



Genetic Loci and Physiologic Pathways Involved in Gestational Diabetes Mellitus Implicated Through Clustering

Camille E. Powe,^{1,2,3} Miriam S. Udler,^{1,2,3} Sarah Hsu,^{1,2} Catherine Allard,⁴ Alan Kuang,⁵ Alisa K. Manning,^{2,3,6} Patrice Perron,⁷ Luigi Bouchard,^{4,8,9} William L. Lowe Jr.,¹⁰ Denise Scholtens,⁵ Jose C. Florez,^{1,2,3} and Marie-France Hivert^{1,3,7,11}

Diabetes 2021;70:268–281 | <https://doi.org/10.2337/db20-0772>

Hundreds of common genetic variants acting through distinguishable physiologic pathways influence the risk of type 2 diabetes (T2D). It is unknown to what extent the physiology underlying gestational diabetes mellitus (GDM) is distinct from that underlying T2D. In this study of >5,000 pregnant women from three cohorts, we aimed to identify physiologically related groups of maternal variants associated with GDM using two complementary approaches that were based on Bayesian nonnegative matrix factorization (bNMF) clustering. First, we tested five bNMF clusters of maternal T2D-associated variants grouped on the basis of physiology outside of pregnancy for association with GDM. We found that cluster polygenic scores representing genetic determinants of reduced β -cell function and abnormal hepatic lipid metabolism were associated with GDM; these clusters were not associated with infant birth weight. Second, we derived bNMF clusters of maternal variants on the basis of pregnancy physiology and tested these clusters for association with GDM. We identified a cluster that was strongly associated with GDM as well as associated with higher infant birth weight. The effect size for this cluster's association with GDM appeared greater than that for T2D. Our findings imply that the genetic and physiologic pathways that lead to GDM differ, at least in part, from those that lead to T2D.

Gestational diabetes mellitus (GDM) and type 2 diabetes (T2D) share common clinical features and have at least some shared genetic architecture (1–4). However, unlike T2D, GDM occurs on the background of dramatic pregnancy-related changes in glycemic physiology (5–8). These include marked increases in both insulin secretory response and insulin resistance as well as dynamic changes in both fasting and postprandial glucose levels across gestation (5–8). Although shared pathophysiologic features of GDM and T2D have been emphasized, it remains unclear whether all physiologic pathways that lead to hyperglycemia outside of pregnancy are equally important during gestation (9,10).

More than 200 genetic loci are known to be associated with T2D and are believed to act through distinct physiologic pathways leading to hyperglycemia (11,12). In previous work, Udler et al. (12) used clustering techniques to group known T2D-associated genetic variants according to physiologic effects in nonpregnant individuals. These genetically anchored clusters (Udler clusters, Table 1) highlight physiologic pathways leading to hyperglycemia in T2D. Polygenic scores that are based on these clusters identify individuals with T2D whose clinical features suggest that one of these pathways plays a significant role in their disease (12).

¹Diabetes Unit, Endocrine Division, Massachusetts General Hospital, Boston, MA

²Broad Institute, Cambridge, MA

³Harvard Medical School, Boston, MA

⁴Centre de Recherche du Centre Hospitalier, Université de Sherbrooke, Sherbrooke, Quebec, Canada

⁵Division of Biostatistics, Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL

⁶Clinical and Translational Epidemiology Unit, Mongan Institute, Massachusetts General Hospital, Boston, MA

⁷Department of Medicine, Université de Sherbrooke, Quebec, Canada

⁸Department of Biochemistry and Functional Genomics, Université de Sherbrooke, Sherbrooke, Quebec, Canada

⁹Department of Medical Biology, CIUSSS Saguenay-Lac-Saint-Jean–Hôpital Universitaire de Chicoutimi, Saguenay, Quebec, Canada

¹⁰Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL

¹¹Department of Population Medicine, Harvard Pilgrim Health Care Institute, Boston, MA

Corresponding author: Camille E. Powe, camille.powe@mgh.harvard.edu

Received 26 July 2020 and accepted 8 October 2020

This article contains supplementary material online at <https://doi.org/10.2337/figshare.13073054>.

© 2020 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <https://www.diabetesjournals.org/content/license>.

To illuminate physiologic pathways leading to GDM, the objective of the current study was to identify physiologically related clusters of maternal genetic variants associated with hyperglycemia in pregnancy using two complementary strategies. First, we tested groups of T2D-associated variants from clusters previously derived on the basis of physiology outside of pregnancy (Udler clusters) for association with GDM and offspring birth weight (reflecting in utero impact of maternal hyperglycemia) (13). Second, we derived clusters of maternal variants on the basis of pregnancy physiology (as assessed by 35 metabolic traits) and tested these novel clusters (pregnancy clusters) for association with the same outcomes.

RESEARCH DESIGN AND METHODS

Cohorts, Traits, and Genotyping

Genetics of Glucose Regulation in Gestation and Growth (Gen3G) is a cohort of pregnant women from Sherbrooke, Quebec, Canada ($N = 1,034$, enrolled 2010–2013) (14). Centre Hospitalier Universitaire de Sherbrooke's institutional review board (IRB) approved the study; participants provided written informed consent. Participants with genetic data available ($n = 574$) did not differ in age, BMI, gravidity, ethnicity, or smoking status from those who were not genotyped. Women enrolled in the first trimester underwent measurement of height, weight, waist circumference, and A1C and had a nonfasting 1-h 50-g glucose challenge test (GCT). Women with first trimester A1C $\geq 6.5\%$ or a GCT result ≥ 186 mg/dL were excluded. At 24–30 weeks' gestation, participants had a fasting 75-g oral glucose tolerance test (OGTT) during which glucose, insulin, and C-peptide levels were measured. Weight, fasting lipids and adipokines, and A1C were also measured. Those with GDM on the basis of the International Association of Diabetes and Pregnancy Study Groups criteria were referred for treatment (15). At delivery, birth weight was recorded. Genotyping, imputation, and quality control (QC) were performed as previously described (4). Principal component (PC) analysis was performed using PLINK.

The Hyperglycemia and Adverse Pregnancy Outcomes (HAPO) study is a multicenter, international, observational study of glycemia in pregnant women ($N = 25,505$, enrolled 2000–2006) (13). The objective of HAPO was to relate OGTT-measured glycemia to adverse pregnancy outcomes (13). Participants provided written informed consent; the IRB at each site approved the study. Participants underwent a blinded 75-g OGTT at 24–32 weeks' gestation. Those with fasting glucose ≥ 105 mg/dL or 2-h postload glucose ≥ 200 mg/dL were excluded; the remaining participants were observed without treatment. At the time of the OGTT, C-peptide, lipids, height, and weight were measured. Birth weight was recorded at delivery. We conducted our analysis in women with genome-wide single nucleotide polymorphism (SNP) data available ($n = 4,431$) (1). Participants selected for genotyping were from a subset of HAPO clinical sites. All participants genotyped from a given site shared the

same self-reported race/ethnicity as follows: Afro-Caribbean participants from Barbados (HAPO-AC); non-Hispanic White participants from the U.K., Canada, and Australia (HAPO-EU); Mexican American participants from California (HAPO-MA); and Thai participants from Bangkok (HAPO-TH). Hereafter, these sampling strata are referred to as subcohorts. Genotyping, imputation, QC, and PC analysis were performed separately within each subcohort, as previously described (1); likewise, we performed our analyses in each subcohort separately and meta-analyzed them. In the HAPO-AC subcohort, we selected proxies ($R^2 > 0.80$) to replace two SNPs for which data were unavailable after QC.

The Massachusetts General Hospital (MGH) Maternal Genetics and Health Study (MGH²) links clinical data from women who received prenatal care at MGH (1998–2015) and genomic data from the Partners Biobank (described below). Clinical data were imported from the electronic medical record into the study database. GDM cases were defined using Carpenter-Coustan criteria applied to a 100-g OGTT after 24 weeks' gestation; controls were women who had a screening GCT result < 140 mg/dL (16). Genotyping and imputation were conducted by the Partners Biobank.

The Partners Biobank contains genomic data and electronic health record–based clinical information from patients affiliated with Partners Healthcare (Boston, MA) (17,18). Participants consented to have their data used in broad-based research. The Partners Biobank and MGH² were approved by the Partners IRB. Genotyping was conducted on Illumina multiethnic arrays (17). Imputation was performed using a Minimac3 HRC (version r1.1 2016) reference panel and SHAPEIT. Variants used in our analysis had an imputation quality R^2 metric ≥ 0.80 , except rs13085136 at *SHQ1* ($R^2 = 0.77$) and rs2066827 at *CDKN1B* ($R^2 = 0.66$). Excluding or replacing these variants with proxies did not change results.

A T2D case-control cohort was selected from the Partners Biobank to test clusters for associations with hyperglycemia outside of pregnancy. T2D status, as previously described, was derived from structured and unstructured electronic medical record data using clinical, computational, and statistical methods (12,19).

Variant-Trait Clusters Derived Outside of Pregnancy (Udler Clusters)

Methods used by Udler et al. (12) to cluster traits and variants with Bayesian nonnegative matrix factorization (bNMF) were previously described in detail. In brief, summary statistics from genome-wide association studies (GWAS) in nonpregnant individuals, describing associations between traits and 94 T2D-associated variants, were assembled in a trait-variant association matrix. The bNMF procedure decomposes this matrix into two component matrices containing the variant and trait weights for each cluster, respectively. A Bayesian framework is used to determine the number of clusters (k) that best fit the data. Each cluster can be described by variants and traits that are most highly weighted (12) (Table 1).

Table 1—Description of Udler clusters

Udler cluster	Key genetic loci	Key phenotypic traits	Proposed mechanism
β-cell	<i>MTNR1B</i> , <i>HHEX</i> , <i>TNF7L2</i> , <i>SLC30A8</i> , <i>HNF1A</i> , <i>HNF1B</i>	↓Fasting insulin, insulin secretory response, BMI ↑Proinsulin	Insulin deficiency: β-cell dysfunction, downstream of proinsulin processing
Proinsulin	<i>ARAP1</i> , <i>SPRY2</i>	↓Fasting insulin, insulin secretory response, BMI, proinsulin	Insulin deficiency: β-cell dysfunction, upstream of proinsulin processing
Obesity	<i>FTO</i> , <i>MC4R</i>	↑Fasting insulin, waist circumference, BMI ↓Insulin sensitivity	Insulin resistance: obesity mediated
Lipodystrophy	<i>PPARG</i> , <i>ANKRD55</i> , <i>ARL15</i> , <i>GRB14</i> , <i>IRS1</i> , <i>LYPLAL1</i>	↑Fasting insulin, triglycerides ↓BMI, insulin sensitivity, HDL	Insulin resistance: fat distribution mediated
Liver-lipid	<i>GCKR</i> , <i>CILP2/TM6SF2</i> , <i>PNPLA3</i>	↑Fasting insulin ↓Triglycerides, insulin sensitivity	Insulin resistance: abnormal hepatic lipid metabolism*

From Udler et al. (12). *Variants in this cluster are associated with nonalcoholic fatty liver disease; functional studies of these variants suggested that they lead to sequestration of the lipids in the liver, lowering the levels in the blood.

Variant-Trait Clusters Derived in Pregnancy

SNP Selection

We selected 222 T2D-associated SNPs (primary signal at each locus, minor allele frequency $\geq 5\%$) from a large GWAS meta-analysis conducted in nonpregnant individuals (11) (Supplementary Table 1). We added four SNPs (from independent loci near *HKDC1*, *G6PC2*, *PCSK1*, and *PPP1R3B*) specifically associated with glycemia in pregnancy (1). We tested the glucose-raising allele at each locus for association with 35 traits measured in pregnant women from Gen3G using linear regression (Supplementary Table 2). Most traits were measured at 24–30 weeks' gestation, although we included BMI, body fat percentage, GCT result, A1C, and waist circumference measured in the first trimester as well as gestational weight gain (first trimester through delivery).

bNMF Clustering

One hundred sixty-four SNPs whose glucose-raising alleles had nominal associations ($P < 0.05$) with at least one trait in Gen3G were used in bNMF clustering (Supplementary Table 2). A matrix of standardized coefficients from linear regression relating SNPs to traits was inputted into the bNMF clustering procedure, which was run with 100 iterations using previously described methods (12).

Cluster Polygenic Scores

The top weighted variants in each cluster were used to build cluster polygenic scores (20). Polygenic scores consisted of the sum of the number of glucose-raising alleles carried by an individual, each multiplied by their weight in the cluster (1,11). For Udler clusters, the threshold for inclusion of variants in polygenic scores was previously described (12). For pregnancy clusters, we included variants that had cluster weights in the top 5% of all variant cluster weights. We standardized polygenic scores within each cohort/subcohort to aid in cross-cohort interpretation.

Statistical Analysis

We tested for relationships between cluster polygenic scores and physiologic traits (35 in Gen3G, 14 in HAPO, 3 in MGH²) in individual cohorts using linear regression. Adjusted models included PCs (first two PCs in each HAPO subcohort, first four PCs in Gen3G, first six PCs in MGH²) as covariates. In MGH², adjusted models also included genotyping/imputation batch. HAPO results were synthesized across subcohorts using fixed-effects inverse variance-weighted meta-analysis. Associations between polygenic scores and traits were considered suggestive if $P < 0.05$, given the descriptive nature of this part of the analysis.

The primary outcome was GDM, defined in Gen3G and HAPO using the International Association of Diabetes and Pregnancy Study Groups criteria and in MGH² using Carpenter-Coustan criteria (15,16). Secondary outcomes included birth weight percentile for gestational age and large-for-gestational-age (LGA) birth weight (defined as ≥ 90 th percentile) (21). We tested for associations between cluster polygenic scores and outcomes using logistic (for GDM and LGA) or linear (for birth weight percentile) regression in each cohort/subcohort. Adjusted models included maternal age, PCs (as above), and genotyping/imputation batch (in MGH²) as covariates.

We performed fixed-effects inverse variance-weighted meta-analysis to synthesize results across cohorts (including HAPO subcohorts). We excluded Gen3G from meta-analyses evaluating the association between pregnancy cluster polygenic scores and GDM because Gen3G glucose levels were used to generate clusters. Because of known limitations in translating polygenic scores across populations with different ancestries (22,23), we also conducted meta-analyses for each cluster and outcome using the cohorts with predominantly non-Hispanic White participants (HAPO-EU, MGH², plus Gen3G in the Udler cluster analyses), whose genetic ancestry was presumed to be most similar to those in whom the clusters were derived

(i.e., European predominant) (24). In meta-analyses, we examined heterogeneity using I^2 and Q metrics (25,26). Associations between polygenic scores were tested at $\alpha = 0.01$ (Bonferroni corrected) to account for multiple testing of the five polygenic scores in each of our two independent analyses (Udler clusters and pregnancy clusters). A power calculation suggested that with the available sample sizes, we had 80% power to detect effect sizes for associations between polygenic scores and GDM that produced odds ratios (ORs) <0.87 or >1.14 in three-cohort meta-analyses ($\alpha = 0.01$) (27).

Effect sizes (β or OR) for models relating polygenic scores to traits and outcomes are given for a 1-SD increase in polygenic score, with SD defined within each cohort. Statistical analyses were conducted in R 3.5.3 and 3.6.1 statistical software.

Data and Resource Availability

The Gen3G and MGH² data sets analyzed in this study are not publicly available because of IRB restrictions. These data sets are available from the corresponding author upon reasonable request and both institutional and IRB approval. The HAPO genotype data and accompanying phenotype data are currently available through the database of Genotypes and Phenotypes (<https://www.ncbi.nlm.nih.gov/gap>).

RESULTS

Participant Characteristics

Characteristics of pregnant participants are given in Table 2. In addition to having higher glucose levels, women with GDM were older and had higher BMI. In Gen3G and HAPO (where these measurements were available), women with GDM had lower insulin sensitivity and higher fasting triglycerides.

Udler Cluster Associations With Glycemic Traits in Pregnancy

Associations between cluster polygenic scores for T2D trait-variant clusters derived outside of pregnancy (Udler clusters) and selected traits in each pregnancy cohort are given in Table 3. Trait associations were considered suggestive if $P < 0.05$.

The β -cell cluster polygenic score was associated ($P < 0.05$) with higher postload glucose and lower insulin secretory response in Gen3G. In HAPO, there were no trait associations with the β -cell cluster polygenic score, although there were trends toward association with higher fasting and 2-h postload glucose. The fasting glucose trend was driven by HAPO-EU ($\beta = 0.45$ mg/dL, $P = 0.01$). In MGH² the β -cell cluster polygenic score was associated with lower first trimester BMI.

The proinsulin cluster was associated ($P < 0.05$) with a higher GCT result in MGH². It was not associated with traits in Gen3G or HAPO.

The obesity cluster polygenic score was associated ($P < 0.05$) with higher BMI in each of the cohorts. In Gen3G,

the obesity cluster was also associated with first trimester waist circumference ($\beta = 0.62$ cm, $P = 0.01$), body fat percentage (first trimester: $\beta = 0.78\%$, $P = 0.03$; 24–28 weeks' gestation: $\beta = 0.57\%$, $P = 0.04$), A1C at 24–30 weeks' gestation ($\beta = 0.03\%$, $P = 0.02$), and lower gestational weight gain. In HAPO, in addition to BMI, the obesity cluster polygenic score was associated with higher fasting ($\beta = 0.04$, $P = 0.008$) and 1-h postload C-peptide z-score and lower insulin sensitivity.

In Gen3G, the lipodystrophy-like cluster was associated ($P < 0.05$) with higher postload glucose (including first trimester GCT: $\beta = 2.3$ mg/dL, $P = 0.047$), lower insulin sensitivity, higher postload insulin (1-h postload: $\beta = 95.9$ pg/mL, $P = 0.04$; 2-h postload: $\beta = 120.3$ pg/mL, $P = 0.02$), higher triglycerides, lower HDL ($\beta = -1.75$ mg/dL, $P = 0.01$), and higher first trimester (but not 24–30-week) A1C ($\beta = 0.02\%$, $P = 0.02$). In HAPO, there were associations with higher 1-h postload C-peptide z-score, lower 1-h postload nonesterified fatty acid (NEFA) z-score ($\beta = -0.08$, $P = 0.002$), and a trend toward association with lower insulin sensitivity.

The liver-lipid cluster was associated ($P < 0.05$) with lower triglycerides and fasting NEFA z-score ($\beta = -0.10$, $P = 0.01$) in Gen3G. In HAPO, this cluster was associated with higher fasting glucose, higher fasting ($\beta = 0.06$, $P = 3.71 \times 10^{-5}$) and 1-h postload C-peptide z-score, lower insulin sensitivity, and lower fasting NEFA z-score ($\beta = -0.06$, $P = 0.04$). Associations between Udler clusters and traits in HAPO subcohorts are given in Supplementary Table 3A.

Udler Cluster Associations With Outcomes

The β -cell cluster was not significantly associated with GDM in the all-cohort meta-analysis (adjusted OR 1.08, $P = 0.06$) (Fig. 1A). In the European-predominant meta-analysis, there was a significant association ($P < 0.01$) in the unadjusted model (OR 1.19, $P = 0.009$) and a nominal association ($P < 0.05$) in the adjusted model (OR 1.18, $P = 0.015$) (Fig. 1B). The association was driven by a strong effect in Gen3G (adjusted OR 1.55, $P = 0.006$), with a consistent direction of effect in HAPO-EU (adjusted OR 1.15, $P = 0.10$).

The liver-lipid cluster was significantly associated ($P < 0.01$) with GDM in the all-cohort meta-analysis (adjusted OR 1.15, $P = 0.002$) (Fig. 1A), driven by associations in HAPO-EU (OR 1.29, $P = 0.004$) and HAPO-TH (OR 1.24, $P = 0.005$), and a consistent direction of effect in MGH² (adjusted OR 1.22, $P = 0.23$). There was evidence of heterogeneity in this meta-analysis ($I^2 = 0.59$, $Q = 12.1$, $P_Q = 0.03$), likely because of the opposite direction of effect in Gen3G (adjusted OR 0.76, $P = 0.09$). We found a nominal association ($P < 0.05$) between the liver-lipid cluster and GDM in the European-predominant meta-analysis (adjusted OR 1.15, $P = 0.04$) (Fig. 1B).

The proinsulin, obesity, and lipodystrophy-like clusters were not significantly associated with GDM (Fig. 1). None of the Udler clusters were associated with birth weight (Supplementary Fig. 1) or LGA.

Table 2—Characteristics of study participants

	Gen3G		HAPO		MGH ²	
	No GDM (n = 530)	GDM (n = 44)	No GDM (n = 3,718)	GDM (n = 713)	No GDM (n = 568)	GDM (n = 53)
Age (years)	28.1 (4.1)	29.7 (5.9)	28.0 (5.8)	30.7 (5.7)	30.2 (6.4)	33.9 (5.0)
Race/ethnicity, n (%)						
Non-Hispanic White	510 (96.2)	44 (100)	1,169 (31.4)	200 (28.1)	354 (62.3)	32 (60.4)
Non-Hispanic Black	3 (0.6)	0 (0.0)	1,011 (27.2)	92 (12.9)	27 (4.8)	1 (1.9)
Non-Hispanic Asian	1 (0.2)	0 (0.0)	924 (24.9)	231 (32.4)	19 (3.3)	2 (3.8)
Hispanic/Latina	7 (1.3)	0 (0.0)	614 (16.5)	190 (26.6)	139 (24.5)	16 (30.2)
Other/unknown	9 (1.7)	0 (0.0)	0 (0.0)	0 (0.0)	29 (5.1)	2 (3.8)
BMI (kg/m ²)*	25.4 (5.5)	28.6 (7.3)	27.3 (4.9)	30.2 (6.3)	25.6 (5.0)	30.1 (6.4)
Fasting glucose (mg/dL)	75.7 (5.4)	84.7 (11.9)	79.6 (5.4)	89.2 (7.3)		
1-h postload glucose (mg/dL)†	124.1 (25.2)	176.0 (25.4)	126.6 (25.0)	171.5 (30.6)	108.2 (19.7)	164.9 (17.0)
2-h postload glucose (mg/dL)	101.1 (19.8)	145.2 (26.3)	106.8 (18.7)	135.4 (26.0)		
Fasting C-peptide (μg/L)	0.95 (0.02)	1.21 (0.09)	1.8 (0.8)	2.4 (1.1)		
Insulin secretory response**	1,187.2 (452.5)	1,083.6 (454.4)	9.3 (3.0)	11.0 (3.4)		
ISI+	9.0 (5.3)	5.1 (2.74)	4.0 (1.5)	2.6 (1.0)		
Gestational weight gain (lb)	27.1 (9.9)	22.7 (12.1)			27.3 (11.9)	24.0 (8.7)
Triglycerides (mg/dL)	168.3 (53.1)	194.9 (75.3)	195.3 (69.6)	233.1 (93.3)		

Data are mean (SD) unless otherwise indicated. ISI, insulin sensitivity index. *BMI is reported from first trimester study visit for the Gen3G cohort, 24–32 weeks' gestation at OGTT for the HAPO cohorts, and first prenatal visit for the MGH² cohort. †One-hour postload glucose from the fasting 75-g OGTT in the Gen3G and HAPO cohorts; 50-g GCT result for the MGH² cohort. **Insulin secretory response is quantified by the Stumvoll first phase estimate from Gen3G cohort and 1-h C-peptide from HAPO cohorts (43,44). In Gen3G, mean (SD) 1-h C-peptide levels were similar in women without and with GDM (2.9 [1.5] and 3.2 [1.2] μg/L, respectively). +ISI is defined by the Matsuda index in the Gen3G cohort and by a modified Matsuda index using C-peptide in the HAPO cohorts (45,46).

Table 3—Associations between Udlr cluster polygenic scores and glycemic traits in pregnancy

	Fasting glucose (mg/dL)	1-h glucose (mg/dL)*	2-h glucose (mg/dL)	BMI (kg/m ²) ⁺	Insulin secretory response**	ISI†	Gestational weight gain (lb)	Triglycerides (mg/dL)
β-Cell								
Gen3G								
β	0.46	4.57	2.33	0.31	-46.5	-0.073	-0.70	-0.21
P	0.11	1.33 × 10⁻⁴	0.02	0.20	0.02	0.75	0.10	0.93
HAPO								
β	0.15	0.017	0.51	0.04	-0.014	0.014		3.25
P	0.09	0.96	0.05	0.58	0.31	0.52		0.18
MGH ²								
β		0.70		-0.44			0.56	
P		0.50		0.048			0.24	
Proinsulin								
Gen3G								
β	0.03	0.60	0.75	0.15	-4.06	-0.05	-0.12	3.23
P	0.93	0.62	0.45	0.55	0.83	0.82	0.77	0.16
HAPO								
β	-0.15	-0.26	-0.10	0.10	0.007	-0.003		4.14
P	0.09	0.43	0.72	0.17	0.61	0.90		0.09
MGH ²								
β		2.98		0.03			-0.49	
P		0.005		0.89			0.31	
Obesity								
Gen3G								
β	0.54	1.88	1.82	0.62	-9.15	-0.31	-1.09	-1.26
P	0.06	0.12	0.07	0.01	0.64	0.17	0.01	0.59
HAPO								
β	0.10	0.14	0.036	0.35	0.031	-0.07		-0.49
P	0.27	0.68	0.89	7.00 × 10⁻⁷	0.03	0.001		0.84
MGH ²								
β		0.104		0.58			0.28	
P		0.92		0.008			0.55	
Lipodystrophy								
Gen3G								
β	0.065	2.41	1.86	0.036	27.33	-0.63	0.44	5.52
P	0.82	0.04	0.06	0.88	0.15	0.005	0.30	0.02
HAPO								
β	0.022	0.10	-0.30	0.026	0.03	-0.04		-1.21
P	0.81	0.76	0.26	0.71	0.02	0.06		0.63
MGH ²								
β		0.98		-0.20			0.48	
P		0.35		0.37			0.31	
Liver-lipid								
Gen3G								
β	0.40	-0.81	-1.83	0.43	33.52	0.04	0.44	-11.02
P	0.17	0.50	0.07	0.08	0.08	0.87	0.30	1.91 × 10⁻⁶
HAPO								
β	0.30	0.35	-0.11	-0.08	0.05	-0.10		-4.51
P	0.001	0.30	0.67	0.26	3.49 × 10⁻⁴	1.82 × 10⁻⁶		0.06
MGH ²								
β		0.24		0.013			-0.21	
P		0.82		0.95			0.67	

Associations between clusters and traits in Gen3G ($n = 574$), HAPO ($n = 4,431$), and MGH² ($n = 621$) are adjusted for PCs (and genotyping/imputation batch in MGH² only). Associations with $P < 0.05$ were considered suggestive and are highlighted in bold. ISI, insulin sensitivity index. *One-hour postload glucose from the fasting 75-g OGTT in the Gen3G and HAPO cohorts; 50-g GCT result for the MGH² cohort. +BMI from first trimester study visit for the Gen3G cohort, 24–32 weeks' gestation at OGTT for the HAPO cohorts, and the first prenatal visit for the MGH² cohort. **Insulin secretory response is quantified by the Stumvoll first phase estimate from Gen3G cohort and 1-h C-peptide z-score from HAPO cohorts (43,44). †ISI is defined by the Matsuda index in the Gen3G cohort and by a modified Matsuda index using C-peptide concentrations in the HAPO cohorts (45,46).

Pregnancy Cluster Associations With Glycemic Traits

In 100 iterations of bNMF using associations between traits and variants in Gen3G, five pregnancy clusters emerged in the plurality of iterations (49 of 100) (Supplementary Fig. 2). In 48 of the 51 remaining iterations, the number of clusters was four (24 iterations) or six (24 iterations), representing collapsing or splitting of two clusters. Figure 2 depicts the highest weighted variants and traits in each cluster. Table 4 provides associations between pregnancy cluster polygenic scores and selected traits in each pregnancy cohort. Trait associations were considered suggestive if $P < 0.05$.

In pregnancy cluster 1 (Fig. 2A), the highest weighted glycemic traits included higher postload glucose levels, lower disposition index, and higher adiposity measures. Higher fasting glucose also appeared in this cluster but was less strongly weighted than postload glucose. The highest weighted genetic loci included several known or suspected to be associated with diminished β -cell function (*MTNR1B*, *GLP2R*, *CRHR2*) and obesity (*MC4R*, *FTO*), along with others with unknown effects (*PURG*, *MRPS30*, *SHQ1*) (11,12,28–31). Also highly weighted in this cluster was the *SLC2A2* locus, known to be involved in glucose transport into hepatocytes and islets and associated with metformin response, and the *PI3KR1* locus, which is less well characterized but appears to be associated with insulin resistance-related traits outside of pregnancy (32–34). In HAPO, the pregnancy cluster 1 polygenic score was associated ($P < 0.05$) with higher fasting glucose, higher BMI, and lower insulin sensitivity.

Highly weighted traits in pregnancy cluster 2 (Fig. 2B) included those representing increased fasting insulin and greater insulin secretory response (unadjusted and adjusted for insulin sensitivity). In addition, higher levels of tumor necrosis factor- α (TNF- α), which has been linked to insulin resistance in pregnancy, was highly weighted (35,36). Among the highest weighted loci in pregnancy cluster 2 were *PIM3*, *BNIPL*, *GLI2*, and *PHF15*, each of which has evidence for acting through insulin resistance mechanisms outside of pregnancy (11,28,32). In HAPO, this cluster was associated ($P < 0.05$) with lower A1C ($\beta = -0.01$, $P = 0.03$) and higher HDL ($\beta = 1.28$ mg/dL, $P = 0.01$).

Highly weighted traits in pregnancy cluster 3 (Fig. 2C) included those representing favorable glucose metabolism, including higher oral disposition index, greater insulin sensitivity, and lower postload glucose. The highest weighted loci in this cluster contained a mix of those with unknown function as well as some β -cell-, obesity-, and insulin resistance-associated loci (28,32). No traits were associated with this cluster in HAPO or MGH².

Highly weighted traits in pregnancy cluster 4 included those representing reduced adiposity. The three highly weighted loci in this cluster were *TCEA2*, *YWHAH*, and *PPP1R3B*; T2D-associated alleles at these loci have evidence for association with reduced BMI outside pregnancy (32). In HAPO, this cluster was associated ($P < 0.05$) with higher fasting glucose, higher fasting C-peptide z-score, lower postload glucose, lower insulin sensitivity, and lower

NEFA z-scores (fasting: $\beta = -0.08$, $P = 0.004$; 1-h postload: $\beta = -0.08$, $P = 0.002$).

Highly weighted traits in pregnancy cluster 5 included those representing lower fasting insulin and insulin secretory response. Highly weighted loci in this cluster included many known or suspected to be associated with reduced β -cell function, including *ABO*, *CDKN2AB*, *SLC30A8*, *CDKN1B*, and *ST6GAL1* (11,12). Although the pregnancy cluster 5 polygenic score was associated ($P < 0.05$) with lower fasting insulin and C-peptide and lower insulin secretory response in Gen3G, there was no association with glucose levels (Table 4). In HAPO, this cluster was associated with higher postload glucose and a trend toward lower 1-h postload C-peptide z-score as well as higher total cholesterol ($\beta = 4.32$, $P = 0.01$) and LDL ($\beta = 4.14$, $P = 0.008$). Associations between pregnancy clusters and traits in HAPO subcohorts are given in Supplementary Table 3B.

Pregnancy Cluster Associations With Outcomes

Pregnancy cluster 1 was associated with GDM in Gen3G (adjusted OR 1.61, $P = 0.003$). In a meta-analysis including the remaining cohorts, a significant association ($P < 0.01$) between pregnancy cluster 1 and GDM was replicated (adjusted OR 1.24, $P = 6.2 \times 10^{-7}$) (Fig. 3A). Exclusion of the *MTNR1B* locus from the polygenic score did not affect the magnitude or statistical significance of the results (adjusted OR 1.20, $P = 0.004$). The significant association between pregnancy cluster 1 and GDM was also present in the European-predominant meta-analysis (adjusted OR 1.35, $P = 1.91 \times 10^{-5}$) (Fig. 3B). There were no statistically significant associations between GDM and the other pregnancy clusters (Fig. 3A).

In the all-cohort meta-analysis, there was a nominal association ($P < 0.05$) between the pregnancy cluster 1 polygenic score and higher offspring birth weight percentile (adjusted $\beta = 0.91$ percentile points, $P = 0.02$) (Supplementary Fig. 1C) that did not meet criteria for statistical significance. This association was statistically significant ($P < 0.01$) in the European-predominant meta-analysis (adjusted $\beta = 1.54$ percentile points, $P = 0.007$) (Supplementary Fig. 1D). There was no association between pregnancy cluster 1 and LGA in the all-cohort meta-analysis, although there was a nominal (but not statistically significant) association in the European-predominant meta-analysis (adjusted OR 1.18, $P = 0.01$). None of the other pregnancy clusters were nominally or significantly associated with birth weight (Supplementary Fig. 1) or LGA in meta-analyses.

Cluster Associations With T2D

In the T2D case-control set ($n = 4,910$ cases, $n = 28,206$ controls) (Supplementary Table 4), we tested for associations between each cluster and T2D (Fig. 4). Of the Udler clusters, the β -cell, obesity, and lipodystrophy-like clusters were significantly associated ($P < 0.01$) with T2D. The proinsulin and liver-lipid clusters had nominal associations ($P < 0.05$). For the lipodystrophy-like cluster, the effect

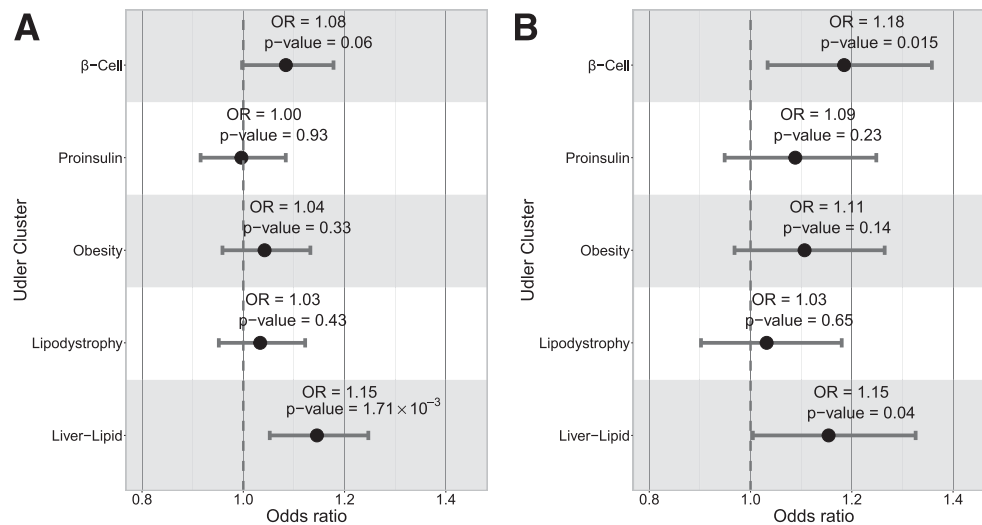


Figure 1—Associations between Udlar cluster polygenic scores and GDM. Shown are the results from meta-analyses of associations between Udlar cluster polygenic scores and GDM. *A*: Meta-analysis of all cohorts (Gen3G, HAPO-AC, HAPO-EU, HAPO-MA, HAPO-TH, and MGH²; $n = 810$ cases, $n = 4,816$ controls). *B*: Meta-analysis of cohorts with presumed European-predominant ancestry (Gen3G, HAPO-EU, and MGH²; $n = 297$ cases, $n = 2,267$ controls). Prior to meta-analysis, associations from logistic regression were adjusted for PCs and age. In the MGH² cohort, we also adjusted for genotyping/imputation batch. ORs, ●. Error bars show the 95% CIs of the ORs. $P < 0.01$ was considered statistically significant.

size for the association with T2D appeared larger than that for GDM (T2D OR 1.17 [95% CI 1.13–1.21]; GDM meta-analysis OR 1.03 [0.95–1.12]). Of the novel pregnancy clusters, pregnancy cluster 1, pregnancy cluster 3, and pregnancy cluster 5 were significantly associated with T2D ($P < 0.01$); pregnancy cluster 2 was nominally associated ($P < 0.05$), and pregnancy cluster 4 was not associated ($P \geq 0.05$). For pregnancy cluster 1, the effect size for association with GDM appeared greater than that for T2D (GDM meta-analysis OR 1.24 [1.14–1.35]; T2D OR 1.11 [1.07–1.14]); the effect size for GDM also appeared greater than that for T2D in the European-predominant GDM meta-analysis (OR 1.35 [1.18–1.55]) (Supplementary Fig. 3).

DISCUSSION

In this study of >5,000 pregnant women, we identify physiologically grouped sets of genetic variants associated with GDM. Observed associations with clusters derived outside of pregnancy imply that β -cell dysfunction and abnormal hepatic lipid metabolism are pathophysiologic mechanisms shared between GDM and T2D. We also identified a strong association between GDM and a novel cluster of variants derived using pregnancy physiology (pregnancy cluster 1). We replicated this association across independent cohorts; notably, the effect size for the association of this cluster with GDM appeared greater than that for T2D. Carrying a greater number of glucose-raising alleles from pregnancy cluster 1 was also nominally associated with having an infant with higher birth weight, presumably as a consequence of maternal hyperglycemia. To our knowledge, we are the first to apply bNMF clustering to the study of GDM and gestational glucose metabolism.

Previous studies have tested T2D-associated variants for association with GDM, but few have probed variants grouped on the basis of more specific physiology (2,4,37–39). In a previous study in Gen3G and HAPO-EU, we found that trait-based polygenic scores (with component variants selected on the basis of effects outside pregnancy) for fasting glucose, fasting insulin, insulin secretion, and insulin sensitivity were associated with GDM (4). Similarly, Moen et al. (37) found associations between polygenic scores for fasting glucose, BMI, and T2D and glucose levels in pregnancy. The current study builds upon this literature by testing variants grouped by physiologic pathway (rather than by effects on a single trait) for associations with pregnancy glycemic traits and GDM; this allows us to identify physiologic pathways that play a role in GDM pathogenesis. A role for β -cell dysfunction and abnormal hepatic lipid metabolism in GDM pathogenesis is implied by the associations we found with the β -cell and liver-lipid cluster polygenic scores. While we may have lacked power to detect associations of modest size between GDM and other Udlar clusters, abnormal fat distribution may be less important in GDM than in T2D, given that the effect size for the lipodystrophy-like cluster polygenic score with GDM was lesser than that for T2D.

Our findings highlight the existence of pathophysiology underlying GDM that appears unique to pregnancy. Specifically, we implicated a novel cluster of T2D-associated genetic variants (pregnancy cluster 1) in pregnancy glycaemia and fetal growth. The robustness of the association of this cluster with GDM is bolstered by replication in independent cohorts. This cluster appeared to have a stronger association with GDM than with T2D. The variants that were highly weighted in pregnancy cluster 1 did not

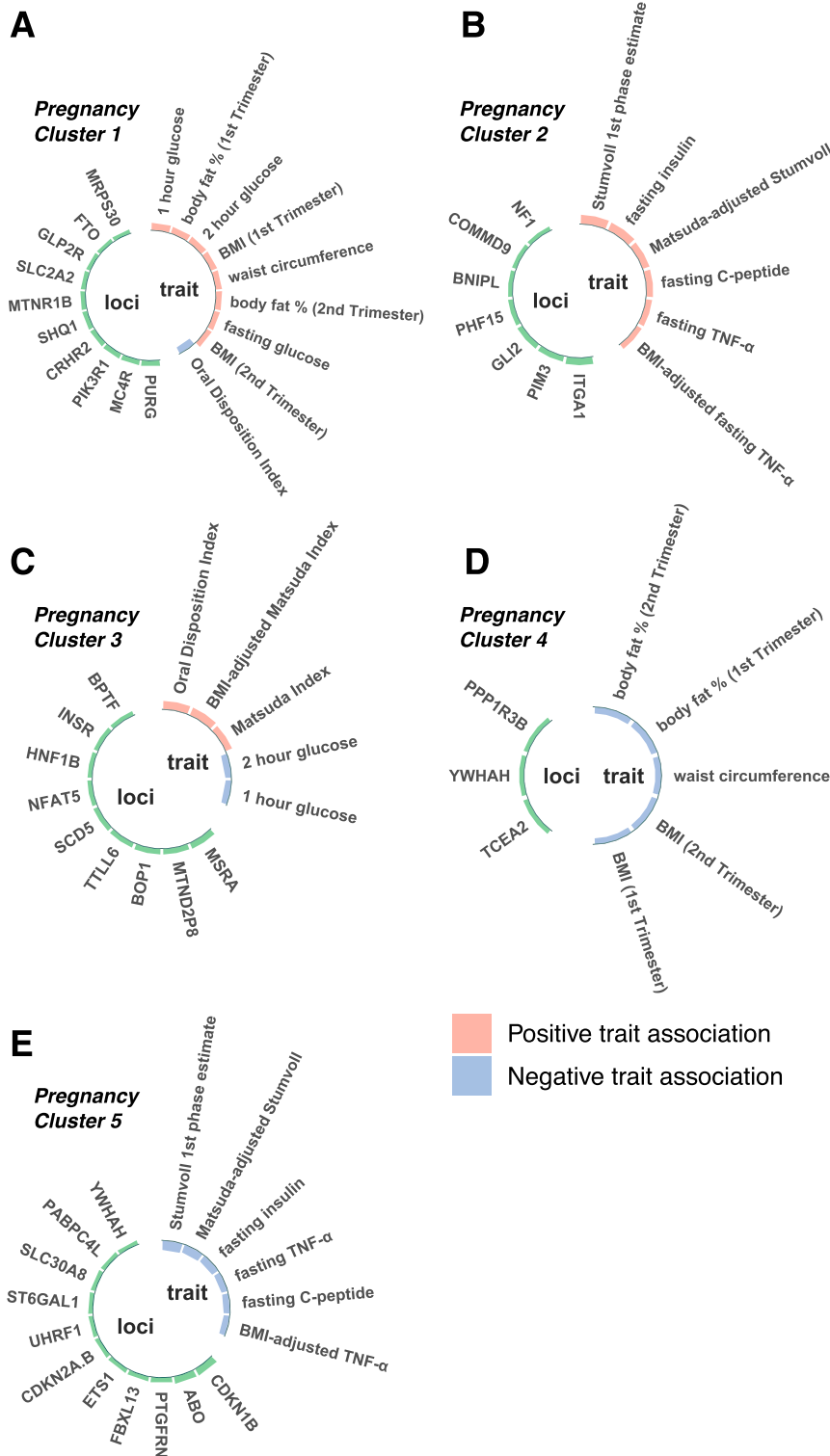


Figure 2—Highly weighted traits and variants in pregnancy clusters. A–E: Highly weighted traits and variants (lying in the top 5% of all cluster weights) are given for newly described pregnancy clusters. The height of the bar for each trait (pink/blue) or locus (green) indicates the strength of the weight in the relevant cluster. Traits with pink bars are positively associated with the cluster; traits with blue bars are negatively associated with the cluster.

appear to represent a single physiologic pathway, leading us to speculate that the commonality resulting in their grouping was their strong effects on pregnancy glycemia. In contrast, despite apparent associations with insulin

resistance and deficiency in pregnant women, polygenic scores for pregnancy clusters 2 and 5 did not seem to be specifically associated with GDM, although larger studies are required to provide more accurate effect estimates and

Table 4—Associations between pregnancy cluster polygenic scores and glycemic traits in pregnancy

	Fasting glucose (mg/dL)	1-h glucose (mg/dL)*	2-h glucose (mg/dL)	BMI (kg/m ²) ⁺	Insulin secretory response**	ISI†	Gestational weight gain (lb)	Triglycerides (mg/dL)
Cluster 1								
Gen3G								
β	1.57	6.07	4.01	1.37	25.5	-0.99	-0.96	4.88
P	3.24 × 10⁻⁸	3.04 × 10⁻⁷	4.39 × 10⁻⁵	1.15 × 10⁻⁸	0.18	9.20 × 10⁻⁶	0.02	0.03
HAPO								
β	0.44	0.58	0.23	0.23	0.012	-0.064		-2.21
P	9.06 × 10⁻⁷	0.08	0.38	8.39 × 10⁻⁴	0.42	0.003		0.38
MGH ²								
β		1.48		-0.084			0.53	
P		0.15		0.70			0.26	
Cluster 2								
Gen3G								
β	0.66	-1.28	-0.73	0.43	107.45	-0.64	0.91	-3.85
P	0.02	0.29	0.46	0.08	1.73 × 10⁻⁸	0.004	0.03	0.10
HAPO								
β	0.84	-0.02	-0.12	0.04	0.004	-0.003		-0.67
P	0.34	0.94	0.63	0.53	0.80	0.90		0.79
MGH ²								
β		1.01		0.31			-0.37	
P		0.32		0.16			0.43	
Cluster 3								
Gen3G								
β	-1.03	-5.37	-4.08	0.064	-11.15	1.31	0.0019	-7.51
P	3.22 × 10⁻⁴	6.08 × 10⁻⁶	3.13 × 10⁻⁵	0.79	0.56	3.15 × 10⁻⁹	>0.99	0.001
HAPO								
β	0.09	0.21	0.19	0.10	-0.02	0.002		-0.85
P	0.30	0.54	0.47	0.15	0.18	0.92		0.72
MGH ²								
β		0.05		0.35			0.58	
P		0.96		0.11			0.90	
Cluster 4								
Gen3G								
β	-0.05	-0.18	-1.03	-0.68	-4.34	0.30	-0.13	2.64
P	0.87	0.88	0.31	0.005	0.82	0.19	0.76	0.26
HAPO								
β	0.31	-0.74	-0.63	-0.009	-0.01	-0.04		0.64
P	4.12 × 10⁻⁴	0.03	0.02	0.89	0.32	0.04		0.79
MGH ²								
β		0.27		0.24			0.06	
P		0.79		0.26			0.90	
Cluster 5								
Gen3G								
β	-0.39	1.45	0.16	-0.57	-129.14	0.74	-0.44	-5.83
P	0.18	0.23	0.88	0.02	1.10 × 10⁻¹¹	0.001	0.31	0.01
HAPO								
β	0.011	0.73	0.57	-0.03	-0.03	0.02		-2.11
P	0.91	0.03	0.03	0.63	0.06	0.28		0.40
MGH ²								
β		0.14		-0.41			-0.53	
P		0.89		0.06			0.27	

Associations between clusters and traits in Gen3G ($n = 574$), HAPO ($n = 4,431$), and MGH² ($n = 621$) are adjusted for PCs (and genotyping/imputation batch in MGH² only). Associations with $P < 0.05$ were considered suggestive and are highlighted in bold. *One-hour postload glucose from the fasting 75-g OGTT in the Gen3G and HAPO cohorts; 50-g glucose loading test result for the MGH² cohort. ⁺BMI from first trimester study visit for the Gen3G cohort, 24–32 weeks' gestation at OGTT for the HAPO cohorts, and the first prenatal visit for the MGH² cohort. **Insulin secretory response is quantified by the Stumvoll first phase estimate from Gen3G cohort and 1-h C-peptide from HAPO cohorts (43,44). †ISI is defined by the Matsuda index in the Gen3G cohort and defined by a modified Matsuda index using C-peptide concentrations in the HAPO cohorts (45,46).

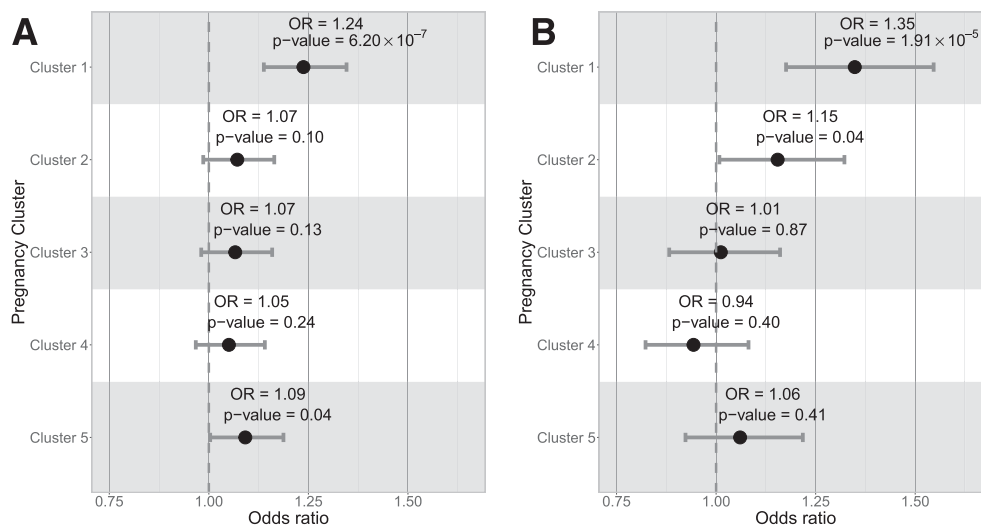


Figure 3—Associations between pregnancy cluster polygenic scores and GDM. Shown are the results from meta-analyses of associations between pregnancy cluster polygenic scores and GDM. Gen3G was excluded given that glucose values in this cohort were used to generate clusters. **A:** Meta-analyses of all remaining cohorts (HAPO-AC, HAPO-EU, HAPO-MA, HAPO-TH, and MGH²; $n = 766$ cases, $n = 4,286$ controls). **B:** Meta-analyses of remaining cohorts with presumed European-predominant ancestry (HAPO-EU and MGH²; $n = 253$ cases, $n = 1,737$ controls). Prior to meta-analysis, associations from logistic regression were adjusted for PCs and age. In the MGH² cohort, we also adjusted for genotyping/imputation batch. ORs, ●. Error bars show the 95% CIs for the ORs. $P < 0.01$ was considered statistically significant.

exclude associations that were too small for our study to detect.

In addition to testing for associations with GDM, we examined relationships between clusters and infant birth weight because fetal overgrowth is a common consequence of hyperglycemia in pregnancy (13). We found that despite associations with GDM, and in contrast to the novel GDM-associated pregnancy cluster 1, the Udler β -cell and liver-lipid clusters showed no evidence of association with infant birth weight, although we may not have had power to detect associations of limited magnitude. Relevant to our observations, there is a known relationship between reduced fetal insulin secretion and lower birth weight (40,41). This could explain why a group of variants associated with GDM through diminished maternal insulin secretion (β -cell cluster) did not increase birth weight to the degree expected from the maternal hyperglycemia alone; indeed, there may be a counterbalancing effect on birth weight in fetuses who inherit these insulin secretion-reducing variants from their mothers. With regard to the liver-lipid cluster, it is plausible that lower triglyceride levels associated with this cluster resulted in less fetal overgrowth, despite an association with maternal hyperglycemia (42). Our findings imply that the group of variants in pregnancy cluster 1, in aggregate, do not diminish fetal insulin secretion and/or reduce maternal circulating lipids in such a way as to limit fetal growth. We note that the association between birth weight and pregnancy cluster 1 did not reach the prespecified level of statistical significance in the all-cohort meta-analysis and thus requires confirmation in future studies.

Investigations of associations between common genetic variants and disease have been limited by the exclusion of

individuals without recent European ancestry (22,23). Our investigation did include participants with diverse geographic ancestries, yet as expected, variants and clusters discovered or derived in populations with recent European ancestry generally displayed less robust associations with glycemic traits and GDM in women without recent European ancestry. Our findings underscore the critical need for diabetes genetic investigations in ethnically diverse populations both in and outside of pregnancy.

A strength of our investigation is the study population, which contained independent cohorts in which to replicate our findings. Furthermore, our focus on pregnancy physiology addresses a gap in the literature. Limitations include our sample size, which while large for a genetic investigation in pregnant women, was likely too small to detect modest genetic effects. In addition, because of the limited number of GWAS with data on glycemia in pregnancy (1,2), the majority of variants included in this investigation were selected on the basis of their effects outside of pregnancy on T2D risk in populations with recent European ancestry. Until we have ethnically diverse genetic cohorts of sufficient size to identify loci across the genome associated with glycemia and related traits during gestation, our ability to study genetic determinants of glucose metabolism unique to pregnancy will be limited. In addition, the inclusion criteria and definitions of GDM differed among cohorts; this may have led to heterogeneity of effects across cohorts. For this reason, we have focused primarily on effects with evidence of consistency across cohorts on the basis of findings in meta-analyses.

In conclusion, we have identified physiologic pathways and groups of genetic variants associated with GDM,

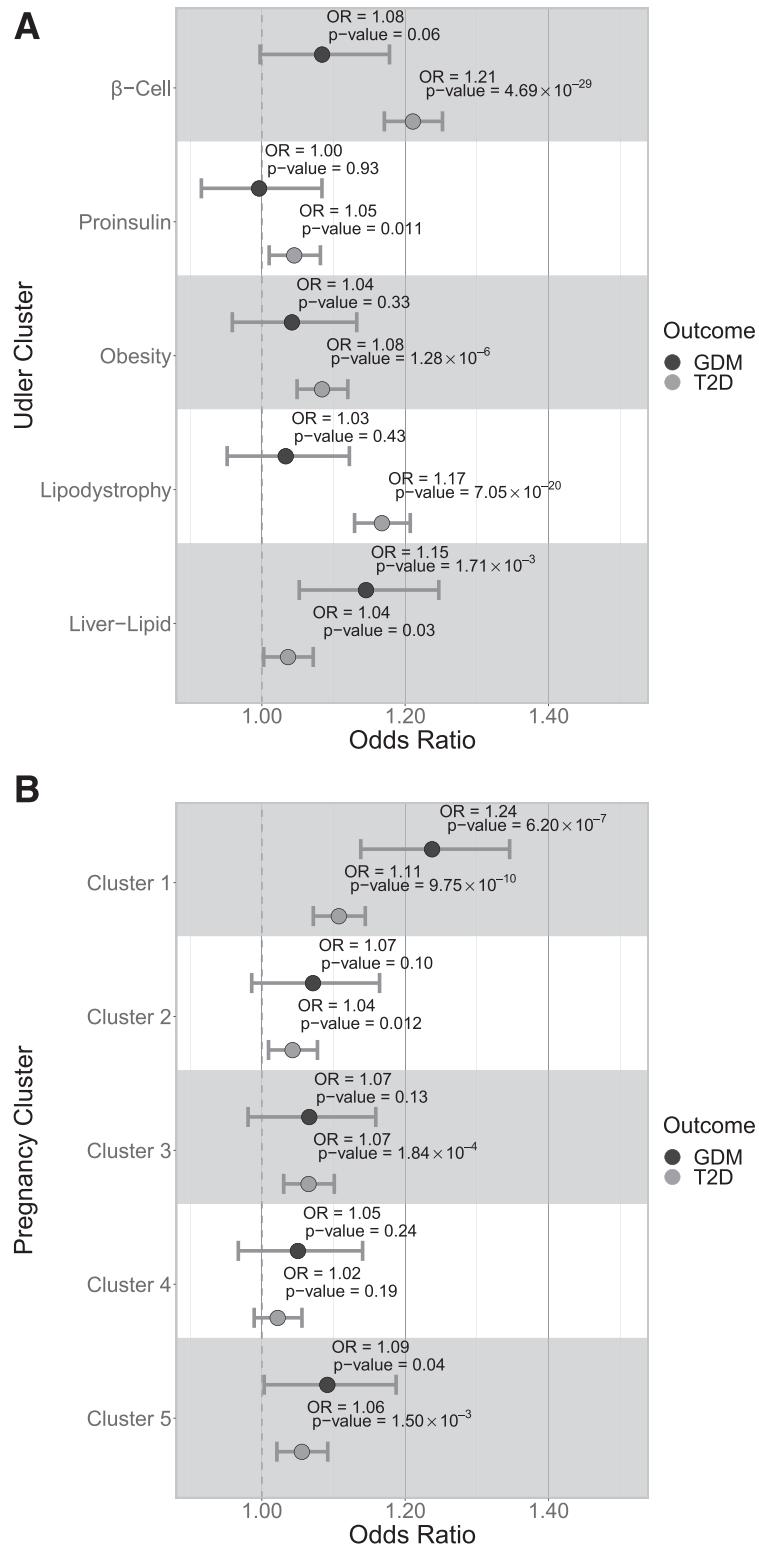


Figure 4—Comparison of cluster polygenetic score associations with GDM and T2D. We compared the association of each cluster—Udler clusters (A) and pregnancy clusters (B)—with GDM (from results of meta-analyses depicted in Fig. 1A [$n = 810$ cases, $n = 4,816$ controls] and Fig. 3A [$n = 766$ cases, $n = 4,286$ controls]) and T2D (from participants in the Partners Biobank [$n = 4,910$ cases, $n = 28,206$ controls]). Associations from logistic regression were adjusted for PCs and age. In the MGH² and Partners Biobank, we also adjusted for genotyping/imputation batch. ORs, ●. Error bars show the 95% CIs for the ORs. $P < 0.01$ was considered statistically significant.

including a newly described cluster of variants that is associated with hyperglycemia and fetal growth in pregnant women. While not the objective of the present investigation, future studies should test whether polygenic scores from our physiology-informed, genetically anchored clusters can be used clinically. For example, it is possible that the pregnancy cluster 1 polygenic score could identify women who would benefit from early GDM screening or that the Udler β -cell and liver-lipid cluster polygenic scores could identify women for whom GDM heralds future T2D. Studies in large, diverse cohorts with genetic and pregnancy phenotypic data will be required to further advance understanding of the genetic determinants of the unique pathophysiology underlying gestational glucose metabolism revealed by our analysis.

Acknowledgments. The authors acknowledge Tanayott Thaweethai (Biostatistics Center, Massachusetts General Hospital, Boston, MA) for assistance with power calculations.

Funding. C.E.P. is supported by National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) grant K23-DK-113218 and the Robert Wood Johnson Foundation Harold Amos Medical Faculty Development Program. M.S.U. is supported by NIDDK grant K23-DK-114551. L.B. is a senior research scholar from the Fond de la recherche du Québec en santé. M.-F.H. and Gen3G are supported by American Diabetes Association Pathways Award #1-15-ACE-26. HAPO was supported by grants from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (HD-34242 and HD-34243), National Human Genome Research Institute (HG-004415), and NIDDK (R01-DK-097534). MGH² is supported by the Massachusetts General Hospital Physician Scientist Development Award (to C.E.P.).

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

Author Contributions. C.E.P., J.C.F., and M.-F.H. conceptualized the project. C.E.P., P.P., L.B., W.L.L., D.S., and M.-F.H. oversaw data collection and/or abstraction. C.E.P., M.S.U., S.H., C.A., and A.K. planned and executed the analysis. C.E.P., M.S.U., S.H., A.K.M., W.L.L., D.S., J.C.F., and M.-F.H. interpreted the results. C.E.P. drafted the manuscript; all other authors critically reviewed and edited the manuscript. C.E.P. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Prior Presentation. Parts of this study were presented in abstract form at the 79th Scientific Sessions of the American Diabetes Association, San Francisco, CA, 7–11 June 2019, and 80th Scientific Sessions of the American Diabetes Association, 12–16 June 2020.

References

- Hayes MG, Urbanek M, Hivert MF, et al.; HAPO Study Cooperative Research Group. Identification of HKDC1 and BACE2 as genes influencing glycemic traits during pregnancy through genome-wide association studies. *Diabetes* 2013;62:3282–3291
- Kwak SH, Kim SH, Cho YM, et al. A genome-wide association study of gestational diabetes mellitus in Korean women. *Diabetes* 2012;61:531–541
- Lowe WL Jr., Scholtens DM, Sandler V, Hayes MG. Genetics of gestational diabetes mellitus and maternal metabolism. *Curr Diab Rep* 2016;16:15
- Powe CE, Nodzinski M, Talbot O, et al. Genetic determinants of glycemic traits and the risk of gestational diabetes mellitus. *Diabetes* 2018;67:2703–2709
- Catalano PM, Tyzbir ED, Roman NM, Amini SB, Sims EA. Longitudinal changes in insulin release and insulin resistance in nonobese pregnant women. *Am J Obstet Gynecol* 1991;165:1667–1672
- Catalano PM, Huston L, Amini SB, Kalhan SC. Longitudinal changes in glucose metabolism during pregnancy in obese women with normal glucose tolerance and gestational diabetes mellitus. *Am J Obstet Gynecol* 1999;180:903–916
- Catalano PM, Tyzbir ED, Wolfe RR, et al. Carbohydrate metabolism during pregnancy in control subjects and women with gestational diabetes. *Am J Physiol* 1993;264:E60–E67
- Powe CE, Huston Presley LP, Locascio JJ, Catalano PM. Augmented insulin secretory response in early pregnancy. *Diabetologia* 2019;62:1445–1452
- McIntyre HD, Catalano P, Zhang C, Desoye G, Mathiesen ER, Damm P. Gestational diabetes mellitus. *Nat Rev Dis Primers* 2019;5:47
- Buchanan TA, Xiang A, Kjos SL, Watanabe R. What is gestational diabetes? *Diabetes Care* 2007;30(Suppl. 2):S105–S111
- Mahajan A, Taliun D, Thurner M, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet* 2018;50:1505–1513
- Udler MS, Kim J, von Grotthuss M, et al.; Christopher D. Anderson on behalf of METASTROKE and the ISGC. Type 2 diabetes genetic loci informed by multi-trait associations point to disease mechanisms and subtypes: a soft clustering analysis. *PLoS Med* 2018;15:e1002654
- Metzger BE, Lowe LP, Dyer AR, et al.; HAPO Study Cooperative Research Group. Hyperglycemia and adverse pregnancy outcomes. *N Engl J Med* 2008;358:1991–2002
- Guillemette L, Allard C, Lacroix M, et al. Genetics of Glucose regulation in Gestation and Growth (Gen3G): a prospective prebirth cohort of mother-child pairs in Sherbrooke, Canada. *BMJ Open* 2016;6:e010031
- Metzger BE, Gabbe SG, Persson B, et al.; International Association of Diabetes and Pregnancy Study Groups Consensus Panel. International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. *Diabetes Care* 2010;33:676–682
- American Diabetes Association. 2. Classification and diagnosis of diabetes: *Standards of Medical Care in Diabetes—2020*. *Diabetes Care* 2020;43(Suppl. 1):S14–S31
- Smoller JW, Karlson EW, Green RC, et al. An eMERGE clinical center at Partners Personalized Medicine. *J Pers Med* 2016;6:5
- Karlson EW, Boutin NT, Hoffnagle AG, Allen NL. Building the Partners HealthCare Biobank at Partners Personalized Medicine: informed consent, return of research results, recruitment lessons and operational considerations. *J Pers Med* 2016;6:2
- Yu S, Liao KP, Shaw SY, et al. Toward high-throughput phenotyping: unbiased automated feature extraction and selection from knowledge sources. *J Am Med Inform Assoc* 2015;22:993–1000
- Udler MS, McCarthy MI, Florez JC, Mahajan A. Genetic risk scores for diabetes diagnosis and precision medicine. *Endocr Rev* 2019;40:1500–1520
- Oken E, Kleinman KP, Rich-Edwards J, Gillman MW. A nearly continuous measure of birth weight for gestational age using a United States national reference. *BMC Pediatr* 2003;3:6
- Martin AR, Gignoux CR, Walters RK, et al. Human demographic history impacts genetic risk prediction across diverse populations. *Am J Hum Genet* 2017;100:635–649
- Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet* 2019;51:584–591
- Bamshad M, Wooding S, Salisbury BA, Stephens JC. Deconstructing the relationship between genetics and race. *Nat Rev Genet* 2004;5:598–609
- Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002;21:1539–1558
- Cochran WG. The combination of estimates from different experiments. *Biometrics* 1954;10:101–129
- Valentine JC, Pigott TD, Rothstein HR. How many studies do you need? A primer on statistical power for meta-analysis. *J Educ Behav Stat* 2010;35:215–247
- Mahajan A, Wessel J, Willems SM, et al.; ExomeBP Consortium; MAGIC Consortium; GIANT Consortium. Refining the accuracy of validated target

- identification through coding variant fine-mapping in type 2 diabetes. *Nat Genet* 2018;50:559–571
29. Prokopenko I, Poon W, Mägi R, et al. A central role for GRB10 in regulation of islet function in man. *PLoS Genet* 2014;10:e1004235
30. Ingelsson E, Langenberg C, Hivert MF, et al.; MAGIC investigators. Detailed physiologic characterization reveals diverse mechanisms for novel genetic Loci regulating glucose and insulin metabolism in humans. *Diabetes* 2010;59:1266–1275
31. Locke AE, Kahali B, Berndt SI, et al.; LifeLines Cohort Study; ADIPOGen Consortium; AGEN-BMI Working Group; CARDIOGRAMplusC4D Consortium; CKDGen Consortium; GLGC; ICBP; MAGIC Investigators; MuTHER Consortium; MIGen Consortium; PAGE Consortium; ReproGen Consortium; GENIE Consortium; International Endogene Consortium. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015;518:197–206
32. Accelerating Medicines Partnership. Type 2 Diabetes Knowledge Portal. Accessed 19 April 2020. Available from <http://www.type2diabetesgenetics.org>
33. Zhou K, Yee SW, Seiser EL, et al.; MetGen Investigators; DPP Investigators; ACCORD Investigators. Variation in the glucose transporter gene SLC2A2 is associated with glycemic response to metformin. *Nat Genet* 2016;48:1055–1059
34. Santer R, Schneppenheim R, Dombrowski A, Götze H, Steinmann B, Schaub J. Mutations in GLUT2, the gene for the liver-type glucose transporter, in patients with Fanconi-Bickel syndrome. *Nat Genet* 1997;17:324–326
35. Kirwan JP, Hauguel-De Mouzon S, Lepercq J, et al. TNF- α is a predictor of insulin resistance in human pregnancy. *Diabetes* 2002;51:2207–2213
36. Guillemette L, Lacroix M, Battista MC, et al. TNF α dynamics during the oral glucose tolerance test vary according to the level of insulin resistance in pregnant women. *J Clin Endocrinol Metab* 2014;99:1862–1869
37. Moen GH, LeBlanc M, Sommer C, et al. Genetic determinants of glucose levels in pregnancy: genetic risk scores analysis and GWAS in the Norwegian STORK cohort. *Eur J Endocrinol* 2018;179:363–372
38. Ding M, Chavarro J, Olsen S, et al. Genetic variants of gestational diabetes mellitus: a study of 112 SNPs among 8722 women in two independent populations. *Diabetologia* 2018;61:1758–1768
39. Stuebe AM, Wise A, Nguyen T, Herring A, North KE, Siega-Riz AM. Maternal genotype and gestational diabetes. *Am J Perinatol* 2014;31:69–76
40. Udler MS, Powe CE, Austin-Tse CA. Case 6-2020: a 34-year-old woman with hyperglycemia. *N Engl J Med* 2020;382:745–753
41. Freathy RM, Bennett AJ, Ring SM, et al. Type 2 diabetes risk alleles are associated with reduced size at birth. *Diabetes* 2009;58:1428–1433
42. Barbour LA, Farabi SS, Friedman JE, et al. Postprandial triglycerides predict newborn fat more strongly than glucose in women with obesity in early pregnancy. *Obesity (Silver Spring)* 2018;26:1347–1356
43. Stumvoll M, Mitrakou A, Pimenta W, et al. Use of the oral glucose tolerance test to assess insulin release and insulin sensitivity. *Diabetes Care* 2000;23:295–301
44. Stumvoll M, Van Haefen T, Fritsche A, Gerich J. Oral glucose tolerance test indexes for insulin sensitivity and secretion based on various availabilities of sampling times. *Diabetes Care* 2001;24:796–797
45. Matsuda M, DeFronzo RA. Insulin sensitivity indices obtained from oral glucose tolerance testing: comparison with the euglycemic insulin clamp. *Diabetes Care* 1999;22:1462–1470
46. Radaelli T, Farrell KA, Huston-Presley L, et al. Estimates of insulin sensitivity using glucose and C-Peptide from the hyperglycemia and adverse pregnancy outcome glucose tolerance test. *Diabetes Care* 2010;33:490–494