Knowledge regarding the genetic risk loci for gestational diabetes mellitus (GDM) is still limited. In this study, we performed a two-stage genome-wide association analysis in Korean women. In the stage 1 genome scan, 468 women with GDM and 1,242 nondiabetic control women were compared using 2.19 million genotyped or imputed markers. We selected 11 loci for further genotyping in stage 2 samples of 931 case and 783 control subjects. The joint effect of stage 1 plus stage 2 studies was analyzed by meta-analysis. We also investigated the effect of known type 2 diabetes variants in GDM. Two loci known to be associated with type 2 diabetes had a genome-wide significant association with GDM in the joint analysis. rs7754840, a variant in CDKAL1, had the strongest association with GDM (odds ratio 1.518; \( P = 6.65 \times 10^{-10} \)). A variant near MTNR1B, rs10830962, was also significantly associated with the risk of GDM (1.454; \( P = 2.49 \times 10^{-13} \)). We found that there is an excess of association between known type 2 diabetes variants and GDM above what is expected under the null hypothesis. In conclusion, we have confirmed that genetic variants in CDKAL1 and near MTNR1B are strongly associated with GDM in Korean women. There seems to be a shared genetic basis between GDM and type 2 diabetes.

There has been a marked increase in our understanding of the genetic predisposition to type 2 diabetes as a result of the technical advances in array-based genotyping and the knowledge derived from the Human Genome Project (1). Recent genome-wide association (GWA) studies and meta-analyses, including the Diabetes Genetics Replication and Meta-analysis+ (DIAGRAM+) Study (2), the Meta-analyses including the Diabetes Genetics Replication and Meta-analysis (DIAGRAM) Study (2), and the recent GWA studies of type 2 diabetes in Asians (4,5), enlisted up to 41 genetic risk loci for type 2 diabetes or glycemic traits (2,3,6–10). However, these loci explain only a limited part of the expected heritability of type 2 diabetes (11). A complementary approach to improve our insight into the genetics of diabetes might involve the identification of genetic risk loci in a different subtype of diabetes, such as gestational diabetes mellitus (GDM). This approach could enable us to compare the genetic risk factors between type 2 diabetes and GDM and relate the similarities and dissimilarities to the pathophysiology of the two closely related diseases.

GDM is defined as abnormal glucose tolerance first recognized during pregnancy (12). The estimated prevalence of GDM varies according to ethnicity (13), and in Korean women, 2–5% of 100 pregnancies are complicated by GDM (14). During pregnancy, women are faced with increased adiposity and increased insulin resistance. The insulin resistance that develops during pregnancy is explained in part by the increased production of human placental lactogen, estrogen, and prolactin (15–17). Those who have limited \( \beta \)-cell capacity for the compensation of insulin resistance are likely to develop GDM (18). Women with GDM are assumed to have decreased \( \beta \)-cell insulin secretory function similar to type 2 diabetes (18). After parturition, nearly one-half of these women progress to type 2 diabetes within 5 years (19–21). Therefore, GDM is often regarded as a herald of type 2 diabetes in later life.

Based on these findings, it can be hypothesized that GDM and type 2 diabetes share a common genetic background. We have previously reported that some of the genetic variants that are strongly associated with type 2 diabetes are similarly associated with GDM risk (22). However, genetic knowledge in the context of GDM is still limited, and systematic approaches such as GWA studies have not been applied to GDM. Thus, it is not known whether there are genetic risk loci specific to GDM. In this study, we performed a GWA study to investigate genetic risk factors for GDM in Korean women and we also compared the genetic basis of GDM and type 2 diabetes.
TABLE 1
Clinical characteristics of the study participants in the GDM GWA analysis

<table>
<thead>
<tr>
<th></th>
<th>GDM</th>
<th>Stage 1 genome scan</th>
<th>Stage 2 genome scan</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nondiabetic control subjects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>468</td>
<td>1.242</td>
<td>931</td>
</tr>
<tr>
<td>Age (years)</td>
<td>31.5 ± 4.0</td>
<td>59.1 ± 5.6</td>
<td>32.5 ± 4.0</td>
</tr>
<tr>
<td>BMI (kg/m^2)</td>
<td>23.3 ± 3.2</td>
<td>24.6 ± 3.2</td>
<td>25.0 ± 4.7</td>
</tr>
<tr>
<td>Systolic blood pressure (mmHg)</td>
<td>116 ± 13</td>
<td>122 ± 19</td>
<td>114 ± 13</td>
</tr>
<tr>
<td>Diastolic blood pressure (mmHg)</td>
<td>70 ± 9</td>
<td>77 ± 11</td>
<td>69 ± 9</td>
</tr>
<tr>
<td>Fasting plasma glucose (mmol/L)</td>
<td>5.4 ± 1.1</td>
<td>4.4 ± 0.4</td>
<td>5.2 ± 1.0</td>
</tr>
<tr>
<td>HbA1c (%)</td>
<td>NA</td>
<td>5.5 ± 0.3</td>
<td>4.9 ± 0.4</td>
</tr>
</tbody>
</table>

Data are means ± SD. Data for women with GDM were measured during the diagnostic 100-g oral glucose tolerance test. NA, not available.

**Stage 2 follow-up subjects.** A total of 931 women with GDM and 783 nondiabetic control women were included in the stage 2 follow-up study (Table 1). The GDM group consisted of two subgroups: 1) 426 women with GDM recruited from the Cheil General Hospital between January 1996 and February 2003 and 2) 505 women with GDM recruited as an independent cohort from the same hospital between March 2003 and February 2008. The diagnostic criteria of GDM in stage 2 were the same as those in stage 1. As control subjects, we enrolled nondiabetic women from four different institutes: 1) Seoul National University Hospital (n = 162), 2) Seoul National University Bundang Hospital (n = 96), 3) Kyungpook National University Hospital (n = 201), and 4) Korea National Institutes of Health (from the Health T2D Study) (n = 324). The criteria for selecting control subjects were as follows: age ≥60 years, no previous history of type 2 diabetes, no first-degree relatives with type 2 diabetes, fasting plasma glucose level <5.6 mmol/L, and HbA1c <6.0%.

The institutional review board of the Clinical Research Institute at Seoul National University Hospital approved the study protocol (institutional review board no. H-0412-138-017), and written informed consent was obtained from each subject. All clinical investigations were conducted according to the principles expressed in the Declaration of Helsinki.

**Biochemical measurements.** The anthropometric and metabolic phenotypes of the women with GDM were measured at the time of the 100-g oral glucose tolerance test. Plasma glucose concentration was determined using a human-specific radioimmunoassay kit (Linco Research, St. Charles, MO). The homeostasis model assessment (HOMA) was used to estimate insulin resistance (HOMA-IR) and β-cell function (HOMA-B) with paired fasting glucose and insulin concentrations (24).

**Genotyping.** We extracted genomic DNA from peripheral leukocytes. Genotyping for the stage 1 genome scan was performed using the Affymetrix Genome-Wide Human SNP Array 6.0 (25). Overall, the TaqMan genotyping success rate was 99.3%, and the concordance rate based on 15 blind duplicate comparisons was 100% in the stage 2 study. For the SNP genotyping in Health T2D Study, standard quality-control measures were applied and imputation was conducted using IMPUTE version 1.0 (http://mathgen.stats.ox.ac.uk/impute/impute.html) (27).

### Statistical analyses
Most of the association testing was performed using PLINK version 1.07 (http://pngu.mgh.harvard.edu/purcell/plink/) (28), and genotype imputation was conducted using MACH software (http://www.sph.umich.edu/csg/abecasis/mach/) (29). The Manhattan plot, which depicts the negative log10 of the p-values distribution of the observed p-values, was used to identify regions of significant association. The significance threshold for the stage 2 follow-up study was p < 0.05. A joint analysis of stage 1 plus stage 2 results was performed using METAL (http://www.sph.umich.edu/csg/abecasis/metal/), using an inverse-variance method assuming fixed effects (30). An additional association analysis was performed using the GnomAD data (https://gnomad.broadinstitute.org) (31). The QQ plot, which shows the distribution of the observed p-values of the logistic regression analysis against the expected distribution under the null hypothesis, was generated using the R statistical package (http://www.r-project.org/). The Manhattan plot, which depicts the negative log10 of p-values derived from the logistic regression analysis, was plotted against the chromosomal position using the R statistical package. The dense regional association results of stage 1, stage 2, and the joint analysis of stage 1 plus stage 2 were plotted using LocusZoom software (http://csg.sph.umich.edu/locuszoom/) (32). To compare the effect size of known type 2 diabetes variants in GDM and type 2 diabetes, the β-coefficient of the logistic regression analysis derived from our stage 1 genome scan and from the stage 1 meta-GWA analysis of the Asian Genetic Epidemiology Network (AGEN) type 2 diabetes report (33) was plotted.

### RESULTS

**Stage 1 genome scan.** In the stage 1 genome scan, logistic regression analysis using an additive genetic model was used to test for the association between the genotypes and GDM. The negative log10 of the p-values from the
association test were plotted against their genomic position in Fig. 1B. The association test results for SNPs with $P < 0.0001$ in the stage 1 genome scan are listed in Supplementary Table 1. A total of nine independent (pairwise linkage disequilibrium [LD] $r^2 < 0.5$) SNPs were suggestive of an association according to our predefined arbitrary threshold of $P < 2.0 \times 10^{-5}$. Among these, variants in CDKAL1 (rs7754840) and near MTNR1B (rs10830962) showed the strongest association with GDM risk. We added two additional variants (rs10757261 near CDKN2A/2B and rs10882066 in IDE) from the type 2 diabetes risk loci, which had suggestive $P$ values in our stage 1 results. Among the 11 SNPs, 4 were substituted to imputed SNPs (rs6499500, rs12715106, rs9395950, and rs187230), which had more significant $P$ values and had strong LD ($r^2 > 0.8$) with the original SNPs. To eliminate hidden population stratification and cryptic relatedness, the EMMAX, a variance component approach accounting for hidden sample structure, was used to test the association between genetic variants and GDM. The $P$ values of the EMMAX are listed in Supplementary Table 1. The results of the EMMAX were similar to those of the original logistic regression analysis. The Pearson correlation coefficient for the log$_{10}$ of the $P$ values of both analyses was 0.921 ($P = 1.51 \times 10^{-288}$). This implies that population stratification was minimal in our study samples.

**Stage 2 follow-up and joint analysis.** In stage 2 follow-up, we genotyped 11 SNPs in 931 case and 783 control subjects. The results of the stage 2 association test are shown in Table 2. Among these, rs7754840 in CDKAL1 (odds ratio [OR] 1.396 [95% CI 1.222–1.594]; $P = 2.90 \times 10^{-5}$), rs10830962 near MTNR1B (1.442 [1.250–1.651]; $P = 6.95 \times 10^{-5}$), rs1470579 in IGF2BP2 (1.236 [1.068–1.430]; $P = 0.0042$), and rs10882066 in IDE (1.203 [1.013–1.428]; $P = 0.035$) were significantly associated with GDM in the stage 2 samples. None of the other SNPs showed evidence...
## TABLE 2
Association between SNPs and GDM in stage 1, stage 2, and joint stage 1 plus stage 2 analysis

<table>
<thead>
<tr>
<th>SNP</th>
<th>Risk CHR</th>
<th>Risk allele</th>
<th>Nearest gene</th>
<th>Stage 1 genome scan</th>
<th>Stage 2 follow-up</th>
<th>Stage 1 plus stage 2</th>
<th>n for 80% power</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>n (GDM subjects/control subjects) = 468/1,242</td>
<td>n (GDM subjects/control subjects) = 931/783</td>
<td>n (GDM subjects/control subjects) = 1,399/2,025</td>
<td></td>
</tr>
<tr>
<td>rs7754840</td>
<td>6</td>
<td>C</td>
<td>CDKAL1</td>
<td>0.575 0.445 1.707 (1.459–1.997) 2.15 × 10⁻¹¹ 0.556 0.467 1.396 (1.222–1.594) 2.90 × 10⁻⁷</td>
<td>1.518 (1.372–1.660) 6.65 × 10⁻¹⁶</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs10830962</td>
<td>11</td>
<td>G</td>
<td>MTNR1B</td>
<td>0.529 0.430 1.469 (1.266–1.705) 5.02 × 10⁻⁷ 0.537 0.444 1.236 (1.259–1.651) 6.95 × 10⁻⁸</td>
<td>1.332 (1.315–1.668) 2.49 × 10⁻¹³</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1470579</td>
<td>3</td>
<td>C</td>
<td>IGF2BP2</td>
<td>0.377 0.293 1.465 (1.266–1.705) 5.02 × 10⁻⁷ 0.337 0.291 1.396 (1.222–1.594) 2.90 × 10⁻⁷</td>
<td>1.332 (1.315–1.668) 2.49 × 10⁻¹³</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs6499500*</td>
<td>16</td>
<td>C</td>
<td>MTNR1B</td>
<td>0.377 0.293 1.465 (1.266–1.705) 5.02 × 10⁻⁷ 0.337 0.291 1.396 (1.222–1.594) 2.90 × 10⁻⁷</td>
<td>1.332 (1.315–1.668) 2.49 × 10⁻¹³</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs12898654</td>
<td>15</td>
<td>G</td>
<td>LBXCOR1</td>
<td>0.215 0.167 1.376 (1.136–1.667) 0.0011 0.210 0.181 1.203 (1.134–1.451) 1.81 × 10⁻⁴</td>
<td>2.853 (1.110–1.390) 1.59 × 10⁻⁴</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs9395950*</td>
<td>6</td>
<td>A</td>
<td>TINAG</td>
<td>0.770 0.703 1.376 (1.136–1.667) 0.0011 0.210 0.181 1.203 (1.134–1.451) 1.81 × 10⁻⁴</td>
<td>2.853 (1.110–1.390) 1.59 × 10⁻⁴</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs10882066</td>
<td>10</td>
<td>G</td>
<td>IDE/CDKN2A</td>
<td>0.215 0.167 1.376 (1.136–1.667) 0.0011 0.210 0.181 1.203 (1.134–1.451) 1.81 × 10⁻⁴</td>
<td>2.853 (1.110–1.390) 1.59 × 10⁻⁴</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs10757261</td>
<td>9</td>
<td>G</td>
<td>CDKN2A</td>
<td>0.732 0.671 1.330 (1.128–1.568) 0.0007 0.699 0.680 1.095 (0.946–1.267) 0.221</td>
<td>1.135 (1.070–1.331) 0.0015</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs7513574</td>
<td>1</td>
<td>G</td>
<td>CACHD</td>
<td>0.352 0.274 1.482 (1.262–1.740) 1.75 × 10⁻⁵ 0.277 0.283 0.973 (0.838–1.129) 0.711</td>
<td>1.232 (1.135–1.515) 4.74 × 10⁻⁵</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs12898654</td>
<td>15</td>
<td>G</td>
<td>LBXCOR1</td>
<td>0.215 0.167 1.376 (1.136–1.667) 0.0011 0.210 0.181 1.203 (1.134–1.451) 1.81 × 10⁻⁴</td>
<td>2.853 (1.110–1.390) 1.59 × 10⁻⁴</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs9395950*</td>
<td>6</td>
<td>A</td>
<td>TINAG</td>
<td>0.770 0.703 1.376 (1.136–1.667) 0.0011 0.210 0.181 1.203 (1.134–1.451) 1.81 × 10⁻⁴</td>
<td>2.853 (1.110–1.390) 1.59 × 10⁻⁴</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs10882066</td>
<td>10</td>
<td>G</td>
<td>IDE/CDKN2A</td>
<td>0.215 0.167 1.376 (1.136–1.667) 0.0011 0.210 0.181 1.203 (1.134–1.451) 1.81 × 10⁻⁴</td>
<td>2.853 (1.110–1.390) 1.59 × 10⁻⁴</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs8187304*</td>
<td>3</td>
<td>C</td>
<td>PLD1</td>
<td>0.731 0.655 1.439 (1.216–1.703) 2.33 × 10⁻⁵ 0.669 0.681 0.816 (0.616–1.096) 0.471</td>
<td>1.182 (1.060–1.319) 0.0027</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data are OR for the risk allele (95% CI), unless otherwise indicated. The risk allele is indexed to the positive strand of the National Center for Biotechnology Information build 36. The nearest gene is defined as the gene closest to the SNP or within the boundary of the 100-kb window from the SNP. RAF (GDM) and RAF (control) refer to the risk allele frequencies in women with GDM and nondiabetic control women, respectively. P values for stage 1 and stage 2 were calculated using logistic regression under an additive genetic model. P values for the joint stage 1 plus stage 2 were calculated by METAL using the inverse-variance method under a fixed-effect model. CHR, chromosome. *These were imputed SNPs using MACH. The number of samples required to achieve 80% power (\( \alpha = 5 \times 10^{-5} \)) was calculated based on the OR in the joint stage 1 plus stage 2 analysis, the risk allele frequency in stage 1 control subjects, and a GDM prevalence of 5%.
of an association. In the joint analysis of stage 1 plus stage 2 results (Table 2), rs7754840 in CDKAL1 (1.518 [1.372–1.680]; \(P = 6.65 \times 10^{-16}\)) and rs10830962 near MTNR1B (1.454 [1.315–1.608]; \(P = 2.49 \times 10^{-13}\)) reached genome-wide significance for an association with GDM. One additional variant, rs1470579 in IGF2BP2 (1.332 [1.197–1.484]; \(P = 1.67 \times 10^{-7}\)), showed a near genome-wide significant association with GDM. The regional association plot of SNPs near CDKAL1 and MTNR1B, including those that have been imputed, are depicted in Fig. 2.

**Association with glucose and insulin-related traits.** To obtain additional insight into the role of these two variants, we performed an association analysis between these variants and quantitative traits of fasting glucose, fasting

---

**FIG. 2.** Dense regional association plot near CDKAL1 (A) and MTNR1B (B). The hash marks above the panel represent the position of each SNP that was genotyped or imputed. The negative log10 of \(P\) values from logistic regression are shown in the panel. The blue diamond indicates the SNP with the most significant association in the stage 1 genome scan. The green and red diamonds represent the results of the SNP in stage 2 and joint stage 1 plus stage 2 analysis, respectively. Their corresponding \(P\) values are indicated on the right. Estimated recombination rates are plotted to reflect recombination hot spots. The SNPs in LD with the most significant SNP are color coded to represent their strength of LD. The locations of genes, exons and introns are shown in the lower panel (taken from the Human Genome hg18 build). (A high-quality color representation of this figure is available in the online issue.)
insulin, HOMA-IR, and HOMA-B in women with GDM and control subjects, separately (Table 3). The rs7754840 C allele of CDKAL1, which was the risk variant of GDM, was significantly associated with decreased fasting insulin concentration ($\beta = -0.026; P = 0.00051$) and decreased HOMA-B ($\beta = -0.034; P = 0.00085$) in women with GDM. The rs10830962 G allele near MTNR1B was nominally associated with decreased fasting insulin concentrations ($\beta = -0.018; P = 0.029$) in women with GDM. This variant was also marginally associated with increased fasting glucose concentrations ($\beta = 0.025; P = 0.041$) in control subjects.

Comparison of genetic risk loci between GDM and type 2 diabetes. Among 41 known type 2 diabetes loci, we were able to examine the association signals for 34 loci that were directly genotyped or imputed (Table 4). There was an excess of small $P$ values compared with the expected distribution under the null hypothesis (Fig. 3). When these known type 2 diabetes risk variants and markers in LD ($r^2 > 0.8$) were excluded, the QQ plot of our stage 1 genome scan was similar to that expected under the null hypothesis (Fig. 1A).

We compared the $\beta$-coefficient in the logistic regression analysis of known type 2 diabetes variants in GDM (our stage 1 genome scan) and type 2 diabetes (AGEN type 2 diabetes stage 1 meta-GWA results [26]) (Fig. 4 and Supplementary Table 2). There was a significant positive correlation between the $\beta$-coefficients of GDM and type 2 diabetes (Pearson correlation coefficient 0.442; $P = 0.0062$). However, some variants, such as rs10830963 of MTNR1B and rs11708067 of ADCY15, showed considerable differences in effect size between GDM and type 2 diabetes.

Finally, we examined the association results for rs7754840 of CDKAL1 and rs10830962 of MTNR1B in the stage 1 meta-GWA study of the AGEND type 2 diabetes report (28). The $C$ allele of rs7754840 was significantly associated with an increased risk of type 2 diabetes ($n = 18,732$; OR 1.20 [95% CI 1.14–1.25]; $P = 9.63 \times 10^{-15}$). However, the G allele of rs10830962 was not associated with type 2 diabetes risk ($n = 10,754$; 1.040 [0.98–1.11]; $P = 0.209$).

**DISCUSSION**

The main finding of this study is that variants in CDKAL1 and near MTNR1B are associated with GDM at a genome-wide significance level ($P < 5.0 \times 10^{-8}$). Our findings are confirmatory of previous candidate gene studies of GDM (22,34,35). Previously, we selected 18 SNPs, known to be associated with the risk of type 2 diabetes, for association testing in GDM (22). Among them, two SNPs (rs77560992 and rs7754840) in CDKAL1 were associated with GDM risk at a genome-wide significance level, and one SNP (rs10811661) near CDKN2A/2B also was strongly associated with GDM. Another study in Europeans tested 11 known SNPs of type 2 diabetes and found a modest association between GDM and TCPTP/R, CDKAL1, and TCF2 variants (34). Recently, Kim et al. (35) reported that two SNPs in MTNR1B (rs19387153 and rs10830963) were strongly associated with GDM in Korean women. The study by Kim et al. was led by one of our investigators but was performed independently of this study, and 505 subjects with GDM in that study overlapped with our stage 2 follow-up subjects. However, only a limited number of genetic variants known to be associated with type 2 diabetes have been tested thus far, and a systematic approach such as GWA analysis has not been applied to GDM until now. To the best of our knowledge, this is the first study to investigate the genetic association of GDM using dense SNP markers across the whole genome.

Variants in CDKAL1 are known to be strongly associated with type 2 diabetes (8–10,36,37). In our previous studies, the variant rs7754840 in CDKAL1 was significantly associated with GDM risk in 1,501 Koreans (23) as well as type 2 diabetes in 6,719 Asians including Koreans (38). The mechanism by which this variant confers type 2 diabetes and GDM risk is not yet clearly understood. CDKAL1 (cyclin-dependent kinase 5 [CDK5] regulatory subunit-associated protein 1-like 1) has been suggested to interact with CDK5 because it has homology with CDK5RAP1, a known inhibitor of CDK5 (39). CDK5 is expressed in pancreatic $\beta$-cells, and it has recently been reported to promote $\beta$-cell survival (40). Pregnancy is a state in which increased insulin resistance results in stress that adversely affects $\beta$-cell survival (41). Therefore, it may be postulated that variants in CDKAL1 might alter the function of CDK5 in $\beta$-cell compensation during pregnancy. The rs7754840 C variant of CDKAL1 was significantly associated with decreased insulin concentration and HOMA-B in our subjects with GDM, which implies compromised $\beta$-cell compensation. It also should be noted that variants in CDKAL1 are associated with decreased birth weight, which could be explained by reduced fetal insulin secretion (42).

Variants in MTNR1B were first identified as genetic risk factors strongly associated with elevated fasting glucose (43–45). Although it also was associated with type 2 diabetes risk in Europeans, the strength of association was rather weak compared with its association with fasting glucose (44,45). It is interesting that the rs10830962 variant in MTNR1B was the second most strongly associated loci in GDM but was not associated with type 2 diabetes risk in a large-scale meta-GWA study of East Asians (AGEN type 2 diabetes meta-GWA study [26]). The effect size of the MTNR1B rs10830962 variant was considerably different in our GDM GWA study (\(\beta = 0.374\)) compared with the AGEN type 2 diabetes meta-GWA study (\(\beta = 0.040\)). The risk allele of rs10830962 G was enriched in women with GDM (risk allele frequency 0.529 and 0.537 in stage 1 and stage 2 subjects, respectively) compared with AGEN type 2 diabetic subjects (risk allele frequency range for participating studies 0.40–0.46). It is not clear why this variant is more strongly associated with the risk of GDM compared with type 2 diabetes. However, it could be cautiously speculated that there might be differences in genetic susceptibility between type 2 diabetes and GDM conferred by the variants of MTNR1B. Additional investigations are required to confirm these findings.

It was recently revealed that MTNR1B is expressed in pancreatic $\beta$-cells (44). Melatonin is thought to inhibit insulin secretion from $\beta$-cells through the activation of MTNR1B, which is a G-protein–coupled receptor (46). The rs10830962 variant is in strong LD ($r^2 = 0.98$, in our study subjects) with rs10830963, which originally was reported to be associated with fasting glucose concentrations, insulin concentrations, and type 2 diabetes. It is interesting that the rs1083092 variant is located 4.36 kb upstream of MTNR1B and in the vicinity of the region where monomethylation of lysine 4 of histone 3 (H3K4me1), known as an enhancer-associated histone mark, was enriched in the Encyclopedia of DNA Elements database (http://genome.uchsc.edu/encode/) (47).

It would be of worth to study the functional role of each variant as well as to search for the functional role of each variant near this locus, especially in the promoter and/or enhancer region.
### Table 3

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Stage 1</th>
<th>Stage 1 Plus Stage 2</th>
<th>( \beta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>log10(HOMA-B)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fasting glucose (mmol/L)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDKAL1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CACNA1G</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD226</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Association**

Values for joint stage 1 plus stage 2, number of subjects with nonmissing phenotypes. Regression coefficients from linear regression models with adjustment for age and sex. The significance values in parentheses are adjusted for multiple comparisons. \( \beta \) values indicate the change in the phenotype for a one-unit change in the genotype.
Because we performed a GWA analysis using 2.19 million genotyped or imputed markers, we were able to test the hypothesis that GDM and type 2 diabetes might share a similar genetic background. Among 11 variants listed in Table 2, five SNPs were located near the known type 2 diabetes loci. Three variants that showed an association at or near genome-wide significance, rs7754840 in CDKAL1, rs10830962 near MTNRI1B, and rs17407579 in IGFBP2, were identical or in strong LD with known type 2 diabetes variants. Two other variants, rs10757261 in CDKN2A/2B and rs10882066 in IDE, were not in LD with known type 2 diabetes variants ($r^2 = 0.013$ between rs10757261 and rs10882066 and $r^2 = 0.002$ between rs10882066 and rs5015480). Given that many of the type 2 diabetes risk loci were associated with GDM, unbiased GWA studies of GDM showed genome-wide significant associations confined to known type 2 diabetes loci, and the effect size of the type 2 diabetes risk variants in GDM and type 2 diabetes were mostly comparable, it would be acceptable to state that GDM and type 2 diabetes share a similar genetic background.

Among the known type 2 diabetes–related genes, those that are thought to modulate pancreatic ß cell function were preferentially associated with GDM (CDKAL1, MTNRI1B, IGFBP2, CDKN2A/2B, SLC30A8, IDE, KCNQ1, and CENTD2) (Table 4). In contrast, loci relevant to insulin resistance, such as FTO, PPARG, IRS1, KLF14, and GCRK were not significantly associated with GDM (Table 4).
These findings support the previous notion that defective pancreatic β-cell compensation to overcome increased insulin resistance during pregnancy might be the core pathophysiology of GDM.

There are certain limitations in our study. First, the sample size was relatively modest for a GWA, and we had limited power in detecting possible variants with small effect size. Assuming a 2.2% prevalence of GDM (48), a disease allele frequency of 0.3, and a genotype relative risk of 1.5 in a multiplicative model, our study had 93% power. However, when genotype relative risk was assumed to be 1.4, the power dropped to 68%. It is possible that variants with a small effect would not have been detected in our study. Therefore, we might have missed genetic variants that have a small effect but are specific to GDM. Regarding the study design, the control group was not fully evaluated for their glucose tolerance status during pregnancy and parity. It is possible that women with GDM might have been included in the control group. However, because the prevalence of GDM was estimated to be 2.2% in Korean women (48), and approximately one-half of these women progress to type 2 diabetes within 10 years (22), only ~1% of women with GDM would have been included in the control group. Therefore, we speculate that the ascertainment bias would have been minimal. In studying genetic variants of GDM, the ideal control group would be pregnant women with normal glucose tolerance. However, we were only able to use nondiabetic control women as control subjects. This could be one of the reasons that variants in CDKAL1 and MTNR1B were most significantly associated with GDM in our study. In this regard, the results of our study should be interpreted cautiously.

In conclusion, we have performed a GWA study in the context of GDM for the first time and confirmed that variants in CDKAL1 and MTNR1B confer the risk of GDM with genome-wide significance. By comparing genetic variants in GDM and type 2 diabetes, we provide evidence that they share a similar genetic background. We hope that even larger GWA studies of GDM and meta-GWA studies could further reveal the genetic risk loci for GDM.

FIG. 3. QQ plot of the association between known type 2 diabetes variants and GDM. Comparison of the effect size of known type 2 diabetes variants in GDM and type 2 diabetes. The gray zone indicates the 95% CI. The known type 2 diabetes variants tested for association are listed in Table 4.

FIG. 4. Comparison of the effect size of known type 2 diabetes variants in GDM and type 2 diabetes. Effect size (β-coefficient from logistic regression analysis) of the known type 2 diabetes variants in GDM (y-axis) and type 2 diabetes (x-axis) are plotted with their corresponding P values (A: P values in GDM; B: P values in type 2 diabetes): red, P < 0.0001; orange, 0.0001 ≤ P < 0.01; yellow, 0.01 ≤ P < 0.10; green, 0.10 ≤ P < 0.50; blue, 0.50 ≤ P. The β-coefficient for GDM was derived from our stage 1 genome scan and that for type 2 diabetes was derived from the AGEN type 2 diabetes study (26). The two CDC123/CAMKID variants are distinguished by CDC123/CAMKID for rs10906115 and CDC123/CAMKID* for rs12779790. The two KCNQ1 variants are distinguished by KCNQ1 for rs231362 and KCNQ1* for rs2237892. (A high-quality color representation of this figure is available in the online issue.)
ACKNOWLEDGMENTS

This work was supported by the Korea Health 21 R&D Project, Korean Ministry of Health and Welfare (grant no. 00-PJ3-PG6-GN07-001); the Korea Healthcare Technology R&D Project, Korean Ministry of Health and Welfare (grant no. A102041); and the World Class University Project of the Ministry of Education, Science and Technology and the Korea Science and Engineering Foundation (grant no. R31-2008-000-10103-0). This work also was supported by grants from the Korea Center for Disease Control and Prevention (grant nos. 4845-301, 4851-302, and 4851-307) and an intramural grant from the Korea National Institutes of Health (grant no. 2010-N73002-00), Republic of Korea.

No potential conflicts of interest relevant to this article were reported.

S.H.K. researched data, contributed to the discussion, and drafted the manuscript. S.-H.K., N.H.C., I.K.L., and B.-G.H. provided study samples. Y.M.C., S.H.C., M.K.M., H.S.J., H.D.S., and S.Y.C. provided study samples. Y.S.C. provided study samples and researched data. H.C.J. and K.S.P. provided statistical advice, contributed to the discussion, and edited the manuscript. H.C.J. and K.S.P. provided data. Y.C. provided study samples and researched data. Y.S.C. provided study samples and researched data. H.C.J. and K.S.P. are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

REFERENCES

8. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:616–678