, 1.71; 95% CI, 1.08 to 2.71; $P=0.0220$) increased risk of AGT compared to the remaining individuals (Fig. 5 and Supplementary Table 11). This risk was nearly 3-fold (OR, 2.98; 95% CI, 1.51 to 5.86; $P=1.6\times10^{-3}$) higher compared to women with a low PRS (bottom quintile). However, adjusting for either a history of GDM ($P=0.0571$) or a PRS derived from 286 T2DM-related variants ($P=0.0408$) attenuated the association between PRS and AGT (Supplementary Table 11).

The AUC for predicting postpartum AGT using only the PRS was 0.60 (95% CI, 0.54 to 0.65) (Supplementary Table 12). Incorporating the PRS into the base model, which included PCs, age and BMI, significantly increased the AUC by 1.7% ($P=0.0187$) and improved risk reclassification for AGT, as evidenced by an NRI index of 31.7% (95% CI, 12.6% to 49.6%). When either the GDM history or the T2DM-related PRS was included in the base model, the addition of the PRS did not impact the AUC. However, a significant NRI index was still observed, suggesting an improvement in reclassification ($25.5\% < \text{NRI} < 25.6\%$; $P<0.05$).

DISCUSSION

This study reported a meta-analysis of three GWASs for GDM in Chinese women, with subsequent replication studies in Chinese and multi-ethnic populations. We identified four susceptibility loci at *TBR1-SLC4A10*, *MTNR1B*, *CDKAL1*, and *INS-IGF2-KCNQ1,* using both genome-wide scan and candidate gene approaches. Moreover, the PRS derived from these loci showed improved risk stratification for Chinese women who are at risk of developing GDM or AGT postpartum.

Three out of the four GDM-associated loci identified in this study have also been implicated in glycemic traits and the development of T2DM in non-gravid populations. Aligning with previous GWASs in Korean [18], Finnish [9], and multi-ethnic populations [8], as well as candidate gene studies across different populations [21,22], our findings further support the associations of *CDKAL1*, *MTNR1B*, and *KCNQ1* loci with GDM risk, with comparable effect sizes estimated in Chinese pregnant women. To the best of our knowledge, this is the first study showing a genome-wide significant association between the *KCNQ1* variant and GDM risk. These three genes have been found to impact insulin secretion and glucose homeostasis, contributing to the development of diabetes [23-25]. Although the role of *CDKAL1* in pancreatic β-cell function remains incompletely understood, it is known to be expressed in human adult pancreatic islets [26]. In mice, *CDKAL1* gene deletion has been observed to cause a modest impairment in insulin secretion during high-fat feeding [27]. Moreover, emerging evidence indicates that functional loss of *CDKAL1* affects the accuracy of protein translation, resulting in abnormal insulin synthesis and subsequently triggering endoplasmic reticulum stress in β-cells [28]. These findings provide evidence linking the molecular function of *CDKAL1* with GDM. *MTNR1B* plays a crucial role in regulating circadian rhythm. Disruptions in circadian rhythms have been observed to promote systemic metabolic dysfunction, including obesity or diabetes, in both humans and rodents [29]. Consequently, altered *MTNR1B* function may influence circadian rhythm regulation, potentially affecting glucose metabolism during pregnancy. The *KCNQ1* gene encodes for a voltage-gated potassium channel (Kv7.1) that is expressed in multiple organs/tissues, including pancreas [30]. In the pancreas, this channel drives an electrical signal in the β-cells that facilitates glucose-stimulated insulin secretion [30]. Changes of *KCNQ1* function may impact the proper functioning of pancreatic β-cells, leading to impaired insulin secretion and glucose regulation in GDM women. Overall, our findings suggest shared genetic determinants between GDM and T2DM. The lack of association observed between other known T2DM-related loci and GDM may indicate partial differences in the genetic architecture of glucose metabolism between the pregnant and non-pregnant states, and/or limited power in the current study.

While many previous studies have identified susceptibility loci associated with both GDM and T2DM, two recent large-scale GWASs involving 143K–353K participants, predominantly of European descent, have uncovered a total of nine novel loci (i.e., *HKDC1*, glucokinase regulator [*GCKR*], SPC25 component of NDC80 kinetochore complex [*SPC25*], adenylate cyclase 5 [*ADCY5*], proprotein convertase subtilisin/kexin type 1 [*PCSK1*], estrogen receptor 1 [*ESR1*], cyclin D2 [*CCND2*], NEDD1 gamma-tubulin ring complex targeting factor [*NEDD1*], c-Maf inducing protein [*CMIP*], and mitogen-activated protein kinase kinase kinase 15 [*MAP3K15*]) specifically linked to

glucose regulation during pregnancy [8,9]. These findings motivate further investigation into the distinct genetic mechanisms underlying GDM compared to T2DM, particularly within non-European populations. In this study, we have discovered a novel common variant rs117781972 located between the *TBR1* and *SLC4A10* gene regions in Chinese population. This variant has not been reported in GWASs for T2DM or glycemic traits. We found that the GDM-associated risk allele of rs117781972 was also associated with higher postprandial glucose levels (i.e., 1-hour glucose and 2-hour glucose) during pregnancy. Interestingly, inconsistent associations have been observed among replication studies, which may be due to differences in genetic background. For instance, the allele frequency of rs117781972 is significantly higher in East Asian populations (minor allele frequency [MAF]=0.227) compared to European (MAF=0.002 to 0.014), South Asian (MAF=0.01), and African (MAF=0.003) populations (https://gnomad.broadinstitute.org/). This genetic heterogeneity can be further modified by differences in phenotypes, lifestyles, and treatments. Our study, along with other research, suggests that while metabolic changes during pregnancy may resemble those observed in the progression of T2DM, and despite the shared genetic risk factors between T2DM and GDM, the underlying mechanisms driving the development of hyperglycemia during pregnancy may have some important differences.

The product of *TBR1* is vital for the development and functioning of brain regions, including cerebral cortex, hippocampus, and amygdala [31,32]. *SLC4A10*, a member of the sodium-coupled bicarbonate transporters (NCBTs) family, regulates neuronal pH and excitability, and the secretion of cerebrospinal fluid [33]. Both genes are primarily expressed in the brain (data source: GTEx Analysis Release V8). *TBR1* protein and multiple variants at *SLC4A10* (rs4500960, rs7604885, and rs7580486) have demonstrated associations with major neurological disorders such as autism spectrum disorder (ASD) [32], cognitive ability [34], schizophrenia [35], and intelligence [36]. In line with the observational data linking exposure to maternal diabetes *in utero* with ASD risk [37], the effects of *TRB1-SLC4A10* locus on GDM may also influence the fetal brain development, potentially leading to ASD after birth. However, little is known about the GDM-associated locus at these genes. A few recent GWASs have linked common variants at the *TBR1-SLC4A10* locus (rs55732192, rs62189012, rs2053865, rs77882688, and rs1449631) to cardiometabolic traits, including BMI, mean arterial pressure, and systolic blood pressure [38-41]. Our identified variant rs117781972, strongly associated with GDM, appears to be independent of these reported variants ($r^2<0.1$ in East Asian population). Furthermore, our data did not show any association between rs117781972 and blood pressure or BMI during pregnancy (Supplementary Table 13). To pinpoint the relevant gene and functional variant(s), further fine-mapping in larger multi-ancestry cohorts may be required.

Similar to T2DM, the risk of GDM may arise from variants in multiple loci, each conferring modest individual effects. Therefore, we computed a PRS using variants identified in this study to evaluate their cumulative impact on predicting an individual's risk of developing hyperglycemia during and after pregnancy. Previous studies in Caucasian, Chinese and Indian populations have examined the association between PRS and GDM risk, using PRSs involving six to 14,971,357 variants [11,42-45]. However, there is limited data available for evaluating the performance of PRS in risk prediction [11,42,46]. Most of these studies derived PRSs using common variants associated with T2DM or glycemic traits, and tested associations in small sample sizes (n=296 to 8,722) [10]. Consistent with previous observations [11,42,46], our PRS, consisting of four variants, showed a significant association with GDM in Chinese and Hispanic populations. Our analysis further demonstrated that this PRS modestly contributed to GDM prediction when combined with well-recognized risk factors, although its overall value in identifying GDM cases at the population level was limited. It is worth noting that this PRS effectively stratified the risk of GDM in Chinese women, independent of clinical risk factors such as age and BMI. We found that women with a high PRS (top 20% of the distribution) had over a three-fold increased GDM risk compared to their counterparts with a lower PRS (remaining 80% of the distribution). More importantly, our PRS improved risk prediction for post-pregnancy hyperglycemia, showing a 71% higher risk among women with a high PRS compared to those with a lower PRS. This relationship was attenuated when adjusting for GDM history, further supporting the causal link between GDM and future T2DM progression. Overall, our findings underscore the importance of the polygenic contributions to GDM, and demonstrated the potential utility of genetic information in clinical practice, such as early genomic screening and personalized risk assessment for pregnant women.

This study has several limitations. First, the sample size for our genome-wide scan analysis was relatively modest, limiting our ability to discover new loci and detect variants with small

genetic effects. Additionally, the HAPO Study have excluded women with severe GDM, further reducing the study's power. Nevertheless, we were able to identify a novel locus associated with GDM within our cohorts. Second, our discovery study was exclusively conducted among Han Chinese women only, with limited generalizability. Third, the glucose tolerance status of controls in the TGDM-NDM Study was not assessed during pregnancy, potentially leading to the inclusion of some women with GDM as controls. Fourth, the predictive value of our PRS may be significantly diminished due to the inclusion of only a small number of genetic variants. Fifth, there is a possibility of over-fitting when evaluating the PRS in the discovery cohort used in the GWAS; however, it has been validated in independent cohorts. Furthermore, the impact of genetic variants may vary between GDM and AGT, and certain loci may be unique to each disease. We anticipate that constructing our PRS using weights from larger GWASs will enhance its predictive accuracy.

In summary, this study identified associations of four genetic loci with GDM among Chinese women. The *MTNR1B*, *CDKAL1*, and *KCNQ1* loci, known for their role in hyperglycemia in non-pregnant populations, also have a significant impact during pregnancy, suggesting shared underlying pathology between T2DM and GDM. The *TBR1-SLC4A10* locus, on the other hand, showed no association with hyperglycemia in non-pregnant populations, highlighting its specific importance in glucose metabolism during pregnancy. Finally, incorporating GDM-related PRS with clinical risk factors enhanced the prediction of GDM and post-pregnancy AGT. This demonstrates the clinical potential of integrating genetic information into risk assessment tools for identifying pregnant women at risk of developing these diseases, thereby providing opportunities to delaying disease onset.

SUPPLEMENTARY MATERIALS

Supplementary materials related to this article can be found online at https://doi.org/10.4093/dmj.2024.0139.

CONFLICTS OF INTEREST

Cadmon King-poo Lim, Juliana Chung-ngor Chan, and Ronald Ching-wan Ma are co-founders of GemVCare, a technology start-up initiated with support from the Hong Kong Government Innovation and Technology Commission and its Technology Start-up Support Scheme for Universities (TSSSU). The other authors declare that there is no duality of interest associated with this manuscript. Ronald Ching-wan Ma is a member of the international editorial board of *Diabetes & Metabolism Journal*.

AUTHOR CONTRIBUTIONS

Conception or design: C.H.T., Y.W., C.C.W., J.C.C., W.H.T., X.Y., R.C.M.
Acquisition, analysis, or interpretation of data: all authors.
Drafting the work or revising: C.H.T., C.C.W., X.Y., R.C.M.
Final approval of the manuscript: all authors.

ORCID

Claudia Ha-ting Tam *https://orcid.org/0000-0002-9169-0013*
Xilin Yang *https://orcid.org/0000-0002-9462-7992*
Ronald Ching-wan Ma *https://orcid.org/0000-0002-1227-803X*

FUNDING

This work was funded by the RGC Theme-based Research Scheme (T12-402/13 N); the Research Impact Fund (R4012-18), and the University Grants Committee Research Grants Matching Scheme. The follow-up of the HAPO Study at Hong Kong field center was supported by the General Research Fund of the Research Grants Council of the Hong Kong SAR, China (CUHK 473408, 471713, 14118718, 14102719). The HAPO Study was funded by the National Institute of Child Health and Human Development (grant no. R01-HD34242) and the National Institute of Diabetes and Digestive and Kidney Diseases (grant no. R01-HD34243). The funding sources do not have any role in the design, interpretation of the study, or the decision to publish the results.

ACKNOWLEDGMENTS

Ronald Ching-wan Ma and Wing Hung Tam are the principal investigators of the HAPO-HK Study and TGDM-NDM Study; Ying Wang and Xilin Yang are the principal investigators of the Tianjin Study.

We thank the HAPO Study steering committee for initiating and conducting the original study, and for their kind help and support. We are also grateful to all the study participants for their contribution. Special thanks are extended to all medical

and nursing staff at all participating centers for their dedication and professionalism. We thank all team members for their kind assistance, and their efforts on the recruitment of patients and data collection. We acknowledge the participants and investigators of the GenDIP Consortium for making their data available for analysis.

REFERENCES

1. Wang H, Li N, Chivese T, Werfalli M, Sun H, Yuen L, et al. IDF Diabetes Atlas: estimation of global and regional gestational diabetes mellitus prevalence for 2021 by International Association of Diabetes in Pregnancy Study Group's criteria. Diabetes Res Clin Pract 2022;183:109050.
2. HAPO Study Cooperative Research Group; Metzger BE, Lowe LP, Dyer AR, Trimble ER, Chaovarindr U, et al. Hyperglycemia and adverse pregnancy outcomes. N Engl J Med 2008;358:1991-2002.
3. Vounzoulaki E, Khunti K, Abner SC, Tan BK, Davies MJ, Gillies CL. Progression to type 2 diabetes in women with a known history of gestational diabetes: systematic review and meta-analysis. BMJ 2020;369:m1361.
4. Lowe WL Jr, Lowe LP, Kuang A, Catalano PM, Nodzenski M, Talbot O, et al. Maternal glucose levels during pregnancy and childhood adiposity in the Hyperglycemia and Adverse Pregnancy Outcome Follow-up Study. Diabetologia 2019;62:598-610.
5. Lowe WL Jr, Scholtens DM, Kuang A, Linder B, Lawrence JM, Lebenthal Y, et al. Hyperglycemia and Adverse Pregnancy Outcome Follow-up Study (HAPO FUS): maternal gestational diabetes mellitus and childhood glucose metabolism. Diabetes Care 2019;42:372-80.
6. Tam WH, Ma RC, Ozaki R, Li AM, Chan MH, Yuen LY, et al. In utero exposure to maternal hyperglycemia increases childhood cardiometabolic risk in offspring. Diabetes Care 2017;40:679-86.
7. Powe CE, Kwak SH. Genetic studies of gestational diabetes and glucose metabolism in pregnancy. Curr Diab Rep 2020;20:69.
8. Pervjakova N, Moen GH, Borges MC, Ferreira T, Cook JP, Allard C, et al. Multi-ancestry genome-wide association study of gestational diabetes mellitus highlights genetic links with type 2 diabetes. Hum Mol Genet 2022;31:3377-91.
9. Elliott A, Walters RK, Pirinen M, Kurki M, Junna N, Goldstein JI, et al. Distinct and shared genetic architectures of gestational diabetes mellitus and type 2 diabetes. Nat Genet 2024;56:377-82.
10. Tian Y, Li P. Genetic risk score to improve prediction and treatment in gestational diabetes mellitus. Front Endocrinol (Lausanne) 2022;13:955821.
11. Kawai VK, Levinson RT, Adefurin A, Kurnik D, Collier SP, Conway D, et al. A genetic risk score that includes common type 2 diabetes risk variants is associated with gestational diabetes. Clin Endocrinol (Oxf) 2017;87:149-55.
12. Leng J, Shao P, Zhang C, Tian H, Zhang F, Zhang S, 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.
13. Tam WH, Ma RCW, Ozaki R, Li AM, Chan MHM, Yuen LY, et al. Antenatal treatment of gestational diabetes and offspring's future cardiometabolic risk. The 9th International Symposium on Diabetes, Hypertension and Metabolic Syndrome and in Pregnancy; 2017 Mar 8-12; Barcelona, Spain.
14. Ko GT, Chan JC, Chan AW, Wong PT, Hui SS, Tong SD, et al. Association between sleeping hours, working hours and obesity in Hong Kong Chinese: the 'better health for better Hong Kong' health promotion campaign. Int J Obes (Lond) 2007;31:254-60.
15. Wu L, Song Y, Zhang Y, Liang B, Deng Y, Tang T, et al. Novel genetic variants of PPARγ2 promoter in gestational diabetes mellitus and its molecular regulation in adipogenesis. Front Endocrinol (Lausanne) 2021;11:499788.
16. International Association of Diabetes and Pregnancy Study Groups Consensus Panel; Metzger BE, Gabbe SG, Persson B, Buchanan TA, Catalano PA, 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.
17. ElSayed NA, Aleppo G, Aroda VR, Bannuru RR, Brown FM, Bruemmer D, et al. 2. Classification and diagnosis of diabetes: standards of care in diabetes-2023. Diabetes Care 2023;46(Suppl 1):S19-40.
18. Kwak SH, Kim SH, Cho YM, Go MJ, Cho YS, Choi SH, et al. A genome-wide association study of gestational diabetes mellitus in Korean women. Diabetes 2012;61:531-41.
19. Changalidis AI, Maksiutenko EM, Barbitoff YA, Tkachenko AA, Vashukova ES, Pachuliia OV, et al. Aggregation of genome-wide association data from FinnGen and UK Biobank replicates multiple risk loci for pregnancy complications. Genes (Basel) 2022;13:2255.
20. Mahajan A, Spracklen CN, Zhang W, Ng MC, Petty LE, Kitaji-

ma H, et al. Multi-ancestry genetic study of type 2 diabetes highlights the power of diverse populations for discovery and translation. Nat Genet 2022;54:560-72.
21. Huerta-Chagoya A, Vazquez-Cardenas P, Moreno-Macias H, Tapia-Maruri L, Rodriguez-Guillen R, Lopez-Vite E, et al. Genetic determinants for gestational diabetes mellitus and related metabolic traits in Mexican women. PLoS One 2015;10:e0126408.
22. Ao D, Wang HJ, Wang LF, Song JY, Yang HX, Wang Y. The rs2237892 polymorphism in KCNQ1 influences gestational diabetes mellitus and glucose levels: a case-control study and meta-analysis. PLoS One 2015;10:e0128901.
23. Yasuda K, Miyake K, Horikawa Y, Hara K, Osawa H, Furuta H, et al. Variants in KCNQ1 are associated with susceptibility to type 2 diabetes mellitus. Nat Genet 2008;40:1092-7.
24. Pascoe L, Tura A, Patel SK, Ibrahim IM, Ferrannini E, Zeggini E, et al. Common variants of the novel type 2 diabetes genes CDKAL1 and HHEX/IDE are associated with decreased pancreatic beta-cell function. Diabetes 2007;56:3101-4.
25. Lyssenko V, Nagorny CL, Erdos MR, Wierup N, Jonsson A, Spegel P, et al. Common variant in MTNR1B associated with increased risk of type 2 diabetes and impaired early insulin secretion. Nat Genet 2009;41:82-8.
26. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science 2007;316:1336-41.
27. Okamura T, Yanobu-Takanashi R, Takeuchi F, Isono M, Akiyama K, Shimizu Y, et al. Deletion of CDKAL1 affects high-fat diet-induced fat accumulation and glucose-stimulated insulin secretion in mice, indicating relevance to diabetes. PLoS One 2012;7:e49055.
28. Wei FY, Suzuki T, Watanabe S, Kimura S, Kaitsuka T, Fujimura A, et al. Deficit of tRNA(Lys) modification by Cdkal1 causes the development of type 2 diabetes in mice. J Clin Invest 2011;121:3598-608.
29. Shimizu I, Yoshida Y, Minamino T. A role for circadian clock in metabolic disease. Hypertens Res 2016;39:483-91.
30. Yamagata K, Senokuchi T, Lu M, Takemoto M, Fazlul Karim M, Go C, et al. Voltage-gated K+ channel KCNQ1 regulates insulin secretion in MIN6 β-cell line. Biochem Biophys Res Commun 2011;407:620-5.
31. Deriziotis P, O'Roak BJ, Graham SA, Estruch SB, Dimitropoulou D, Bernier RA, et al. De novo TBR1 mutations in sporadic autism disrupt protein functions. Nat Commun 2014;5:4954.
32. den Hoed J, Sollis E, Venselaar H, Estruch SB, Deriziotis P, Fisher SE. Functional characterization of TBR1 variants in neurodevelopmental disorder. Sci Rep 2018;8:14279.
33. Jacobs S, Ruusuvuori E, Sipila ST, Haapanen A, Damkier HH, Kurth I, et al. Mice with targeted Slc4a10 gene disruption have small brain ventricles and show reduced neuronal excitability. Proc Natl Acad Sci U S A 2008;105:311-6.
34. Davies G, Lam M, Harris SE, Trampush JW, Luciano M, Hill WD, et al. Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. Nat Commun 2018;9:2098.
35. Trubetskoy V, Pardinas AF, Qi T, Panagiotaropoulou G, Awasthi S, Bigdeli TB, et al. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. Nature 2022;604:502-8.
36. Williams CM, Labouret G, Wolfram T, Peyre H, Ramus F. A general cognitive ability factor for the UK Biobank. Behav Genet 2023;53:85-100.
37. Xiang AH, Wang X, Martinez MP, Walthall JC, Curry ES, Page K, et al. Association of maternal diabetes with autism in offspring. JAMA 2015;313:1425-34.
38. Evangelou E, Warren HR, Mosen-Ansorena D, Mifsud B, Pazoki R, Gao H, et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. Nat Genet 2018;50:1412-25.
39. Sakaue S, Kanai M, Tanigawa Y, Karjalainen J, Kurki M, Koshiba S, et al. A cross-population atlas of genetic associations for 220 human phenotypes. Nat Genet 2021;53:1415-24.
40. Huang J, Huffman JE, Huang Y, Do Valle I, Assimes TL, Raghavan S, et al. Genomics and phenomics of body mass index reveals a complex disease network. Nat Commun 2022;13:7973.
41. Zhu Z, Guo Y, Shi H, Liu CL, Panganiban RA, Chung W, et al. Shared genetic and experimental links between obesity-related traits and asthma subtypes in UK Biobank. J Allergy Clin Immunol 2020;145:537-49.
42. Popova PV, Klyushina AA, Vasilyeva LB, Tkachuk AS, Vasukova EA, Anopova AD, et al. Association of common genetic risk variants with gestational diabetes mellitus and their role in GDM prediction. Front Endocrinol (Lausanne) 2021;12:628582.
43. Powe CE, Nodzenski M, Talbot O, Allard C, Briggs C, Leya MV, et al. Genetic determinants of glycemic traits and the risk of gestational diabetes mellitus. Diabetes 2018;67:2703-9.
44. Shen Y, Jia Y, Li Y, Gu X, Wan G, Zhang P, et al. Genetic determinants of gestational diabetes mellitus: a case-control study in two independent populations. Acta Diabetol 2020;57:843-52.
45. Ding M, Chavarro J, Olsen S, Lin Y, Ley SH, Bao W, et al. Ge-

netic variants of gestational diabetes mellitus: a study of 112 SNPs among 8722 women in two independent populations. Diabetologia 2018;61:1758-68.
46. Lamri A, Mao S, Desai D, Gupta M, Pare G, Anand SS. Fine-tuning of genome-wide polygenic risk scores and prediction of gestational diabetes in South Asian Women. Sci Rep 2020;10:8941.

SUPPLEMENTARY METHODS

Cohort descriptions

Discovery cohort: HAPO-HK Study

The study methods employed in the Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong (HAPO-HK) Study have been previously documented [1]. The original HAPO Study is an international, prospective population-based study conducted at multiple centers. It aims to clarify the risks of adverse pregnancy outcomes linked to various degrees of maternal hyperglycemia occurring below the threshold for diagnosing overt diabetes mellitus in approximately 25,000 pregnant women during their third trimester of gestation, with medical caregivers 'blinded' to their glucose tolerance status [2]. In the HAPO-HK Study, we recruited a total of 1,674 pregnant women with singleton pregnancy at the Prince of Wales Hospital from 2000 to 2006 as part of the HAPO Study. These participants underwent a 75-g oral glucose tolerance test (OGTT) during weeks 24 to 32 of gestation, and the diagnosis of GDM was retrospectively applied based on the International Association of the Diabetes and Pregnancy Study Groups (IADPSG) criteria established in 2010. Standardized questionnaires were used to collect data on smoking and alcohol use, family history of diabetes and hypertension, as well as demographic characteristics. Individuals of non-Chinese ancestry, and those with glucose measurements outside the setting of the HAPO Study were excluded. Eligible participants were then invited to attend a follow-up assessment between 2009 and 2013, which took place approximately 7 years after childbirth. During this follow-up visit, all participants underwent a 75-g OGTT to ascertain their status of glucose tolerance, DNA samples were collected from these women. All participants provided written informed consent at the time of enrolment. The study received ethical approval (The approval numbers for the HAPO-HK study are: CRE-2002.119, CRE-2008.017, CRE-2013.042, CRE-2015.473) from the Clinical Research Ethics Committee of the Chinese University of Hong Kong.

Discovery cohort: Tianjin Study

A detailed description of the study's design, cohort, and methods was provided elsewhere [3]. Between October 2010 and August 2012, a prospective cohort study was conducted at Tianjin's Women and Children Health Center (TWCHC), recruiting a total of 22,302 pregnant women during their initial antenatal care visit through a universal screening and management system for gestational diabetes mellitus (GDM). These women participated a two-step screening procedure used to identify cases of GDM. Initially, pregnant women at 24 to 28 weeks of gestation, visiting a primary hospital, underwent a 1-hour 50-g glucose challenge test (GCT) in a non-fasting state. Women with a GCT result of ≥7.8 mmol/L were then referred to the GDM clinic at TWCHC for a 2-hour 75-g OGTT conducted in the morning after fasting for at least 8 hours. From July 2011 to June 2012, 2,991 pregnant women provided overnight fasting blood samples at the primary care hospitals. Among them, a total of 243 women with GDM and 243 women without GDM were chosen for a nested case-control study. Of them, 455 women (229 GDM cases vs. 226 non-GDM controls) with good quality DNA data were used in the current analysis. GDM was diagnosed according to the IADPSG criteria. The selection was based on matching maternal age (±1 year). The Ethics Committee for Clinical Research of Tianjin Women and Children's Health Center granted ethics approval (Ethics Approval number: 2009-02). Written informed consent was obtained from all pregnant women involved in the study.

Discovery cohort: Treated GDM Cases vs. Non-diabetes Controls (TGDM-NDM) Study

This is a case-control study that includes participants from two distinct sources. Firstly, the GDM cases were selected from an additional 141 pregnant women who were diagnosed with GDM based on OGTT at the Hong Kong center around the time of the HAPO Study. These women received antenatal treatment and were invited to return for follow-up assessment at around the same time as participants of the HAPO-HK follow-up study [4]. After excluding 47 women who were diagnosed with GDM but did not meet the criteria set by the IADPSG, along with seven women lacking genotype data and one woman with failed genotype data, a total of 86 women with GDM were included in the analysis.

Secondly, the non-diabetes and non-pregnant controls were selected from hospital staff and the "Better Health for Better Hong Kong (BHBHK)" Campaign, a community-based health awareness and promotion program [5]. The BHBHK Campaign was initiated between 2000 and 2002 and underwent re-evaluation between 2010 and 2014. Its primary objective is to raise awareness among the low-income working population about the importance of maintaining a healthy lifestyle through various educational and health screening strategies. Participants were randomly selected using stratified random sampling

based on the occupational groups documented in the 1996 Hong Kong Population By-Census Report. A total of 561 participants underwent detailed clinical and laboratory assessment at the Prince of Wales Hospital. To determine the glycemic status, these individuals underwent a 75-g OGTT using the criteria established by the American Diabetes Association. Women without diabetes at baseline and throughout the 10-year follow-up period were utilized as controls in the analysis. DNA samples from both the GDM cases and non-diabetes controls were genotyped using different genotyping platforms.

Replication cohort: Guangzhou Study

Pregnant women in their second trimester who attended the antenatal clinic of the Third Affiliated Hospital of Sun Yat-Sen University in Guangzhou between September 2013 and September 2015 for OGTT were included [6]. The prevalence of GDM during the study period was 16.8%. The diagnostic criteria for GDM were based on 75-g OGTT during weeks 24 to 28 of gestation, according to the IADPSG criteria. A total of 564 pregnant women diagnosed with GDM and 572 women without GDM were included in the study. Subjects with type 1 or type 2 diabetes mellitus, cardiovascular diseases (including hypertension and arteriosclerosis) and other metabolic disorders (including obesity body mass index [BMI] >30 kg/m^2 and polycystic ovary syndrome) were excluded. Whole blood samples collected during the OGTT were stored at –80°C prior to analysis. The study was approved by the Ethics Committee of the Third Affiliated Hospital of Sun Yat-Sen University, and all experimental procedures were performed in accordance with the relevant guidelines and regulations. Written informed consent was obtained from each individual to collect additional blood samples for genetic test.

Replication cohort: The FinnGen Project

The FinnGen Project, initiated in 2017, is a study that combines genome information with digital health care data. It operates as a public-private partnership between Finnish universities, biobanks, hospital districts, and several international pharmaceutical companies [7]. The primary objective of this project is to collect biological samples from 500,000 individuals in Finland over 6 years. Through genetic research, the project aims to enhance human health by identifying novel therapeutic targets and diagnostic tools for a wide range of diseases. The analysis in this study is based on the FinnGen release R8, which includes data from 342,499 individuals and disease endpoints. Clinical endpoints were derived from the register codes using the Finnish version of International Classification of Diseases, 10th revision (ICD-10) diagnosis codes and harmonized with definitions from ICD-8 and ICD-9. The comprehensive details on the methodology have been described elsewhere [8].

Replication cohort: The GenDIP Consortium

This study from the GENetics of Diabetes In Pregnancy (GenDIP) Consortium was a trans-ancestry meta-analysis of genetic associations for GDM in 5,485 women with GDM and 347,856 women without GDM. The descriptions for individual cohorts are available in previous reports [9].

Replication cohorts: Thai and Hispanic populations of the HAPO Study (HAPO-Thai and HAPO-Hispanic Studies)

The HAPO Study is an international epidemiological investigation with the objective of addressing unanswered questions regarding the associations between various levels of glucose intolerance and the risks of adverse outcomes during pregnancy [2]. Its general goal is to significantly advance our understanding of glucose levels during pregnancy that pose increased risks to the mother, fetus, and neonate. The study involves the collection of standardized, high-quality data on 25,000 women of diverse ethnic-racial and sociodemographic backgrounds from 16 centers worldwide. Additionally, it aims to generate data on the associations between glucose levels with the risks of adverse outcomes during pregnancy, which can be utilized to develop internationally acceptable guidelines for diagnosis and classification of GDM. For this particular study, we have exclusively included data from the Thai and Hispanic populations within the HAPO Study to validate the results of the polygenic risk score (PRS) analysis. This decision was based on the fact that only these two populations possessed complete genetic data necessary for the analysis.

The data/analyses presented in the current publication are based on the use of study data obtained from database of Genotypes and Phenotypes (dbGaP), a publicly accessible repository for genetic and phenotypic data. The specific dataset used in this study was titled "Hyperglycemia and Adverse Pregnancy Outcome (HAPO) Study: maternal glycemia and birthweight GEI Study," with accession number phs000096.v4.p1 (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000096.v4.p1). The entire dataset contains 1,500 infants and their mothers of European descent, 1,250 Af-

ro-Caribbean infants and mothers, 800 Hispanic (Mexican-American) infants and mothers, and 1,200 Thai infants and mothers. It was made available by the HAPO Study steering committee. The data download from dbGaP followed the guidelines and regulations for data access and usage as outlined by the repository.

Clinical and laboratory measurements

Descriptions of the assessment methods utilized in the HAPO-HK Study was provided elsewhere [1]. Briefly, all women underwent a standard 75-g OGTT between 24 and 32 weeks of gestation. Height, weight, and blood pressure were measured during the test visit. Data regarding smoking habits, family history of diabetes and hypertension, and demographic characteristics were collected using standardized questionnaires. Fasting blood samples were also taken to measure C-peptide levels and glycosylated hemoglobin.

Maternal weight at delivery was abstracted by research staff from the medical records. BMI was calculated as weight (kg) divided by the square of the measured height (m^2). Gestational weight gain (GWG) was computed as the mother's weight at delivery minus her pre-pregnancy weight, divided by the gestational age at delivery. Area under the curve (AUC) for glucose during OGTT at 0 to 120 minutes was calculated using the trapezoid rule. The homeostatic model assessment 2 (HOMA2) indices were computed based on the measurements of fasting glucose and C-peptide (https://www.dtu.ox.ac.uk/homacalculator/).

Genotyping, quality control, and imputation

For the genome-wide scan analysis, DNA samples from the three discovery cohorts (i.e., HAPO-HK Study, Tianjin Study, and TGDM-NDM Study) were genotyped using one of three arrays: (1) HumanOmniZhongHua-8 BeadChip, (2) Infinium Global Screening Array, and (3) Illumina Omni2.5+Exome Array. Samples from the Guangzhou Study were genotyped using the matrix-assisted laser desorption/ionization time-of-flight mass spectrometry platform.

We applied uniform quality control (QC) procedures to each set of genome-wide data. The QC of individual genotypes consisted of four steps: (1) verification of sex based on genotype calls from chromosomes X and Y; (2) identification of low-quality samples based on call rate and heterozygosity rate; (3) detection of potential familial relationships or duplicated individuals using estimates of identity-by-descent; and (4) assessment of population stratification through principal component (PC) analysis (Supplementary Fig. 5). Only biallelic autosomal single nucleotide polymorphisms (SNPs) were included in the per-marker QC. SNPs were excluded from further analysis if they met any of the following criteria: (1) Hardy–Weinberg equilibrium $P<1\times10^{-4}$; (2) minor allele frequency (MAF) <1%; or (3) call rate <95%. In particular, SNPs with MAF ≥1% but ≤5% are excluded if their call rate was <99%.

Within each cohort, the genotype data were imputed to the 1000 Genomes Project phase III reference panel (October 2014) using the Michigan Imputation Server [10]. SNPs with a MAF <1%, or an imputation quality score Rsq <0.5 were excluded. In total, 6,322,337 SNPs that were common across the discovery cohorts were included in the meta-analysis of genome-wide association studies (GWASs) for GDM.

Construction of polygenic risk score

We derived a PRS for GDM based on the variants identified in the current study. For each individual, the PRS was computed using the "score" command in PLINKv2.0 [11], represented as $\beta_1\chi_1+\beta_2\chi_2+\beta_3\chi_3+...+\beta_k\chi_k+...+\beta_n\chi_n$, where χ_k is the number of effect alleles at the k^{th} SNP, β_k is the corresponding β-coefficient for GDM risk estimated in this study, and n is the total number of SNPs involved in the PRS. The PRS was subsequently transformed into a z-score and classified into five categories using the quintile thresholds defined in the HAPO-HK Study, which is the largest cohort included in this study. Within this study, we assessed the performance of our PRS in two aspects: (1) detecting women who developed GDM during pregnancy and abnormal glucose tolerance (AGT) after pregnancy, and (2) improving the prediction of GDM and AGT after pregnancy compared to the clinical risk factors. In addition, we calculated a type 2 diabetes mellitus (T2DM)-related PRS based on 286 T2DM-related variants which were previously reported by the Diabetes Meta-Analysis of Trans-Ethnic association studies (DIAMANTE) Consortium [12]. We investigated the clinical utility of our GDM-related PRS in predicting the risk of AGT after pregnancy by comparing its performance with the T2DM-related PRS.

Statistical analyses

All analyses were performed using PLINK v1.9 and v2.0 [11], IBM SPSS Statistics version 26 (IBM Co., Armonk, NY, USA), and R 3.4.4 (http://www.r-project.org/, 31st December, 2019) unless specified otherwise. Meta-analysis was implemented by METAL software [13]. FUMA was used to annotate, prioritize,

and interpret the GWAS results [14]. Regional plot around genome-wide locus were visualized using LocusZoom (http://csg.sph.umich.edu/locuszoom/).

Data are expressed as percentage (*n*), mean and standard deviation, or median (interquartile range) as appropriate. Differences between groups were tested with chi-squared test, Student's *t*-test, or Mann-Whitney test, as appropriate.

Within the cohort, we performed logistic regression to examine the association of genetic markers (e.g., an individual SNP under an additive genetic model or the PRS) with the risks of GDM and AGT after pregnancy, with the adjustments for PCs, age and/or BMI. The results obtained from individual cohorts were combined through meta-analysis using an inverse-variance weighted approach under a fixed-effects model. Heterogeneity of effect across studies was assessed using Cochran's *Q* test. To address potential population stratification and relatedness among individuals, we adjusted for PCs in all association tests, and applied genomic control correction during the meta-analysis analysis. Associations of identified variants with glycemic and metabolic traits measured during pregnancy were tested by linear regression, adjusting for PCs, age, and/or BMI. Odds ratios with their 95% confidence intervals (CIs), or β±standard error were presented in these analyses. *P* values <0.05 and $<5.0\times10^{-8}$ were considered significant and genome-wide significant, respectively. In the candidate gene analysis, we adjusted for multiple testings using Bonferroni correction. Individuals with missing data points for any variables included in the logistic or linear regression model were removed from the analysis.

The area under the receiver operating characteristic curve (AUROC) and continuous net reclassification improvement (NRI) index were used to evaluate the incremental predictive value of PRS in GDM and AGT after pregnancy, over the clinical risk factors and PCs. We calculated the AUC and NRI index based on the predicted risk obtained from logistic regression, using respectively the "concordance.index" and "nricens" functions in R package. Bootstrapping with 10,000 iterations were used to estimate the 95% CI for the NRI index. We compared two correlated AUCs using the paired *t*-test implemented by the "cindex.comp" function in R package.

SUPPLEMENTARY RESULTS

Sensitivity analysis

In the HAPO-HK Study which included data on comprehensive clinical assessment during pregnancy, the associations between the four identified variants and GDM risk persisted after multivariate adjustment for BMI, GWG, blood pressure, smoking status, education year, parity, family history of diabetes and hypertension (Supplementary Table 14).

Associations for glycemic and metabolic traits during pregnancy

In the linear regression analysis adjusted for PCs and age, we observed several significant associations in the HAPO-HK Study: (1) the A-allele of T-box brain transcription factor 1 (*TBR1*)-solute carrier family 4 member 10 (*SLC4A10*) rs117781972 was associated with elevated levels of 1-hour glucose, 2-hour glucose and AUC_{glu} at 0 to 120 minutes ($1.7\times10^{-4}<P<0.0495$); (2) the C-allele of *CDKAL1* rs7754840 demonstrated an association with increased levels of 1-hour glucose and AUC_{glu} at 0 to 120 minutes ($0.0245<P<0.0249$); (3) there was a notable elevation in levels of 1-hour glucose, 2-hour glucose, AUC_{glu} at 0 to 120 minutes, fasting C-peptide and HOMA2 of insulin resistance (HOMA-IR) index per copy of the C-allele of *INS-IGF2*-potassium voltage-gated channel subfamily Q member 1 (*KCNQ1*) rs2237897 ($1.2\times10^{-4}<P<0.0109$); and (4) the C-allele carriers of melatonin receptor 1B (*MTNR1B*) rs7945617 had higher levels of fasting glucose, 1-hour glucose, 2-hour glucose, and AUC_{glu} at 0 to 120 minutes, as well as a reduced of HOMA2 of β-cell function (HOMA2-β) index ($1.1\times10^{-4}<P<0.0119$) (Supplementary Table 13). Adjustment for BMI did not further change these associations (Supplementary Table 13).

SUPPLEMENTARY REFERENCES

1. Tam WH, Ma RC, Ozaki R, Li AM, Chan MH, Yuen LY, et al. In utero exposure to maternal hyperglycemia increases childhood cardiometabolic risk in offspring. Diabetes Care 2017;40:679-86.
2. HAPO Study Cooperative Research Group; Metzger BE, Lowe LP, Dyer AR, Trimble ER, Chaovarindr U, et al. Hyperglycemia and adverse pregnancy outcomes. N Engl J Med 2008;358:1991-2002.
3. Leng J, Shao P, Zhang C, Tian H, Zhang F, Zhang S, 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.
4. Tam WH, Ma RCW, Ozaki R, Li AM, Chan MHM, Yuen LY, et al. Antenatal treatment of gestational diabetes and offspring's

future cardiometabolic risk. The 9th International Symposium on Diabetes, Hypertension and Metabolic Syndrome and in Pregnancy; 2017 Mar 8-12; Barcelona, Spain.
5. Ko GT, Chan JC, Chan AW, Wong PT, Hui SS, Tong SD, et al. Association between sleeping hours, working hours and obesity in Hong Kong Chinese: the 'better health for better Hong Kong' health promotion campaign. Int J Obes (Lond) 2007;31:254-60.
6. Wu L, Song Y, Zhang Y, Liang B, Deng Y, Tang T, et al. Novel genetic variants of PPARγ2 promoter in gestational diabetes mellitus and its molecular regulation in adipogenesis. Front Endocrinol (Lausanne) 2021;11:499788.
7. Kurki MI, Karjalainen J, Palta P, Sipila TP, Kristiansson K, Donner K, et al. FinnGen: unique genetic insights from combining isolated population and national health register data [Preprint]. Posted 2022 Mar 6. medRxiv https://doi.org/10.1101/2022.03.03.22271360.
8. Elliott A, Walters RK, Pirinen M, Kurki M, Junna N, Goldstein JI, et al. Distinct and shared genetic architectures of gestational diabetes mellitus and type 2 diabetes. Nat Genet 2024;56:377-82.
9. Pervjakova N, Moen GH, Borges MC, Ferreira T, Cook JP, Allard C, et al. Multi-ancestry genome-wide association study of gestational diabetes mellitus highlights genetic links with type 2 diabetes. Hum Mol Genet 2022;31:3377-91.
10. Das S, Forer L, Schonherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. Nat Genet 2016;48:1284-7.
11. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 2015;4:7.
12. Mahajan A, Spracklen CN, Zhang W, Ng MC, Petty LE, Kitajima H, et al. Multi-ancestry genetic study of type 2 diabetes highlights the power of diverse populations for discovery and translation. Nat Genet 2022;54:560-72.
13. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010;26:2190-1.
14. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun 2017;8:1826.
15. Huerta-Chagoya A, Vazquez-Cardenas P, Moreno-Macias H, Tapia-Maruri L, Rodriguez-Guillen R, Lopez-Vite E, et al. Genetic determinants for gestational diabetes mellitus and related metabolic traits in Mexican women. PLoS One 2015;10:e0126408.

Supplementary Table 1. Clinical characteristics of all individuals from the three discovery cohorts according to the status of gestational diabetes

Characteristic	HAPO-HK Study (population-based study) HumanOmniZhongHua-8 BeadChip			Tianjin Study (case-control study) Global Screening Array			TGDM-NDM Study (case-control study) HumanOmniZhongHua-8 BeadChip/Infinium Omni2.5+exome array		
	GDM case (n=149)	Non-GDM control (n=811)	P value	GDM case (n=229)	Non-GDM control (n=226)	P value	GDM case (n=86)	Non-diabetes control (n=180)	P value
Gestational age at OGTT, wk	27.6±1.34	27.7±1.30	0.3930	26.6±1.44	-	-	28.0±2.01	-	-
Age, yr	32.8±4.24	30.7±4.65	<0.0001	29.2±2.73	29.2±3.32	0.8522	35.2±3.97	41.4±8.11	<0.0001
Body mass index, kg/m^2	25.6±2.93	24.4±2.95	<0.0001	27.1±3.66	25.3±3.52	<0.0001	24.8±4.71	22.6±3.77	0.0003
Fasting plasma glucose, mmol/L	4.70±0.56	4.32±0.29	<0.0001	5.21±0.70	-	-	4.82±0.62	4.68±0.42	-
1-hr plasma glucose, mmol/L	9.98±1.68	7.29±1.32	<0.0001	10.1±1.61	-	-	-	-	-
2-hr plasma glucose, mmol/L	8.67±1.09	6.30±1.03	<0.0001	8.37±1.62	-	-	9.10±0.99	5.69±1.30	-
Glucose area under the curve, mmol/L	1,000±112	756±101	<0.0001	1,016±139	-	-	-	-	-
Fasting C-peptide, pmol/L	666 (500–866)	533 (433–674)	<0.0001	-	-	-	-	-	-
Fasting plasma insulin, mU/L	-	-	-	13.07±7.09	-	-	-	-	-
HOMA2-IR	1.45 (1.10–1.89)	1.14 (0.91–1.43)	<0.0001	-	-	-	-	-	-
HOMA2-% β	143 (125–164)	143 (127–167)	0.4908	-	-	-	-	-	-
Maternal smoking during pregnancy	1.4 (11)	2.7 (4)	0.2295	0.9 (2)	0.4 (1)	1.0000	-	-	-

Values are presented as mean±standard deviation, median (interquartile range), or percentage (number). Between-group comparisons were performed by unpaired Student's t-test, the Wilcoxon Rank Sum test, or the chi-square test.

HAPO-HK, Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong; TGDM-NDM, Treated GDM Cases vs. Non-diabetes Controls; GDM, gestational diabetes mellitus; OGTT, oral glucose tolerance test; HOMA2-IR, homeostatic model assessment 2 of insulin resistance; HOMA2-%β, homeostatic model assessment 2 of β-cell function.

Supplementary Table 2. Clinical characteristics of all individuals from the three replication cohorts according to the status of gestational diabetes

Characteristic	Guangzhou Study (case-control study) *de novo* replication MALDI-TOF mass spectrometry platform			HAPO-Thai Study (population-based study) *in silico* replication Illumina HumanOmni1-Quad v1.0 BeadChip			HAPO-Hispanic Study (population-based study) *in silico* replication Illumina Human1M-Duov3 BeadChip		
	GDM case (*n*=564)	Non-GDM control (*n*=572)	*P* value	GDM case (*n*=260)	Non-GDM control (*n*=933)	*P* value	GDM case (*n*=207)	Non-GDM control (*n*=596)	*P* value
Gestational age at OGTT, wk	-	-	-	28.1±1.86	28.2±1.77	0.5086	26.8±2.02	26.8±2.04	0.9901
Age, yr	29.2±4.16[a]	31.4±4.43[a]	<0.0001	29.3±5.64	27.3±5.39	<0.0001	30.7±5.38	28.3±5.29	<0.0001
Body mass index, kg/m^2	25.7±2.94[a]	26.0±3.04[a]	0.0992	26.6±3.91	25.4±3.49	<0.0001	33.1±6.52	29.1±4.90	<0.0001
Fasting plasma glucose, mmol/L	4.73±0.74	4.24±0.33	<0.0001	4.75±0.45	4.35±0.29	<0.0001	5.16±0.40	4.50±0.29	<0.0001
1-hr plasma glucose, mmol/L	9.97±1.70	7.21±1.40	<0.0001	10.4±1.32	7.62±1.30	<0.0001	9.49±1.90	7.02±1.49	<0.0001
2-hr plasma glucose, mmol/L	9.06±1.68	6.41±1.03	<0.0001	8.12±1.46	6.22±1.05	<0.0001	7.27±1.56	5.79±1.05	<0.0001
Glucose area under the curve, mmol/L	1,012±144	752±108	<0.0001	1,011±99.2	774±98.6	<0.0001	943±143	730±113	<0.0001
Fasting C-peptide, pmol/L	-	-	-	633 (466–799)	466 (366–599)	<0.0001	966 (766–1199)	666 (533–866)	<0.0001
HOMA2-IR	-	-	-	1.33 (0.99–1.73)	0.97 (0.76–1.25)	<0.0001	2.12 (1.68–2.68)	1.40 (1.10–1.86)	<0.0001
HOMA2-% β	-	-	-	133 (110–152)	126 (109–149)	0.2647	150 (129–173)	150 (130–182)	0.1761
Maternal smoking during pregnancy	-	-	-	1.2 (3)	0.5 (5)	0.2803	0 (0)	0.3 (2)	0.4040

Values are presented as mean±standard deviation, median (interquartile range), or percentage (number). Between-group comparisons were performed by unpaired Student's *t*-test, the Wilcoxon Rank Sum test, or the chi-square test.

MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight; HAPO, Hyperglycemia and Adverse Pregnancy Outcome; GDM, gestational diabetes mellitus; OGTT, oral glucose tolerance test; HOMA2-IR, homeostatic model assessment 2 of insulin resistance; HOMA2-%β, homeostatic model assessment 2 of β-cell function.

[a]Data on age and body mass index were available for only 348 women with GDM and 439 women without GDM.

Supplementary Table 3. Associations of *TBR1-SLC4A10* rs117781972 with gestational diabetes in all discovery and replication cohorts

Stage	Study	Ethnicity	Population	Study design	Covariates	Imputation quality, Rsq	Number		Risk allele frequency		Association for gestational diabetes					
							GDM case	Non-GDM control	GDM case	Non-GDM control	OR (95% CI)	$P_{additive}$	P_{GC}	P_{Q}	$P_{ancestry}$	$P_{residual}$
Discovery	HAPO-HK Study	East Asian	Chinese pregnant women	Population-based	Age and PCs	0.806	149	811	0.313	0.193	2.26 (1.64–3.11)	5.1×10^{-7}	-	-	--	--
	Tianjin Study	East Asian	Chinese pregnant women	Case-control	Age and PCs	0.527	229	226	0.189	0.143	1.97 (1.18–3.27)	9.4×10^{-3}	-	-	--	--
	TGDM-NDM Study	East Asian	Cases: Chinese pregnant women; Controls: Chinese non-pregnant women	Case-control	Age and PCs	0.789	86	180	0.233	0.180	1.59 (0.90–2.81)	0.1081	-	-	--	--
	Meta-analysis of discovery cohorts	East Asian	Chinese women	-	-	-	464	1,217	-	-	2.05 (1.61–2.62)	7.6×10^{-9}	1.1×10^{-8}	0.5624	--	--
De novo replication	Guangzhou Study	East Asian	Chinese pregnant women	Case-control	None	-	564	572	0.230	0.241	0.94 (0.78–1.14)	0.5465	-	-	--	--
In silico replication	FinnGen Study [8]	European	Finnish pregnant women	Case-control	Age, PCs and genotyping batch	-	12,332	131,109	0.010		0.98 (0.85–1.13)	0.7850	-	-	--	--
	GenDIP Consortium [9]	Multi-ethnicities	72.2% European, 13.4% East Asian, 9.9% South Asian, 2.8% Hispanic/Latino, and 1.7% African	Meta-analysis	-	-	6,857		0.071		1.31 (1.07–1.60)	0.0164	-	-	0.1690	4.5×10^{-4}

OR and 95% CI were estimated according to the GDM-related risk allele (A-allele) of *TBR1-SLC4A10* rs117781972. $P_{additive}$ was obtained from either logistic regression model with the adjustments of covariates, or meta-analysis under a fixed effect model. P_{GC} was obtained from meta-analysis under a fixed effect model after genomic control. P_{Q} was obtained from heterogeneity test (Cochran's Q test). For the GenDIP analysis, $P_{additive}$, $P_{ancestry}$, and $P_{residual}$ were obtained from meta-regression implement in Meta-Regression of Multi-AncEstry Genetic Association (MR-MEGA). The combined OR and 95% CI was estimated by meta-analysis.

TBR1, T-box brain transcription factor 1; SLC4A10, solute carrier family 4 member 10; GDM, gestational diabetes mellitus; OR, odds ratio; CI, confidence interval; HAPO-HK, Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong; PC, principal component; TGDM-NDM, Treated GDM Cases vs. Non-diabetes Controls; GenDIP, GENetics of Diabetes In Pregnancy.

Supplementary Table 4. Associations of *MTNR1B* rs7945617 with gestational diabetes in all discovery and replication cohorts

Stage	Study	Ethnicity	Population	Study design	Covariates	Imputation quality, Rsq	Number		Risk allele frequency		Association for gestational diabetes					
							GDM case	Non-GDM control	GDM case	Non-GDM control	OR (95% CI)	$P_{additive}$	P_{GC}	P_{Q}	$P_{ancestry}$	$P_{residual}$
Discovery	HAPO-HK Study	East Asian	Chinese pregnant women	Population-based	Age and PCs	0.883	149	811	0.572	0.476	1.61 (1.23–2.12)	6.0×10^{-4}	-	-	--	--
	Tianjin Study	East Asian	Chinese pregnant women	Case-control	Age and PCs	0.949	229	226	0.535	0.434	1.48 (1.14–1.93)	3.7×10^{-3}	-	-	--	--
	TGDM-NDM Study	East Asian	Cases: Chinese pregnant women; Controls: Chinese non-pregnant women	Case-control	Age and PCs	0.893	86	180	0.557	0.411	2.40 (1.52–3.79)	1.9×10^{-4}	-	-	--	--
	Meta-analysis of discovery cohorts	East Asian	Chinese women	-	-	-	464	1,217	-	-	1.64 (1.38–1.96)	2.6×10^{-8}	4.6×10^{-8}	0.1988	--	--
De novo replication	Guangzhou Study	East Asian	Chinese pregnant women	Case-control	None	-	564	572	0.524	0.477	1.21 (1.02–1.43)	0.0247	-	-	--	--
In silico replication	FinnGen Study [8]	European	Finnish pregnant women	Case-control	Age, PCs and genotyping batch	-	12,332	131,109	0.406		1.34 (1.31–1.38)	1.4×10^{-95}	-	-	-	--
	GenDIP Consortium [9]	Multi-ethnicities	72.2% European, 13.4% East Asian, 9.9% South Asian, 2.8% Hispanic/Latino, and 1.7% African	Meta-analysis	-	-	334,576		0.400		1.29 (1.24–1.35)	6.3×10^{-35}	-	-	2.7×10^{-3}	0.3162

OR and 95% CI were reported according to the GDM-related risk allele (C-allele) of *MTNR1B* rs7945617. $P_{additive}$ was obtained from either logistic regression model with the adjustments of covariates, or meta-analysis under a fixed effect model. P_{GC} was obtained from meta-analysis under a fixed effect model after genomic control. P_{Q} was obtained from heterogeneity test (Cochran's Q test). For the GenDIP analysis, $P_{additive}$, $P_{ancestry}$, and $P_{residual}$ were obtained from meta-regression implement in Meta-Regression of Multi-AncEstry Genetic Association (MR-MEGA). The combined OR and 95% CI was estimated by meta-analysis.

MTNR1B, melatonin receptor 1B; GDM, gestational diabetes mellitus; OR, odds ratio; CI, confidence interval; HAPO-HK, Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong; PC, principal component; TGDM-NDM, Treated GDM Cases vs. Non-diabetes Controls; GenDIP, GENetics of Diabetes In Pregnancy.

Supplementary Table 5. Associations of type 2 diabetes mellitus-related variants and gestational diabetes in Chinese women

Supplementary Table 6. Associations of *KCNQ1* rs2237897 with gestational diabetes in all discovery and replication cohorts

Stage	Study	Ethnicity	Population	Study design	Covariates	Imputation quality, Rsq	Number		Risk allele frequency		Association for gestational diabetes					
							GDM case	Non-GDM control	GDM case	Non-GDM control	OR (95% CI)	$P_{additive}$	P_{GC}	P_{Q}	$P_{ancestry}$	$P_{residual}$
Discovery	HAPO-HK Study	East Asian	Chinese pregnant women	Population-based	Age and PCs	0.968	149	811	0.741	0.634	1.76 (1.31–2.36)	1.7×10^{-4}	-	-	-	-
	Tianjin Study	East Asian	Chinese pregnant women	Case-control	Age and PCs	0.995	229	226	0.715	0.661	1.29 (0.97–1.72)	0.0806	-	-	-	-
	TGDM-NDM Study	East Asian	Cases: Chinese pregnant women; Controls: Chinese non-pregnant women	Case-control	Age and PCs	0.963	86	180	0.695	0.644	1.40 (0.88–2.25)	0.1582	-	-	-	-
	Meta-analysis of discovery cohorts	East Asian	Chinese women	-	-	-	464	1,217	-	-	1.48 (1.23–1.79)	3.8×10^{-5}	4.7×10^{-5}	0.3238	-	-
De novo replication	Guangzhou Study	East Asian	Chinese pregnant women	Case-control	None	-	564	572	0.717	0.660	1.33 (1.10–1.60)	2.7×10^{-3}	-	-	-	-
In silico replication I	FinnGen Study [8]	European	Finnish pregnant women	Case-control	Age, PCs and genotyping batch	-	12,332	131,109	0.919		1.13 (1.07–1.19)	8.0×10^{-6}	-	-	-	-
	GenDIP Consortium [9]	Multi-ethnicities	72.2% European, 13.4% East Asian, 9.9% South Asian, 2.8% Hispanic/Latino, and 1.7% African	Meta-analysis	-	-	345,656		0.951		1.19 (1.10–1.28)	2.2×10^{-4}	-	-	0.6109	0.1958
In silico replication II	Mexican Study [15]	Hispanic	Mexican pregnant women	Case-control	Age, pregestational BMI, dummy reference hospital and PCs	-	408	342	0.753	0.656	1.85 (1.39–2.50)	4.5×10^{-5}	-	-	-	-
Lookup from literatures	Overall meta-analysis	Multi-ethnicities	1.9% Chinese, 0.5% Mexican, and 97.6% Finnish	-	-	-	13,768	133,240	-	-	1.16 (1.10–1.22)	2.4×10^{-9}	-	0.0405	-	-

A total of three studies (i.e., the HAPO-HK Study, the Tianjin Study, and the TGDM-NDM Study) were included in the "meta-analysis of discovery cohorts." A total of six studies (i.e., the HAPO-HK Study, the Tianjin Study, the TGDM-NDM Study, the Guangzhou Study, the FinnGen Study, and the Mexican Study) were included in the "overall meta-analysis." OR and 95% CI were reported according to the type 2 diabetes mellitus-related risk allele (C-allele) of *KCNQ1* rs2237897. $P_{additive}$ was obtained from either logistic regression model with the adjustments of covariates, or meta-analysis under a fixed effect model. P_{GC} was obtained from meta-analysis under a fixed effect model after genomic control. P_{Q} was obtained from heterogeneity test (Cochran's Q test). For the GenDIP analysis, $P_{additive}$, $P_{ancestry}$, and $P_{residual}$ were obtained from meta-regression implement in Meta-Regression of Multi-AncEstry Genetic Association (MR-MEGA). The combined OR and 95% CI was estimated by meta-analysis under a fixed effect model.

KCNQ1, potassium voltage-gated channel subfamily Q member 1; GDM, gestational diabetes mellitus; OR, odds ratio; CI, confidence interval; HAPO-HK, Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong; PC, principal component; TGDM-NDM, Treated GDM Cases vs. Non-diabetes Controls; GenDIP, GENetics of Diabetes In Pregnancy; BMI, body mass index.

Supplementary Table 7. Conditional analysis for gestational diabetes at the *CDKAL1* and *MTNR1B* loci in Chinese women

SNP	Chr	Position (Build 37)	Nearest gene(s)	Risk/ other alleles	Conditional SNP	Cohort	LD r^2	Number		Risk allele frequency		Unconditional analysis			Conditional analysis		
								Case	Control	Case	Control	OR (95% CI)	*P* value	P_Q	OR (95% CI)	*P* value	P_Q
rs7754840	6	20661250	*CDKAL1*	C/G	rs9348441	HAPO-HK Study	0.809	149	811	0.451	0.361	1.47 (1.14–1.90)	3.4E-03	-	0.82 (0.43–1.55)	0.5415	-
						Tianjin Study	0.812	229	226	0.465	0.381	1.41 (1.08–1.84)	0.0110	-	1.61 (0.86–3.01)	0.1390	-
						TGDM-NDM Study	0.807	86	180	0.353	0.346	0.92 (0.61–1.38)	0.6836	-	0.33 (0.13–0.86)	0.0232	-
						Meta-analysis	-	464	1,217	-	-	1.33 (1.13–1.58)	7.9E-04	0.1400	0.92 (0.61–1.39)	0.6981	0.0228
rs9348441	6	20661250	*CDKAL1*	C/G	rs7754840	HAPO-HK Study	0.809	149	811	0.445	0.341	1.59 (1.22–2.06)	4.9E-04	-	1.91 (1.00–3.64)	0.0500	-
						Tianjin Study	0.812	229	226	0.445	0.376	1.34 (1.02–1.75)	0.0340	-	0.87 (0.46–1.63)	0.6567	-
						TGDM-NDM Study	0.807	86	180	0.403	0.341	1.16 (0.79–1.71)	0.4565	-	3.01 (1.20–7.53)	0.0187	-
						Meta-analysis	-	464	1,217	-	-	1.40 (1.18–1.65)	9.4E-05	0.3808	1.51 (1.01–2.27)	0.0469	0.0603
rs10830962	11	92698427	*MTNR1B*	G/C	rs7945617	HAPO-HK Study	0.899	149	811	0.522	0.439	1.52 (1.16–2.00)	2.4E-03	-	0.69 (0.29–1.61)	0.3899	-
						Tianjin Study	0.928	229	226	0.505	0.421	1.36 (1.05–1.76)	0.0180	-	0.26 (0.07–0.97)	0.0445	-
						TGDM-NDM Study	0.877	86	180	0.504	0.384	2.01 (1.28–3.16)	2.6E-03	-	0.26 (0.06–1.10)	0.0677	-
						Meta-analysis	-	464	1,217	-	-	1.51 (1.27–1.79)	3.1E-06	0.3381	0.45 (0.24–0.86)	0.0148	0.3373
rs7945617	11	92700287	*MTNR1B*	C/T	rs10830962	HAPO-HK Study	0.899	149	811	0.572	0.476	1.61 (1.23–2.12)	6.0E-04	-	2.30 (0.98–5.36)	0.0548	-
						Tianjin Study	0.928	229	226	0.535	0.434	1.48 (1.14–1.93)	3.7E-03	-	5.75 (1.48–22.3)	0.0115	-
						TGDM-NDM Study	0.877	86	180	0.557	0.411	2.40 (1.52–3.79)	1.9E-04	-	8.42 (1.99–35.6)	3.8E-03	-
						Meta-analysis	-	464	1,217	-	-	1.64 (1.38–1.96)	2.6E-08	0.1988	3.66 (1.92–6.96)	7.9E-05	0.2383
rs7945617	11	92700287	*MTNR1B*	C/T	rs10830963	HAPO-HK Study	0.827	149	811	0.572	0.476	1.61 (1.23–2.12)	6.0E-04	-	2.21 (1.14–4.29)	0.0188	-
						Tianjin Study	0.879	229	226	0.535	0.434	1.48 (1.14–1.93)	3.7E-03	-	1.67 (0.74–3.79)	0.2179	-
						TGDM-NDM Study	0.776	86	180	0.557	0.411	2.40 (1.52–3.79)	1.9E-04	-	3.40 (1.07–10.8)	0.0378	-
						Meta-analysis	-	464	1,217	-	-	1.64 (1.38–1.96)	2.6E-08	0.1988	2.17 (1.35–3.47)	1.3E-03	0.6150
rs10830963	11	92708710	*MTNR1B*	G/C	rs7945617	HAPO-HK Study	0.827	149	811	0.498	0.426	1.51 (1.13–2.01)	5.3E-03	-	0.69 (0.34–1.40)	0.3063	-
						Tianjin Study	0.879	229	226	0.478	0.393	1.47 (1.11–1.96)	7.9E-03	-	0.87 (0.36–2.11)	0.7550	-
						TGDM-NDM Study	0.776	86	180	0.486	0.357	2.22 (1.37–3.60)	1.2E-03	-	0.67 (0.20–2.27)	0.5156	-
						Meta-analysis	-	464	1,217	-	-	1.58 (1.31–1.91)	1.5E-06	0.3257	0.74 (0.45–1.22)	0.2403	0.9103

Association was assessed by logistic regression with adjustments for age and principal components. Results obtained from the individual cohorts were meta-analyzed under a fixed effect model. OR and 95% CI were reported according to the gestational diabetes mellitus-related risk allele. P_Q was obtained from the heterogeneity test (Cochran's Q test).
CDKAL1, CDK5 regulatory subunit-associated protein 1-like 1; *MTNR1B*, melatonin receptor 1B; SNP, single nucleotide polymorphism; Chr, chromosome; LD, linkage disequilibrium; OR, odds ratio; CI, confidence interval; HAPO-HK, Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong; PC, principal component; TGDM-NDM, Treated GDM Cases vs. Non-diabetes Controls.

Supplementary Table 8. Associations between polygenic risk score and risk of gestational diabetes in all discovery and replication cohorts

Stage	Cohort	Ethnicity	Population	Study design	Covariates	Number		Z-score			Top 20% vs. others		
						GDM case	Non-GDM control	OR (95% CI)	*P* value	P_Q	OR (95% CI)	*P* value	P_Q
Discovery	HAPO-HK Study	East Asian	Chinese pregnant women	Population-based	PCs, age, and BMI	149	810	2.07 (1.71–2.50)	1.2×10^{-13}	-	3.49 (2.35–5.19)	6.6×10^{-10}	-
					PCs, age, BMI, GWG, SBP, DBP, smoking status, education year, parity, family history of diabetes and hypertension	145	776	2.08 (1.70–2.53)	5.5×10^{-13}	-	3.54 (2.35–5.34)	1.6×10^{-9}	-
	Tianjin Study	East Asian	Chinese pregnant women	Case-control	PCs, age, and BMI	229	226	1.67 (1.36–2.06)	8.9×10^{-7}	-	2.68 (1.61–4.43)	1.3×10^{-4}	-
	TGDM-NDM Study	East Asian	Cases: Chinese pregnant women; Controls: Chinese non-pregnant women	Case-control	PCs, age, and BMI	86	180	1.68 (1.23–2.29)	1.0×10^{-3}	-	2.58 (1.20–5.58)	0.0158	-
	Meta-analysis of discovery cohorts	East Asian	Chinese women	-	PCs, age, and BMI	464	1,216	1.84 (1.62–2.09)	9.2×10^{-21}	0.2788	3.07 (2.30–4.09)	3.1×10^{-14}	0.6440
De novo replication	Guangzhou Study	East Asian	Chinese pregnant women	Case-control	None	560	566	1.16 (1.03–1.31)	0.0133	-	1.25 (0.96–1.64)	0.0961	-
In silico replication	HAPO-Thai Study	East Asian	Thai pregnant women	Population-based	PCs, age, and BMI	260	933	1.08 (0.94–1.25)	0.2695	-	0.77 (0.49–1.20)	0.2490	-
					PCs, age, BMI, GWG, SBP, DBP, smoking status, education year, parity, family history of diabetes and hypertension	134	459	1.03 (0.84–1.27)	0.7755	-	0.61 (0.30–1.24)	0.1741	-
	HAPO-Hispanic Study	Hispanic	Hispanic pregnant women	Population-based	PCs, age, and BMI	207	596	1.41 (1.18–1.67)	1.2×10^{-4}	-	1.96 (1.14–3.39)	0.0154	-
					PCs, age, BMI, GWG, SBP, DBP, smoking status, education year, parity, family history of diabetes and hypertension	194	556	1.37 (1.14–1.64)	7.8×10^{-4}	-	1.90 (1.07–3.37)	0.0284	-

OR and 95% CI, as well as *P* value were obtained from either logistic regression model with the adjustments for covariates, or meta-analysis under a fixed effect model. P_Q was obtained from heterogeneity test (Cochran's Q test).

GDM, gestational diabetes mellitus; OR, odds ratio; CI, confidence interval; HAPO-HK, Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong; BMI, body mass index; GWG, gestational weight gain; SBP, systolic blood pressure; DBP, diastolic blood pressure; PC, principal component; TGDM-NDM, Treated GDM Cases vs. Non-diabetes Controls.

Supplementary Table 9. Incremental value of polygenic risk score in predicting gestational diabetes, above and beyond the clinical risk factors

Stage	Cohort	Ethnicity	Population	Study design	PCs and clinical factors	Number		ROC curve analysis						Continuous NRI (95% CI)	P_{NRI}
								Model 1: PRS only		Model 2 (base model): PCs and clinical risk factors		Model 3: PRS, PCs and clinical risk factors		Comparison of AUC Model 3 vs. Model 2			
						GDM case	Non-GDM control	AUC (95% CI)	*P* value	AUC (95% CI)	*P* value	AUC (95% CI)	*P* value	Increased AUC	$P_{increase}$		
Discovery	HAPO-HK Study	East Asian	Chinese pregnant women	Population-based	PCs, age, and BMI	149	811	0.692 (0.646–0.738)	4.3×10^{-14}	0.675 (0.630–0.720)	5.3×10^{-12}	0.753 (0.711–0.795)	4.4×10^{-23}	0.0779	1.4×10^{-4}	57.4 (39.6 to 75.3)	<0.05
					PCs, age, BMI, GWG, SBP, DBP, smoking status, education year, parity, family history of diabetes and hypertension	145	776	0.695 (0.649–0.742)	4.0×10^{-14}	0.695 (0.647–0.742)	4.9×10^{-14}	0.767 (0.725–0.809)	9.4×10^{-25}	0.0722	2.7×10^{-4}	58.1 (36.7 to 74.1)	<0.05
	Tianjin Study	East Asian	Chinese pregnant women	Case-control	PCs, age, and BMI	229	226	0.625 (0.574–0.676)	2.0×10^{-6}	0.643 (0.593–0.694)	6.3×10^{-8}	0.693 (0.645–0.741)	5.1×10^{-13}	0.0499	0.0134	35.6 (18.4 to 59.3)	<0.05
	TGDM-NDM Study	East Asian	Cases: Chinese pregnant women; Controls: Chinese non-pregnant women	Case-control	PCs, age, and BMI	86	180	0.643 (0.575–0.712)	8.0×10^{-5}	0.785 (0.731–0.838)	3.0×10^{-14}	0.810 (0.758–0.861)	1.6×10^{-16}	0.0250	0.0596	25.8 (1.4 to 60.8)	<0.05
De novo replication	Guangzhou Study	East Asian	Chinese pregnant women	Case-control	Age, and BMI	344	436	0.543 (0.503–0.584)	0.0192	0.647 (0.608–0.685)	9.5×10^{-13}	0.651 (0.613–0.690)	1.9×10^{-13}	0.0046	0.4568	10.8 (–2.5 to 24.9)	NS
In silico replication	HAPO-Thai Study	East Asian	Thai pregnant women	Population-based	PCs, age, and BMI	260	933	0.542 (0.503–0.581)	0.0191	0.638 (0.600–0.676)	4.5×10^{-12}	0.640 (0.602–0.678)	2.6×10^{-12}	0.0016	0.6199	9.8 (–5.7 to 24.5)	NS
					PCs, age, BMI, GWG, SBP, DBP, smoking status, education year, parity, family history of diabetes and hypertension	134	459	0.531 (0.477–0.584)	0.1414	0.711 (0.661–0.761)	4.7×10^{-14}	0.711 (0.661–0.761)	4.8×10^{-14}	3.25E-05	0.9830	–1.6 (–10.2 to 26.2)	NS
	HAPO-Hispanic Study	Hispanic	Hispanic pregnant women	Population-based	PCs, age, and BMI	207	596	0.570 (0.524–0.616)	1.3×10^{-3}	0.713 (0.673–0.754)	2.8×10^{-20}	0.729 (0.689–0.769)	5.1×10^{-23}	0.0153	0.0595	29.9 (12.7 to 47.1)	<0.05
					PCs, age, BMI, GWG, SBP, DBP, smoking status, education year, parity, family history of diabetes and hypertension	194	556	0.561 (0.513–0.609)	5.6×10^{-3}	0.735 (0.694–0.776)	1.1×10^{-22}	0.742 (0.701–0.783)	5.7×10^{-24}	0.0070	0.2863	26.2 (5.8 to 45.6)	<0.05

ROC curve analysis: Model 1 includes the PRS only; Model 2 (base model) includes the principal components and the clinical risk factors; Model 3 includes both models 1 and 2. AUC (95% CI) and related changes were estimated based on the risk prediction models (i.e., models 1–3) computed by logistic regression. *P* tests the null hypothesis that "AUC is equal to zero." $P_{increase}$ tests the null hypothesis that "difference in AUC between models 2 and 3 is equal to zero."

Reclassification analysis: Old model includes the principal components and the clinical risk factors. New model includes old model and PRS. P_{NRI} tests the null hypothesis that "continuous NRI is equal to zero."

PC, principal component; GDM, gestational diabetes mellitus; ROC, receiver operating characteristic; PRS, polygenic risk score; AUC, area under the curve; NRI, net reclassification improvement; CI, confidence interval; HAPO-HK, Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong; BMI, body mass index; GWG, gestational weight gain; SBP, systolic blood pressure; DBP, diastolic blood pressure; TGDM-NDM, Treated GDM Cases vs. Non-diabetes Controls; NS, not significant.

Supplementary Table 10. Clinical characteristics of all individuals from the HAPO-HK 7-year follow-up study according to the status of abnormal glucose tolerance

Characteristic	HAPO-HK 7-year follow-up study (population-based study) HumanOmniZhongHua-8 BeadChip		
	AGT case (n=129)	Non-AGT control (n=829)	P value
Age, yr	39.6±4.52	37.4±4.61	<0.0001
Body mass index, kg/m^2	25.8±4.02	22.7±3.30	<0.0001
Fasting plasma glucose, mmol/L	5.69±1.71	4.69±0.34	<0.0001
2-hr plasma glucose, mmol/L	8.96±3.40	5.26±1.07	<0.0001
Glucose area under the curve, mmol/L	14.6±4.88	9.95±1.22	<0.0001
Impaired fasting glucose only	22.5 (29)	0 (0)	<0.0001
Impaired glucose tolerance only	51.9 (67)	0 (0)	<0.0001
Impaired fasting glucose and impaired glucose tolerance	11.6 (15)	0 (0)	<0.0001
Diabetes	14.0 (18)	0 (0)	<0.0001
History of gestational diabetes	39.5 (51)	11.8 (98)	<0.0001

Values are presented as mean±standard deviation or percentage (number). Between-group comparisons were performed by unpaired Student's t-test, the Wilcoxon Rank Sum test, or the chi-square test.
HAPO-HK, Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong; AGT, abnormal glucose tolerance.

Supplementary Table 11. Associations between polygenic risk score and risk of abnormal glucose tolerance at 7-year postpartum in the HAPO-HK Study

Variable	Number		PRS derived from 4 GDM-related variants				PRS derived from 286 T2DM-related variants [12]			
			Z-score of PRS		Top 20% vs. others		Z-score of PRS		Top 20% vs. others	
	AGT case	Non-AGT control	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
PCs, age, and BMI	129	829	1.40 (1.15–1.71)	6.7×10^{-4}	1.71 (1.08–2.71)	0.0220	1.78 (1.44–2.19)	8.9×10^{-8}	2.12 (1.36–3.29)	8.3×10^{-4}
PCs, age, BMI, and GDM history	129	828	1.22 (0.99–1.50)	0.0571	1.26 (0.77–2.05)	0.3589	1.63 (1.32–2.03)	8.8×10^{-6}	1.81 (1.15–2.87)	0.0108
PCs, age, BMI, and another PRS[a]	129	828	1.24 (1.01–1.53)	0.0408	1.56 (0.98–2.49)	0.0632	1.68 (1.35–2.08)	3.0×10^{-6}	2.01 (1.28–3.13)	2.2×10^{-3}

The T2DM-related PRS was derived based on 286 T2DM-related variants reported by Diabetes Meta-Analysis of Trans-Ethnic association studies (DIAMANTE) consortium [12]. OR and 95% CI, as well as *P* value were obtained from either logistic regression model with the adjustments for covariates.

HAPO-HK, Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong; AGT, abnormal glucose tolerance; PRS, polygenic risk score; GDM, gestational diabetes mellitus; OR, odds ratio; CI, confidence interval; T2DM, type 2 diabetes mellitus; PC, principal component; BMI, body mass index.

[a]In the analysis for GDM-related PRS, the covariates included PCs, age, BMI, and T2DM-related PRS. In the analysis for T2DM-related PRS, the covariates included PCs, age, BMI, and GDM-related PRS.

Supplementary Table 12. Incremental value of polygenic risk score in predicting abnormal glucose tolerance at 7-year postpartum, over the clinical risk factors, in the HAPO-HK Study

Variable	Number		ROC curve analysis								Continuous NRI (95% CI)	P_{NRI}
			Model 1: PRS only		Model 2 (base model): PCs and clinical risk factors		Model 3: PRS, PCs, and clinical risk factors		Comparison of AUC Model 3 vs. Model 2			
	GDM case	Non-GDM control	AUC (95% CI)	*P* value	AUC (95% CI)	*P* value	AUC (95% CI)	*P* value	Increased AUC	*P* value		
PCs, age, and BMI	129	829	0.595 (0.544–0.645)	2.6×10^{-4}	0.750 (0.701–0.798)	3.5×10^{-20}	0.767 (0.721–0.813)	8.9×10^{-23}	0.0171	0.0187	31.7 (12.6–49.6)	<0.05
PCs, age, BMI, and GDM status	129	828	0.595 (0.544–0.645)	2.6×10^{-4}	0.783 (0.738–0.828)	2.3×10^{-25}	0.787 (0.743–0.831)	4.1×10^{-26}	0.0045	0.1623	25.5 (1.50–42.8)	<0.05
PCs, age, BMI, and T2DM-related PRS	129	828	0.595 (0.544–0.645)	2.6×10^{-4}	0.781 (0.736–0.827)	3.8×10^{-25}	0.788 (0.744–0.833)	2.7×10^{-26}	0.0069	0.0822	25.6 (2.10–42.1)	<0.05

The T2DM-related PRS was derived based on 286 T2DM-related variants reported by Diabetes Meta-Analysis of Trans-Ethnic association studies (DIAMANTE) consortium [2].
ROC curve analysis: Model 1 includes the PRS only; Model 2 (base model) includes the principal components and the clinical risk factors; Model 3 includes both models 1 and 2. AUC (95% CI) and related changes were estimated based on the risk prediction models (i.e., models 1–3) computed by logistic regression. *P* tests the null hypothesis that "AUC is equal to zero." $P_{increase}$ tests the null hypothesis that "difference in AUC between models 2 and 3 is equal to zero."
Reclassification analysis: Old model includes the principal components and the clinical risk factors. New model includes old model and PRS. P_{NRI} tests the null hypothesis that "continuous NRI is equal to zero."
HAPO-HK, Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong; GDM, gestational diabetes mellitus; ROC, receiver operating characteristic; PRS, polygenic risk score; PC, principal component; AUC, area under the curve; CI, confidence interval; NRI, net reclassification improvement; BMI, body mass index.

Supplementary Table 13. Associations of identified variants with glycemic and metabolic traits during pregnancy in Chinese women from the HAPO-HK Study

SNP	Nearest gene(s)	Chr	Position (Build 37)	Risk/non-risk allele	Trait measured during pregnancy	Transformation	Model I: PCs and age				Model II: PCs, age, and BMI			
							No.	Risk allele frequency	β ± SE	*P* value	No.	Risk allele frequency	β ± SE	*P* value
rs117781972	*TBR1-SLC4A10*	2	162,332,226	A/G	Fasting glucose, mmol/L	Winsorization	961	0.212	0.019 ± 0.021	0.3826	960	0.211	0.014 ± 0.021	0.5068
					1-hr glucose, mmol/L	Winsorization	960	0.212	0.200 ± 0.102	0.0495	959	0.211	0.183 ± 0.100	0.0681
					2-hr glucose, mmol/L	Winsorization	961	0.212	0.302 ± 0.080	1.7×10^{-4}	960	0.211	0.288 ± 0.079	2.8×10^{-4}
					AUC_{glu} at 0–120 min	Winsorization	960	0.212	21.82 ± 8.066	7.0×10^{-3}	959	0.211	20.24 ± 7.911	0.0107
					HbA1c, %	Winsorization	886	0.212	–0.006 ± 0.023	0.7944	886	0.212	–0.009 ± 0.023	0.6904
					HOMA2-β	Natural logarithm	956	0.212	0.012 ± 0.014	0.3597	955	0.212	0.007 ± 0.013	0.5774
					HOMA2-IR	Natural logarithm	956	0.212	0.031 ± 0.022	0.1523	955	0.212	0.020 ± 0.019	0.2944
					Fasting C-peptide, pmol/L	Natural logarithm	956	0.212	0.030 ± 0.021	0.1559	955	0.212	0.019 ± 0.018	0.3020
					Body mass index, kg/m^2	Winsorization	960	0.211	0.201 ± 0.184	0.2741	-	-	-	-
					Gestational weight gain, kg/wk	Winsorization	953	0.212	0.003 ± 0.008	0.7224	953	0.212	0.0004 ± 0.008	0.9613
					Systolic blood pressure, mm Hg	Winsorization	960	0.211	0.289 ± 0.602	0.6318	960	0.211	0.095 ± 0.577	0.8687
					Diastolic blood pressure, mm Hg	Winsorization	960	0.211	–0.052 ± 0.456	0.9087	960	0.211	–0.215 ± 0.432	0.6190
rs7754840	*CDKAL1*	6	20,661,250	C/G	Fasting glucose, mmol/L	Winsorization	961	0.374	0.020 ± 0.016	0.2273	960	0.374	0.018 ± 0.016	0.2448
					1-hr glucose, mmol/L	Winsorization	960	0.375	0.174 ± 0.077	0.0245	959	0.375	0.170 ± 0.076	0.0258
					2-hr glucose, mmol/L	Winsorization	961	0.374	0.097 ± 0.061	0.1147	960	0.374	0.093 ± 0.060	0.1237
					AUC_{glu} at 0–120 min	Winsorization	960	0.375	13.80 ± 6.142	0.0249	959	0.375	13.39 ± 6.013	0.0261
					HbA1c, %	Winsorization	886	0.372	0.008 ± 0.017	0.6306	886	0.372	0.007 ± 0.017	0.6950
					HOMA2-β	Natural logarithm	956	0.374	–0.001 ± 0.010	0.9074	955	0.374	–0.003 ± 0.010	0.7912
					HOMA2-IR	Natural logarithm	956	0.374	0.010 ± 0.016	0.5484	955	0.374	0.007 ± 0.015	0.6258
					Fasting C-peptide, pmol/L	Natural logarithm	956	0.374	0.009 ± 0.016	0.5881	955	0.374	0.006 ± 0.014	0.6739
					Body mass index, kg/m^2	Winsorization	960	0.374	0.049 ± 0.140	0.7276	-	-	-	-
					Gestational weight gain, kg/wk	Winsorization	953	0.375	0.010 ± 0.006	0.1084	953	0.375	0.009 ± 0.006	0.1223
					Systolic blood pressure, mm Hg	Winsorization	960	0.374	0.007 ± 0.458	0.9877	960	0.374	–0.040 ± 0.438	0.9277
					Diastolic blood pressure, mm Hg	Winsorization	960	0.374	–0.181 ± 0.347	0.6025	960	0.374	–0.220 ± 0.328	0.5027
rs2237897	*INS-IGF2-KCNQ1*	11	2,858,546	C/T	Fasting glucose, mmol/L	Winsorization	961	0.651	0.033 ± 0.017	0.0501	960	0.651	0.031 ± 0.016	0.0577
					1-hr glucose, mmol/L	Winsorization	960	0.651	0.244 ± 0.080	2.4×10^{-3}	959	0.651	0.237 ± 0.079	2.7×10^{-3}
					2-hr glucose, mmol/L	Winsorization	961	0.651	0.244 ± 0.063	1.2×10^{-4}	960	0.651	0.239 ± 0.062	1.4×10^{-4}
					AUC_{glu} at 0–120 min	Winsorization	960	0.651	23.08 ± 6.349	2.9×10^{-4}	959	0.651	22.49 ± 6.219	3.2×10^{-4}
					HbA1c, %	Winsorization	886	0.650	0.031 ± 0.018	0.0840	886	0.650	0.030 ± 0.018	0.0920
					HOMA2-β	Natural logarithm	956	0.651	0.013 ± 0.011	0.2164	955	0.651	0.010 ± 0.010	0.2965
					HOMA2-IR	Natural logarithm	956	0.651	0.045 ± 0.017	9.0×10^{-3}	955	0.651	0.039 ± 0.015	9.8×10^{-3}
					Fasting C-peptide, pmol/L	Natural logarithm	956	0.651	0.042 ± 0.017	0.0109	955	0.651	0.037 ± 0.015	0.0121
					Body mass index, kg/m^2	Winsorization	960	0.651	0.077 ± 0.145	0.5936	-	-	-	-
					Gestational weight gain, kg/wk	Winsorization	953	0.651	–0.007 ± 0.006	0.2430	953	0.651	-0.008 ± 0.006	0.1821
					Systolic blood pressure, mm Hg	Winsorization	960	0.651	0.459 ± 0.475	0.3342	960	0.651	0.385 ± 0.455	0.3976
					Diastolic blood pressure, mm Hg	Winsorization	960	0.651	0.390 ± 0.360	0.2788	960	0.651	0.328 ± 0.341	0.3364

(*Continued to the next page*)

Supplementary Table 13. Continued

SNP	Nearest gene(s)	Chr	Position (Build 37)	Risk/non-risk allele	Trait measured during pregnancy	Transformation	Model I: PCs and age				Model II: PCs, age, and BMI			
							No.	Risk allele frequency	β ± SE	*P* value	No.	Risk allele frequency	β ± SE	*P* value
rs7945617	*MTNR1B*	11	92,700,287	C/T	Fasting glucose, mmol/L	Winsorization	961	0.490	0.059 ± 0.017	3.5×10^{-4}	960	0.490	0.057 ± 0.016	4.5×10^{-4}
					1-hr glucose, mmol/L	Winsorization	960	0.490	0.250 ± 0.079	1.5×10^{-3}	959	0.490	0.242 ± 0.078	1.8×10^{-3}
					2-hr glucose, mmol/L	Winsorization	961	0.490	0.241 ± 0.062	1.1×10^{-4}	960	0.490	0.233 ± 0.061	1.5×10^{-4}
					AUC_{glu} at 0–120 min	Winsorization	960	0.490	24.04 ± 6.233	1.2×10^{-4}	959	0.490	23.25 ± 6.111	1.5×10^{-4}
					HbA1c, %	Winsorization	886	0.490	0.007 ± 0.018	0.6731	886	0.490	0.004 ± 0.018	0.8152
					HOMA2-β	Natural logarithm	956	0.491	−0.026 ± 0.010	0.0119	955	0.490	−0.030 ± 0.010	2.4×10^{-3}
					HOMA2-IR	Natural logarithm	956	0.491	0.005 ± 0.017	0.7813	955	0.490	−0.002 ± 0.015	0.8797
					Fasting C-peptide, pmol/L	Natural logarithm	956	0.491	0.00004 ± 0.016	0.9980	955	0.490	−0.007 ± 0.014	0.6411
					Body mass index, kg/m^2	Winsorization	960	0.490	0.110 ± 0.143	0.4415	-	-	-	-
					Gestational weight gain, kg/wk	Winsorization	953	0.490	0.006 ± 0.006	0.3395	953	0.490	0.005 ± 0.006	0.4484
					Systolic blood pressure, mm Hg	Winsorization	960	0.490	0.239 ± 0.468	0.6092	960	0.490	0.133 ± 0.447	0.7656
					Diastolic blood pressure, mm Hg	Winsorization	960	0.490	−0.022 ± 0.354	0.9500	960	0.490	−0.111 ± 0.335	0.7408

β and SE were reported according to the GDM-related risk allele. *P* was obtained from linear regression model with the adjustments of covariates.
HAPO-HK, Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong; SNP, single nucleotide polymorphism; Chr, chromosome; PC, principal component; SE, standard error; BMI, body mass index; *TBR1*, T-box brain transcription factor 1; *SLC4A10*, solute carrier family 4 member 10; AUC_{glu}, glucose area under the curve during OGTT; HbA1c, glycosylated hemoglobin; HOMA2-β, homeostatic model assessment 2 of β-cell function; HOMA2-IR, homeostatic model assessment 2 of insulin resistance; *CDKAL1*, CDK5 regulatory subunit-associated protein 1-like 1; *INS-IGF2*, insulin-insulin-like growth factor 2; *KCNQ1*, potassium voltage-gated channel subfamily Q member 1; *MTNR1B*, melatonin receptor 1B.

Supplementary Table 14. Sensitivity analysis for the association of identified variants with the risk of gestational diabetes in Chinese women from the HAPO-HK Study, with different covariates adjustments

SNP	Nearest gene(s)	Chr	Position (Build 37)	Risk/non-risk allele	Adjustment	Number	Risk allele frequency	OR (95% CI)	*P* value
rs117781972	*TBR1-SLC4A10*	2	162,332,226	A/G	PCs, age	960	0.212	2.26 (1.64–3.11)	5.1×10^{-7}
					PCs, age, BMI	959	0.211	2.25 (1.63–3.10)	7.4×10^{-7}
					PCs, age, BMI, GWG	952	0.211	2.28 (1.65–3.16)	5.3×10^{-7}
					PCs, age, BMI, GWG, SBP, DBP	952	0.211	2.30 (1.66–3.19)	5.1×10^{-7}
					PCs, age, BMI, GWG, SBP, DBP, smoking status	952	0.211	2.30 (1.66–3.18)	5.9×10^{-7}
					PCs, age, BMI, GWG, SBP, DBP, smoking status, education year	927	0.212	2.32 (1.67–3.22)	5.6×10^{-7}
					PCs, age, BMI, GWG, SBP, DBP, smoking status, education year, parity	927	0.212	2.32 (1.67–3.22)	5.8×10^{-7}
					PCs, age, BMI, GWG, SBP, DBP, smoking status, education year, parity, family history of diabetes	921	0.212	2.33 (1.67–3.24)	5.4×10^{-7}
					PCs, age, BMI, GWG, SBP, DBP, smoking status, education year, parity, family history of diabetes and hypertension	921	0.212	2.29 (1.64–3.18)	9.4×10^{-7}
rs7754840	*CDKAL1*	6	20,661,250	C/G	PCs, age	960	0.375	1.47 (1.14–1.90)	3.4×10^{-3}
					PCs, age, BMI	959	0.375	1.47 (1.13–1.91)	3.6×10^{-3}
					PCs, age, BMI, GWG	952	0.375	1.47 (1.13–1.91)	4.2×10^{-3}
					PCs, age, BMI, GWG, SBP, DBP	952	0.375	1.47 (1.13–1.92)	4.2×10^{-3}
					PCs, age, BMI, GWG, SBP, DBP, smoking status	952	0.375	1.47 (1.12–1.91)	4.7×10^{-3}
					PCs, age, BMI, GWG, SBP, DBP, smoking status, education year	927	0.373	1.49 (1.14–1.95)	3.2×10^{-3}
					PCs, age, BMI, GWG, SBP, DBP, smoking status, education year, parity	927	0.373	1.51 (1.16–1.97)	2.6×10^{-3}
					PCs, age, BMI, GWG, SBP, DBP, smoking status, education year, parity, family history of diabetes	921	0.374	1.50 (1.15–1.96)	3.2×10^{-3}
					PCs, age, BMI, GWG, SBP, DBP, smoking status, education year, parity, family history of diabetes and hypertension	921	0.374	1.50 (1.15–1.97)	3.1×10^{-3}
rs2237897	*INS-IGF2-KCNQ1*	11	2,858,546	C/T	PCs, age	960	0.349	1.76 (1.31–2.36)	1.7×10^{-4}
					PCs, age, BMI	959	0.349	1.77 (1.31–2.37)	1.7×10^{-4}
					PCs, age, BMI, GWG	952	0.349	1.76 (1.31–2.37)	2.1×10^{-4}
					PCs, age, BMI, GWG, SBP, DBP	952	0.349	1.76 (1.30–2.37)	2.4×10^{-4}
					PCs, age, BMI, GWG, SBP, DBP, smoking status	952	0.349	1.77 (1.31–2.39)	2.0×10^{-4}
					PCs, age, BMI, GWG, SBP, DBP, smoking status, education year	927	0.347	1.79 (1.32–2.42)	1.7×10^{-4}
					PCs, age, BMI, GWG, SBP, DBP, smoking status, education year, parity	927	0.347	1.80 (1.33–2.43)	1.5×10^{-4}
					PCs, age, BMI, GWG, SBP, DBP, smoking status, education year, parity, family history of diabetes	921	0.348	1.79 (1.32–2.42)	1.8×10^{-4}
					PCs, age, BMI, GWG, SBP, DBP, smoking status, education year, parity, family history of diabetes and hypertension	921	0.348	1.80 (1.33–2.44)	1.6×10^{-4}
rs7945617	*MTNR1B*	11	92,700,287	C/T	PCs, age	960	0.490	1.61 (1.23–2.12)	6.0×10^{-4}
					PCs, age, BMI	959	0.490	1.59 (1.21–2.09)	9.0×10^{-4}
					PCs, age, BMI, GWG	952	0.490	1.59 (1.21–2.09)	1.0×10^{-3}
					PCs, age, BMI, GWG, SBP, DBP	952	0.490	1.56 (1.18–2.07)	1.6×10^{-3}
					PCs, age, BMI, GWG, SBP, DBP, smoking status	952	0.490	1.56 (1.18–2.07)	1.6×10^{-3}
					PCs, age, BMI, GWG, SBP, DBP, smoking status, education year	927	0.488	1.58 (1.19–2.09)	1.5×10^{-3}

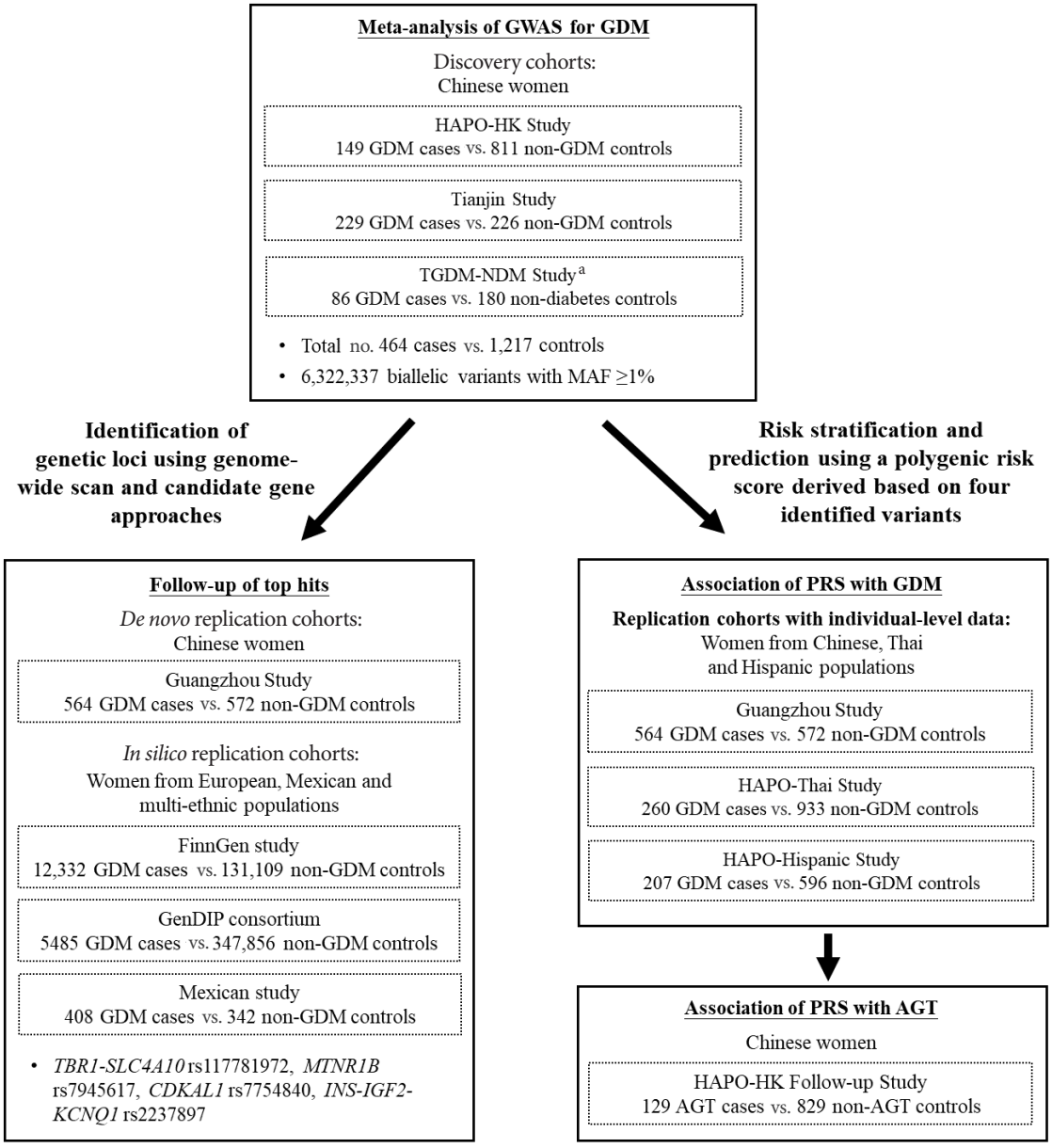
(*Continued to the next page*)

Supplementary Table 14. Continued

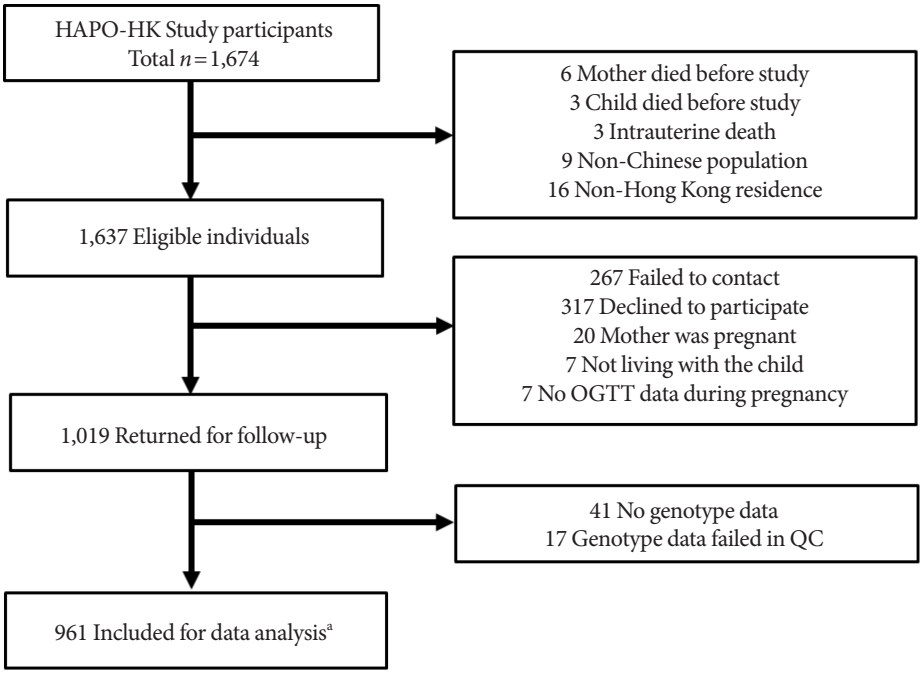
SNP	Nearest gene(s)	Chr	Position (Build 37)	Risk/non-risk allele	Adjustment	Number	Risk allele frequency	OR (95% CI)	*P* value
					PCs, age, BMI, GWG, SBP, DBP, smoking status, education year, parity	927	0.488	1.58 (1.19–2.09)	1.5×10^{-3}
					PCs, age, BMI, GWG, SBP, DBP, smoking status, education year, parity, family history of diabetes	921	0.489	1.58 (1.20–2.10)	1.4×10^{-3}
					PCs, age, BMI, GWG, SBP, DBP, smoking status, education year, parity, family history of diabetes and hypertension	921	0.489	1.56 (1.18–2.07)	2.0×10^{-3}

OR and 95% CI were reported according to the GDM-related risk allele. *P* value was obtained from logistic regression model with the adjustments of covariates.

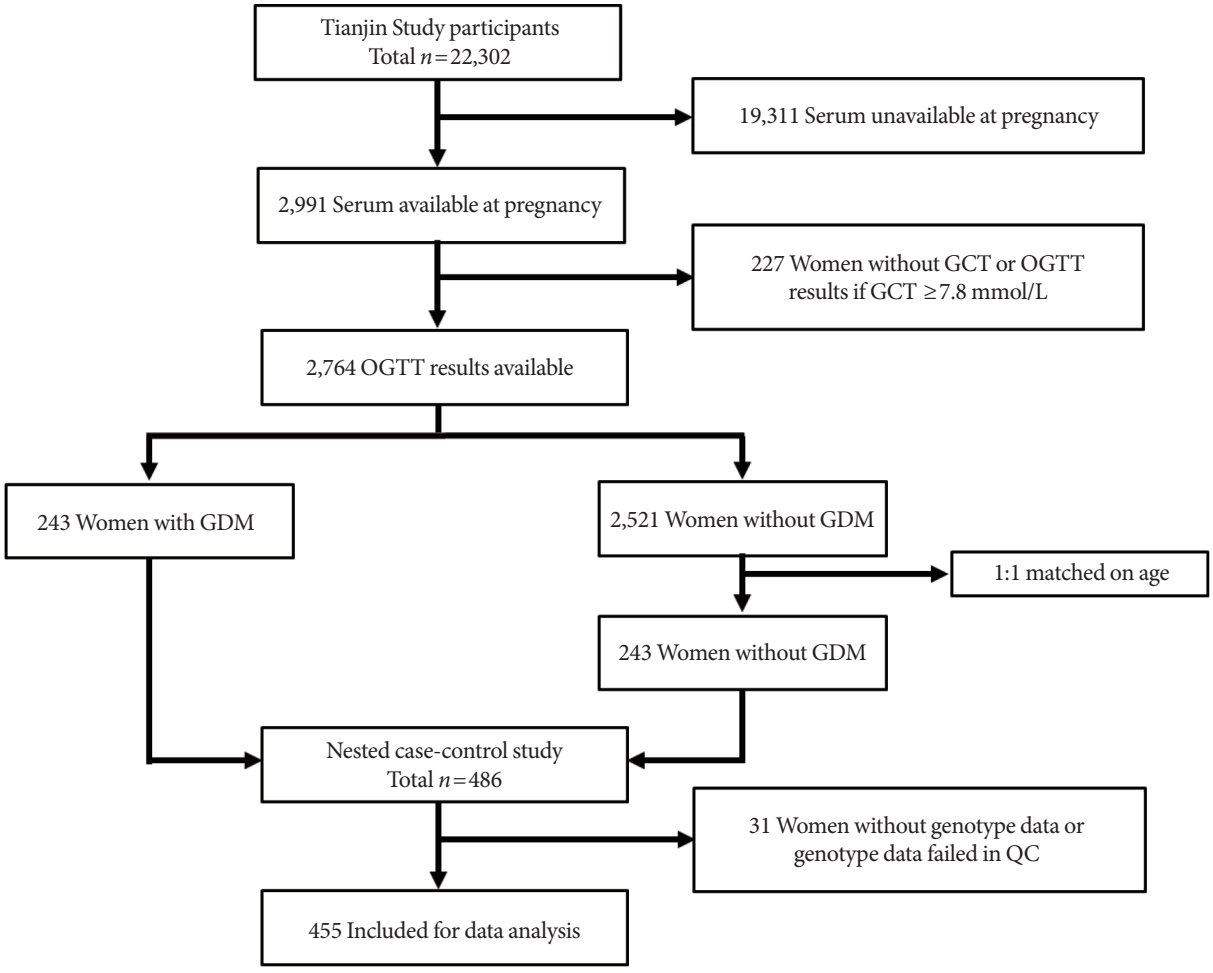
HAPO-HK, Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong; SNP, single nucleotide polymorphism; Chr, chromosome; OR, odds ratio; CI, confidence interval; *TBR1*, T-box brain transcription factor 1; *SLC4A10*, solute carrier family 4 member 10; PC, principal component; BMI, body mass index; GWG, gestational weight gain; SBP, systolic blood pressure; DBP, diastolic blood pressure; *CDKAL1*, CDK5 regulatory subunit-associated protein 1-like 1; *INS-IGF2*, insulin-insulin-like growth factor 2; *KCNQ1*, potassium voltage-gated channel subfamily Q member 1; *MTNR1B*, melatonin receptor 1B.



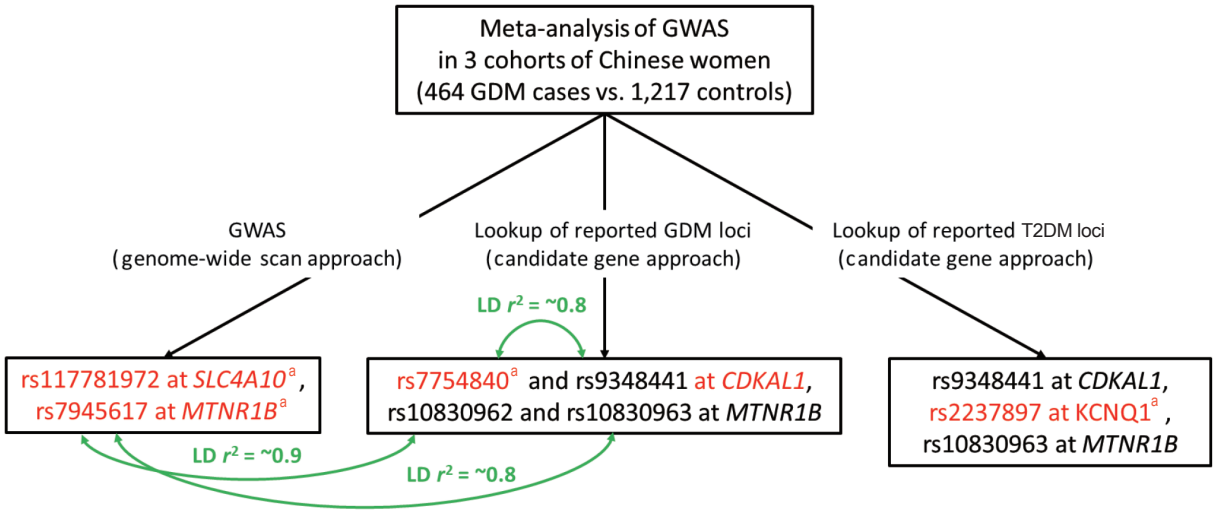
Supplementary Fig. 1. Study design and workflow. Step 1: To identify novel loci associated with an increased risk of gestational diabetes mellitus (GDM), we conducted a meta-analysis of three genome-wide association studies in Chinese women. This was followed by *de novo* replication in an independent Chinese cohort and *in silico* replications in European, Mexican and multi-ethnic populations. Through a combination of genome-wide scan and candidate gene approaches, we identified four loci associated with GDM. Step 2: In order to explore the potential clinical utility of personal genetic information, we derived a polygenic risk score (PRS) for GDM based on the four identified variants from step 1. We evaluated the predictive value of this PRS for GDM in Chinese, Thai and Hispanic populations. Additionally, we assessed its predictive value for abnormal glucose tolerance (AGT) at 7-year postpartum in a Chinese population. GWAS, genome-wide association studies; HAPO-HK, Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong; MAF, minor allele frequency; GenDIP, GENetics of Diabetes In Pregnancy; *TBR1*, T-box brain transcription factor 1; *SLC4A10,* solute carrier family 4 member 10; *MTNR1B*, melatonin receptor 1B; *CDKAL1*, CDK5 regulatory subunit-associated protein 1-like 1; *INS-IGF2*, insulin-insulin-like growth factor 2; *KCNQ1*, potassium voltage-gated channel subfamily Q member 1. [a]Women in the control group of the Treated GDM Cases vs. Non-diabetes Controls (TGDM-NDM) Study were non-diabetic and non-pregnant, and their glucose tolerance status was not assessed during pregnancy.



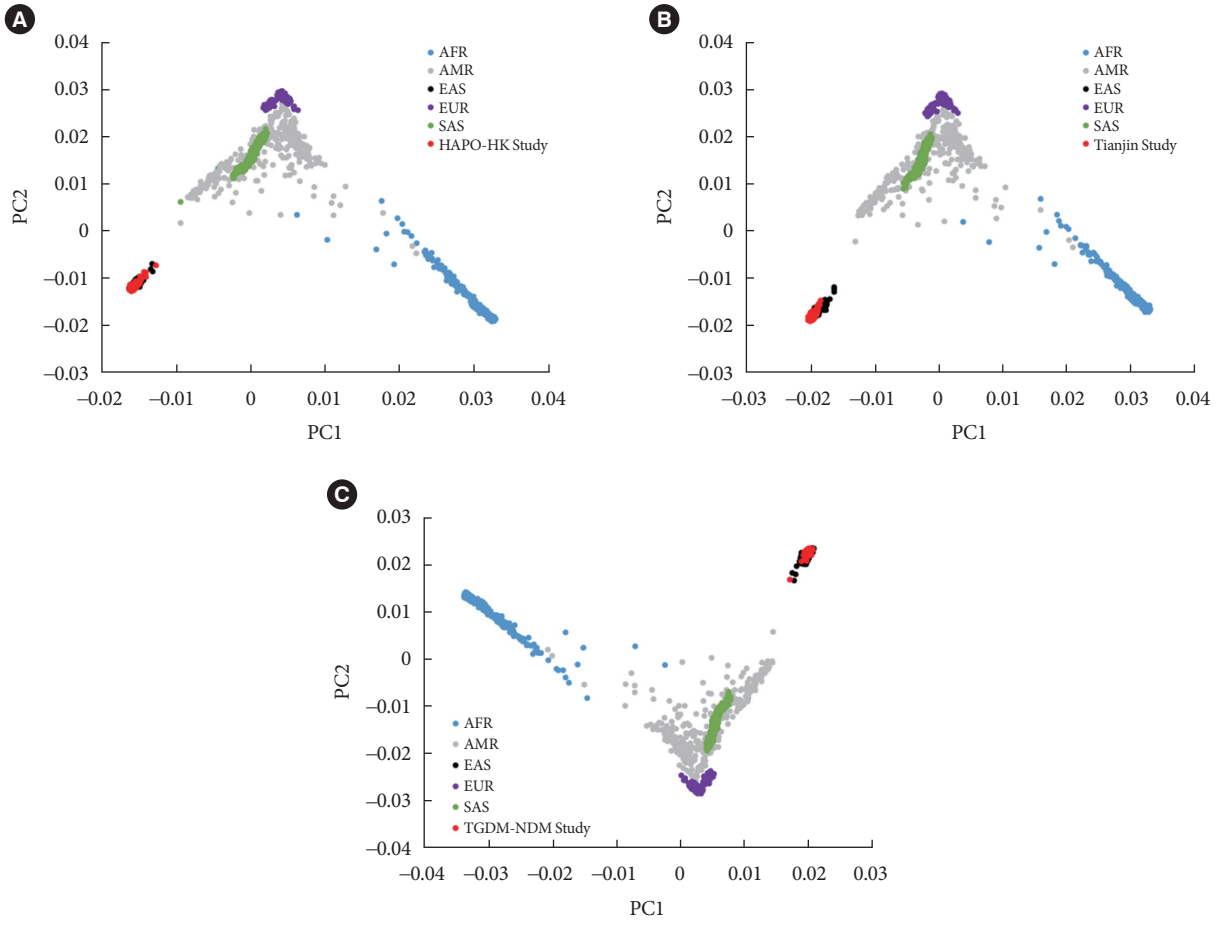
Supplementary Fig. 2. Number of individuals included in the Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong (HAPO-HK) Study. QC, quality control. [a]A total of 17 women who were unblinded to the oral glucose tolerance test (OGTT) results were included in data analysis.



Supplementary Fig. 3. Number of individuals included in the Tianjin Study. GCT, glucose challenge test; OGTT, oral glucose tolerance test; GDM, gestational diabetes mellitus; QC, quality control.



Supplementary Fig. 4. Summary of identified variants for gestational diabetes mellitus (GDM) in the present study. GWAS, genome-wide association studies; T2DM, type 2 diabetes mellitus; *SLC4A10*, solute carrier family 4 member 10; *MTNR1B*, melatonin receptor 1B; *CDKAL1*, CDK5 regulatory subunit-associated protein 1-like 1; *KCNQ1*, potassium voltage-gated channel subfamily Q member 1. [a]Variants highlighted in red colour were selected for the construction of a polygenic risk score. The linkage disequilibrium (LD) r^2 between variants was estimated using the data of the Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong (HAPO-HK) Study, the Tianjin Study, and the Treated GDM Cases vs. Non-diabetes Controls (TGDM-NDM) Study (Supplementary Table 8).



Supplementary Fig. 5. Principal component analysis (PCA) in (A) the Hyperglycemia and Adverse Pregnancy Outcome-Hong Kong (HAPO-HK) Study, (B) the Tianjin Study, and (C) the Treated GDM Cases vs. Non-diabetes Controls (TGDM-NDM) Study. The PCA plots show the first two principal components (PCs), based on genotype data of 26 different populations from the 1000 Genomes Project, as well as each discovery cohort of Chinese women (A: 961 women from the HAPO-HK Study; B: 455 women from the Tianjin Study; and C: 266 women from the TGDM-NDM Study). The 26 populations from the 1000 Genomes Project have been divided into five super populations: African (AFR) includes Yoruba in Ibadan, Nigeria, Luhya in Webuye, Kenya, Gambian in Western Divisions in the Gambia, Mende in Sierra Leone, Esan in Nigeria, Americans of African Ancestry in SW USA, and African Caribbeans in Barbados; Ad Mixed American (AMR) includes Mexican Ancestry from Los Angeles USA, Puerto Ricans from Puerto Rico, Colombians from Medellin, and Colombia, Peruvians from Lima, Peru; South Asian (SAS) includes Gujarati Indian from Houston, Texas, Punjabi from Lahore, Pakistan, Bengali from Bangladesh, Sri Lankan Tamil from the UK, and Indian Telugu from the UK; European (EUR) includes Utah Residents (CEPH) with Northern and Western European Ancestry, Toscani in Italia, Finnish in Finland, British in England and Scotland, and Iberian Population in Spain; and East Asian (EAS) includes Han Chinese in Beijing, China, Japanese in Tokyo, Japan, Southern Han Chinese, Chinese Dai in Xishuangbanna, China, and Kinh in Ho Chi Minh City, Vietnam.