Genetic Information and the Prediction of Incident Type 2 Diabetes in a High-Risk Multiethnic Population

The EpiDREAM genetic study

Sonia S. Anand, md, phd[1,2,3]
David Meyre, phd[1,3,4]
Guillaume Pare, md[1,3,4]
Swneke D. Bailey, phd[5]
Changchun Xie, phd[6]
Xiaohe Zhang, msc[1,2]
Alexandre Montpetit, phd[7]
Dipika Desai, msc[1]
Jackie Bosch, msc[1,8]
Viswanathan Mohan, mbbs[9]
Rafael Diaz, md[10]
Matthew J. McQueen, md, chb, phd[1,2,4]
Heather J. Cordell, dphil[11]
Bernard Keavney, mbbs, dphil[12]
Salim Yusuf, mbbs, dphil[1,2,3]
Daniel Gaudet, md, phd[13]
Hertzel Gerstein, md, msc[1,2,3]
James C. Engert, phd[14,15]
on behalf of the EpiDREAM Genetics Investigators

OBJECTIVE—To determine if 16 single nucleotide polymorphisms (SNPs) associated with type 2 diabetes (T2DM) in Europeans are also associated with T2DM in South Asians and Latinos and if they can add to the prediction of incident T2DM in a high-risk population.

RESEARCH DESIGN AND METHODS—In the EpiDREAM prospective cohort study, physical measures, questionnaires, and blood samples were collected from 25,063 individuals at risk for dysglycemia. Sixteen SNPs that have been robustly associated with T2DM in Europeans were genotyped. Among 15,466 European, South Asian, and Latino subjects, we examined the association of these 16 SNPs alone and combined in a gene score with incident cases of T2DM (n = 1,016) that developed during 3.3 years of follow-up.

RESULTS—Nine of the 16 SNPs were significantly associated with T2DM, and their direction of effect was consistent across the three ethnic groups. The gene score was significantly higher among subjects who developed incident T2DM (cases vs. noncases: 16.47 [2.50] vs. 15.99 [2.56]; P = 0.00001). The gene score remained an independent predictor of incident T2DM, with an odds ratio of 1.08 (95% CI 1.05–1.11) per additional risk allele after adjustment for T2DM risk factors. The gene score in those with no family history of T2DM was 16.02, whereas it was 16.19 in those with one parent with T2DM and it was 16.32 in those with two parents with T2DM (P trend = 0.0004). The C statistic of T2DM risk factors was 0.708 (0.691–0.725) and increased only marginally to 0.714 (0.698–0.731) with the addition of the gene score (P for C statistic change = 0.0052).

CONCLUSIONS—T2DM genetic associations are generally consistent across ethnic groups, and a gene score only adds marginal information to clinical factors for T2DM prediction.

Diabetes Care 36:2836–2842, 2013

Type 2 diabetes (T2DM) has increased dramatically worldwide, a trend that is expected to worsen in the next two decades, in large part because of the increasing prevalence of obesity (1). The clinical factors and health behaviors that predict the development of T2DM are well characterized and include elevated fasting glucose, weight, low levels of physical activity, smoking, and dietary factors (2). Several clinical risk scores have been developed to predict the risk of adults developing T2DM but are not widely used (3). Recent genetic association studies have shown that >60 single nucleotide polymorphisms (SNPs) are associated with T2DM (4). Whereas the association with T2DM of some of these gene variants is consistent across various ethnicities, other SNPs demonstrate an ethnicity-specific heterogeneity of effect size with T2DM (5). Although the individual contribution of each genetic variant to the risk of diabetes is modest, their frequency in the population make their collective impact potentially important. Previous evaluations of the addition of genetic variants to established clinical risk factors have demonstrated little gain in T2DM prediction, but most of these studies have been conducted in European populations (6,7). We sought to determine if 16 SNPs reproducibly associated with T2DM in Europeans also were associated with T2DM in South Asians and Latinos and if use of a gene score derived

From the [1]Population Health Research Institute, McMaster University and Hamilton Health Sciences, Hamilton General Hospital, Hamilton, Ontario, Canada; the [2]Department of Medicine, McMaster University, Hamilton, Ontario, Canada; the [3]Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada; the [4]Department of Pathology and Molecular Medicine, McMaster University, Hamilton, Ontario, Canada; the [5]Ontario Cancer Institute/University Health Network, Toronto, Ontario, Canada; the [6]Department of Environmental Health, College of Medicine, University of Cincinnati, Kettering Laboratory, Cincinnati, Ohio; the [7]McGill University and Genome Quebec Innovation Centre, Montreal, Quebec, Canada; the [8]School of Rehabilitation Science, McMaster University, Hamilton, Ontario, Canada; the [9]Madras Diabetes Research Foundation, Gopalapuram, Chennai, India; [10]Estudios Cardiológicos Latinoamérica, Rosario, Argentina; [11]Newcastle University, Institute of Genetic Medicine, International Centre for Life, Newcastle upon Tyne, U.K.; [12]Newcastle University, Institute of Human Genetics, International Centre for Life, Newcastle upon Tyne, U.K.; the [13]Faculty of Medicine, Université de Montréal, Montreal, Quebec, Canada; the [14]Department of Medicine, McGill University, Montreal, Quebec, Canada; and the [15]Department of Human Genetics, McGill University, Montreal, Quebec, Canada.
Corresponding author: Sonia S. Anand, anands@mcmaster.ca.
Received 7 December 2012 and accepted 25 February 2013.
DOI: 10.2337/dc12-2553
This article contains Supplementary Data online at http://care.diabetesjournals.org/lookup/suppl/doi:10.2337/dc12-2553/-/DC1.

from the 16 SNPs adds to the risk prediction of incident T2DM over clinical factors alone in a multiethnic population at high risk for T2DM.

RESEARCH DESIGN AND METHODS

EpiDREAM cohort

The EpiDREAM study included 25,063 participants from 191 centers who were screened for the DREAM clinical trial (8). Individuals at risk for dysglycemia because of family history, ethnicity, and abdominal obesity, between the ages of 18 and 85 years, were screened using an oral glucose tolerance test from July 2001 to August 2003 as previously described (8,9). Five thousand two hundred sixty-nine individuals with impaired glucose tolerance (IGT) or impaired fasting glucose (IFG) (or both) participated in the DREAM trial (9). Thirteen thousand seven hundred twenty-one individuals who were not eligible for the DREAM trial or who declined to enter the trial agreed to long-term follow-up. Thus, in total, 18,990 individuals were prospectively followed up for a median of 3.3 years and 18,486 individuals who provided a blood sample for DNA were available for analysis.

DNA extraction and genotyping

Buffy coat DNA extractions and genotyping using the Illumina CVD bead chip microarray ITMAT Broad Care (IBC) array was successful for 19,197 subjects (10). Genotyping was performed at the McGill University and Genome Quebec Innovation Centre using the Illumina Bead Studio genotyping module, version 3.2. Of the 19,197 genotyped samples, 711 subjects were excluded because of genotyping failure (defined as <97% of genotypes called), leaving 18,486 for analysis; 15,466 of these (including 1,016 incident cases) were of European, South Asian, and Latino origin and 12,249 were followed prospectively (Supplementary Fig. 1). Because the focus of this article is on the prediction of incident T2DM, only the following three largest ethnic groups with >100 incident cases of T2DM were included: Europeans (9,408, with 599 new T2DM cases), South Asians (2,764, with 194 new T2DM cases), and Latinos (3,294, with 223 new T2DM cases).

To verify self-reported ethnicity, we performed a principal component analysis using the genotype data of all EpiDREAM study participants using the software smartpca, which is part of the eigensoft software package (http://genepath.med.harvard.edu/~reich/software.htm). Samples that failed to cluster with individuals of the same self-reported ethnicity were removed.

Genes selected for the gene score

Sixteen SNPs from 16 genes were selected to be included in the gene score because they have been previously associated with T2DM in Europeans at genome-wide significance ($P < 5 \times 10^{-8}$) (11,12) in the literature and were included in versions 1 and 2 of the IBC array (Supplementary Table 1). We tested for Hardy-Weinberg equilibrium for each of the 16 SNPs within each ethnic group ($P > 0.001$). The call rate for each of the 16 SNPs was >99.5%.

Dysglycemia definitions

The 2003 American Diabetes Association criteria were used to classify participants as normal, as having IFG, as having IGT, or as having T2DM as confirmed by the baseline oral glucose tolerance test. Normoglycemia was defined as a fasting plasma glucose <5.6 mmol/L, IFG was defined as a fasting plasma glucose of 5.6 to 6.9 mmol/L, IGT was defined as a fasting plasma glucose <7.0 mmol/L and a 2-h glucose between 7.8 and 11.0 mmol/L, and diabetes was defined if either the fasting plasma glucose was ≥7.0 mmol/L or the 2-h glucose was ≥11.1 mmol/L (13).

Statistical methods

Statistical analyses were performed using PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/) and SAS (version 9.1.3; SAS Institute). For each of the 16 SNPs, the previously identified risk allele was used as the risk allele in each ethnic group. To maximize our power (range 61–100%) to investigate the association between each SNP and T2DM among ethnic groups, we combined the baseline and incident cases of T2DM (n = 3,209). The association between each SNP and baseline and incident cases of T2DM was calculated using logistic regression. A genotype score was calculated for each individual by allele counting, and thus the score could range from 0 to 32. Baseline factors associated with the development of diabetes were tested versus incident T2DM with logistic regression adjusting for age, sex, and ethnicity. We constructed multivariate logistic regression models to determine the predictive ability of the gene score on incident diabetes in addition to clinical factors, age, sex, and ethnicity. The change in the C statistic was calculated after addition of the gene score to the clinical factors found to be associated with an increased risk of T2DM. The change in the receiver operating characteristic was calculated. Interactions between the gene score and ethnicity were tested. Model calibration was assessed using the Hosmer Lemeshow test. The population-attributable risk (PAR) of the gene score was calculated using the methods of Benichou and Gail (14).

RESULTS

Baseline characteristics

For the 15,466 subjects included in this analysis, the mean age was 53 years, ~60% were women, and more than half of the participants were of European origin. At baseline, 6,697 (43.3%) individuals were classified as normoglycemic, 6,538 (42.3%) had either IFG or IGT (or both), and 2,231 (14.4%) had type 2 diabetes. Additional details of baseline characteristics are found in Table 1. Ethnic comparisons of baseline characteristics revealed that people of South Asian origin were significantly younger, mostly male, and had a lower waist circumference and BMI compared with Latinos and Europeans. Furthermore, they were less likely to smoke, were less physically active, had lower systolic and diastolic blood pressures, had lower apolipoprotein B, and had less hypertension. Despite being younger and having fewer other cardiovascular risk factors, South Asians were more likely to be diabetic and have lower apolipoprotein A compared with the other ethnic groups (Table 2).

Ethnic variation of SNP frequencies and association with T2DM

Each of the 16 SNPs was polymorphic, and the risk allele frequencies in the three ethnic groups are shown in Supplementary Table 1. Of the 16 SNPs, the European minor allele was the risk allele for nine SNPs, and the common allele was the risk allele for seven. Of the 16 SNPs, nine (56%) were significantly associated with T2DM, and the direction of effect was generally consistent across ethnic groups (Supplementary Table 2). Considering incident cases only (n = 1,016), in which our power was lower, six out of 16 SNPs were significantly associated with incident T2DM; however, the direction and magnitude of effects were generally consistent across ethnic groups.

Gene score

The mean gene score overall was 16.06 (SD 2.58), and the median was 16 (interquartile

Table 1—*Baseline characteristics of EpiDREAM subjects*

Characteristics	Overall	Normoglycemic	IFG and/or IGT	Diabetes at baseline	*P* for group difference
N	15,466	6,697	6,538	2,231	—
Age, years (SD)	52.7 (11.4)	49.4 (11.1)	54.98 (11.0)	55.7 (11.0)	<0.001
Male (%)	6,210 (40.1)	2,294 (34.2)	2,836 (43.4)	1,080 (48.4)	<0.0001
Waist, cm (SD)	95.3 (14.5)	91.5 (14.0)	97.6 (14.1)	99.6 (14.4)	<0.0001
Hip, cm (SD)	107.0 (13.2)	105.3 (12.7)	108.3 (13.4)	108.6 (13.6)	<0.0001
BMI, kg/m^2 (SD)	29.8 (5.7)	28.64 (5.6)	30.53 (5.5)	30.91 (5.8)	<0.0001
Waist-to-hip ratio (SD)	0.89 (0.09)	0.87 (0.09)	0.90 (0.09)	0.92 (0.09)	<0.0001
Family history of diabetes (%)	7,443 (48.4)	3,340 (50.1)	3,048 (46.9)	1,055 (47.5)	0.00076
European (%)	9,408 (60.8)	3,669 (54.8)	4,379 (67.0)	1,360 (61.0)	<0.0001
Current smoker (%)	2,258 (14.6)	1,131 (16.9)	829 (12.7)	298 (13.4)	<0.0001
High physical activity (%)	1,947 (12.6)	824 (12.3)	883 (13.5)	240 (10.8)	0.0019
Moderate/high physical activity (%)	11,226 (72.7)	4,739 (70.8)	4,931 (75.6)	1,556 (69.8)	<0.0001
Systolic blood pressure (SD)	132.6 (19.1)	126.8 (18.1)	135.8 (18.3)	140.2 (20.0)	<0.0001
Diastolic blood pressure (SD)	81.9 (11.0)	79.38 (10.9)	83.33 (10.7)	85.01 (10.9)	<0.0001
Apolipoprotein B, g/L (SD)	0.93 (0.23)	0.90 (0.23)	0.95 (0.22)	0.98 (0.24)	<0.0001
History of HTN with Rx (%)	3,559 (23.0)	910 (13.6)	1,943 (29.7)	706 (31.6)	<0.0001
Apolipoprotein A1, g/L (SD)	1.41 (0.31)	1.41 (0.32)	1.41 (0.31)	1.39 (0.30)	0.0096
Mean gene score (SD)	16.06 (2.58)	15.87 (2.58)	16.12 (2.55)	16.45 (2.58)	<0.0001
Nontrial participants (%)	11,940 (77.2)	6,697 (100.0)	3,012 (46.1)	2,231 (100.0)	<0.0001

HTN, hypertension; Rx, medication.

range 14–18). The mean gene score increased progressively from 15.87 (SD 2.58) in normoglycemic subjects to 16.12 (SD 2.55) in subjects with IFG/IGT and to 16.45 (SD 2.58) in those with T2DM at baseline (*P* between-group difference = 5.9×10^{-20}) (Table 1). The mean gene score in those with no family history of T2DM was 16.02 (SE 0.03), whereas in those with one parent with a history of T2DM mean gene score was 16.19 (SE 0.04) and among those with two parents with T2DM it was 16.32 (SE 0.08; *P* trend = 0.0004) when adjusted for age, sex, and ethnicity (Fig. 1). The gene score in South Asians was higher than in Europeans and Latinos (16.65 [SE 0.05] vs. 16.03 [SE 0.03] and 15.66 [SE 0.05], $P < 0.0001$) after adjustment for age and sex.

Table 2—*Baseline characteristics by ethnic group*

Characteristics	Overall	Europeans	South Asians	Latinos	*P* for group difference
N	15,466	9,408	2,764	3,294	—
Age, years (SD)	52.7 (11.4)	55.0 (10.8)	44.9 (9.4)	52.6 (11.6)	<0.001
Male (%)	6,210 (40.1)	3,691 (39.2)	1,426 (51.6)	1,093 (33.2)	<0.0001
Waist, cm (SD)	95.3 (14.5)	96.5 (15.0)	89.3 (11.3)	96.9 (14.1)	<0.0001
Hip, cm (SD)	107.0 (13.2)	108.5 (13.4)	100.3 (10.4)	108.4 (12.9)	<0.0001
BMI, kg/m^2 (SD)	29.8 (5.7)	30.4 (5.7)	26.4 (4.3)	30.8 (5.7)	<0.0001
Waist-to-hip ratio (SD)	0.89 (0.09)	0.89 (0.09)	0.89 (0.09)	0.89 (0.08)	0.031
Family history of diabetes (%)	7,443 (48.4)	4,544 (48.6)	1,312 (47.5)	1,587 (48.5)	0.058
History of HTN with Rx (%)	3,559 (23.0)	2,404 (25.6)	244 (8.8)	911 (27.7)	<0.001
Current smoker (%)	2,258 (14.6)	1,327 (14.1)	251 (9.1)	680 (20.6)	<0.0001
High physical activity (%)	1,947 (12.6)	1,603 (17.1)	111 (4.0)	233 (7.1)	<0.0001
Moderate/high physical activity (%)	11,226 (72.7)	7,702 (82.0)	1,233 (44.6)	2,291 (69.5)	<0.0001
Systolic blood pressure (SD)	132.6 (19.1)	134.1 (18.6)	124.34 (17.6)	135.09 (20.1)	<0.0001
Diastolic blood pressure (SD)	81.9 (11.0)	82.5 (10.7)	78.3 (11.0)	82.9 (11.4)	<0.0001
Apolipoprotein B, g/L (SD)	0.93 (0.23)	0.95 (0.23)	0.86 (0.22)	0.95 (0.23)	<0.001
Apolipoprotein A1, g/L (SD)	1.41 (0.31)	1.48 (0.32)	1.18 (0.23)	1.38 (0.27)	<0.0001
New DM (%)	1,016 (9.5)	599 (10.9)	194 (8.4)	223 (7.8)	<0.0001
IFG, new definition (%)	4,989 (32.3)	3,449 (36.7)	416 (15.0)	1,124 (34.1)	<0.0001
IGT (%)	3,814 (24.7)	2,470 (26.2)	569 (20.6)	775 (23.5)	<0.0001
Area under curve glucose (SD)	13.2 (4.2)	13.2 (3.7)	13.32 (5.8)	12.8 (3.8)	<0.0001
Mean gene score (SD)	16.1 (2.6)	16.0 (2.6)	16.7 (2.4)	15.7 (2.5)	<0.0001
Nontrial participants (%)	11,940 (77.2)	7,135 (75.8)	2,372 (85.8)	2,433 (73.9)	<0.0001

Includes patients with diabetes at baseline. DM, diabetes mellitus; HTN, hypertension; Rx, medication.

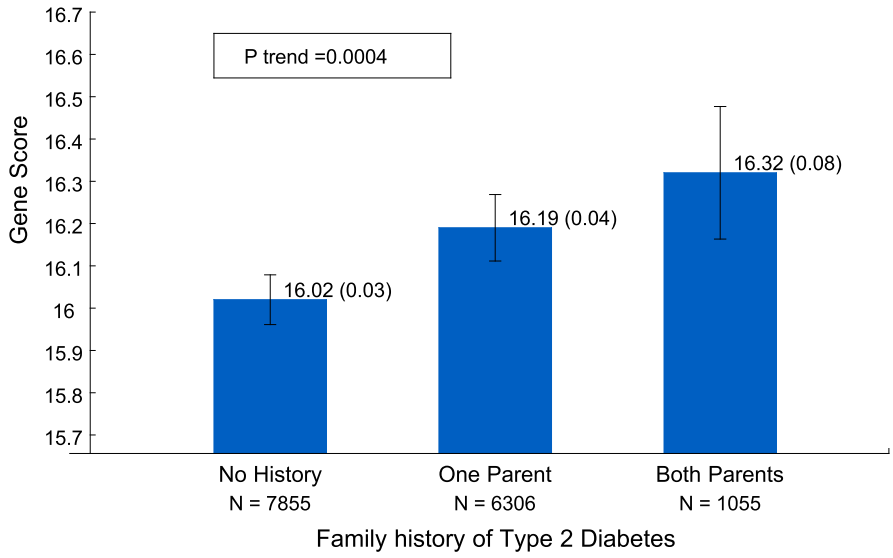

Figure 1—*Mean gene score and SE (y-axis) by category of reported family history (no parental history of T2DM, one parent with T2DM, and two parents with T2DM) on the x-axis.*

Association between gene score and incident T2DM

Individuals classified with T2DM at baseline were excluded (n = 2,231). Of the 10,640 subjects who completed the follow-up, 1,016 (9.55%) developed incident T2DM after a mean of 3.3 years. Characteristics of those who did and did not develop T2DM are shown in Supplementary Table 3. The gene score was significantly higher among subjects who developed incident T2DM (cases) compared with those who did not (cases vs. noncases: 16.47 [2.50] vs. 15.99 [2.56]; $P < 0.00001$). Almost 9% (n = 1,378) of subjects possessed a gene score ≥20, and the relative risk of incident T2DM was 4.66 (95% CI 1.87–11.62) compared with a gene score ≤10. We constructed a multivariate model for incident T2DM and included factors that differed significantly between incident T2DM cases and nondiabetic subjects on univariate analysis to assess whether the gene score would remain an independent predictor of T2DM. After adjustment for age, sex, ethnic group, BMI, waist circumference, family history of T2DM, smoking, moderate or high activity, apolipoprotein B, apolipoprotein A1, hypertension with medication, and trial status, the gene score remained a significant predictor of incident T2DM (Table 3). Specifically, for each additional risk allele, the risk of T2DM increased by 8% on average (odds ratio 1.08 [95% CI 1.05–1.10]). This finding is consistent across ethnic groups (P heterogeneity = 0.46) (Fig. 2).

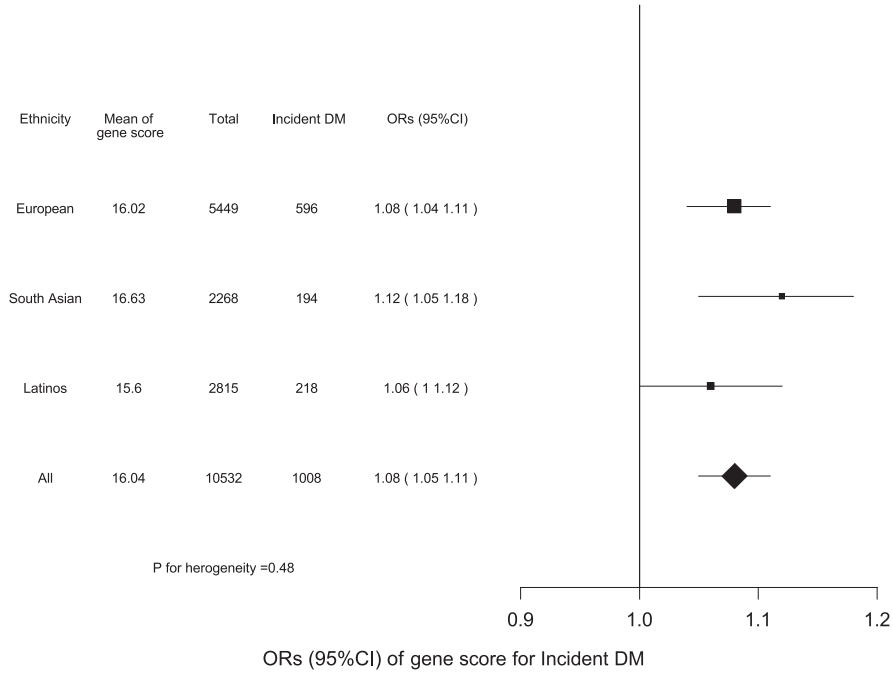

Figure 2—*Mean gene score by ethnic group and in all groups combined, the number of subjects in each group, the number of incident T2DM events per group, and the per allele odds ratio (OR) and 95% CI of incident T2DM per group and overall. DM, diabetes mellitus.*

Prediction of incident T2DM

We used the C statistic to assess the incremental value of adding the gene score to the clinical factors. When the clinical factors (BMI, waist circumference, family history of T2DM, smoking, moderate or high physical activity, apolipoprotein B, apolipoprotein A1, hypertension with medication, and trial status) together with age, sex, and ethnicity were included in the model, the C statistic equaled 0.708 (95% CI 0.691–0.725), and the gene score increased the discrimination only marginally to 0.714 (0.698–0.731), which represents a change in the area under curve of 0.006 (P = 0.0052). The C statistic was 0.599 (0.580–0.617) with only gene score, age, sex, and ethnicity in the model.

PAR

The PAR of the gene score ≥20 on incident T2DM after adjustment for all clinical predictors is 5.6% (95% CI 3.6–8.6). The PAR for the clinical factors is 55.4% (45.2–65.2), and with the addition of the gene score ≥20 the PAR increases by 2.8–58.2% (48.3–67.5).

Consistency in subgroups

We investigated effect modification of the gene score by certain phenotypic characteristics, including age, sex, ethnic group, family history of T2DM, BMI, level of physical activity, and presence of IFG or IGT at baseline. No significant interactions between the gene score and any of these parameters on incident T2DM were observed (Supplementary Fig. 2). In

Table 3—*Multivariate logistic regression results predicting incident diabetes*

Factor	Beta (SE)	Odds ratio (95% CI)	*P*
Age	0.01 (0.004)	1.01 (1.00–1.02)	0.003
Sex, female vs. male	−0.25 (0.089)	0.78 (0.65–0.92)	0.004
Ethnicity, Latino vs. European	−0.22 (0.09)	0.80 (0.68–0.96)	0.01
Ethnicity, South Asian vs. European	0.43 (0.12)	1.53 (1.21–1.93)	0.0003
Trial status, EpiDREAM vs. DREAM	−1.17 (0.07)	0.31 (0.27–0.36)	<0.0001
Gene score*	0.08 (0.01)	1.08 (1.05–1.11)	<0.0001
BMI	0.05 (0.01)	1.05 (1.03–1.07)	<0.00001
WC	0.002 (0.004)	1.00 (0.99–1.01)	0.06
Family history of DM	0.16 (0.07)	1.18 (1.02–1.36)	0.02
Smoking	0.17 (0.08)	1.19 (1.03–1.38)	0.02
Moderate/high activity	0.04 (0.08)	1.04 (0.88–1.22)	0.06
Apolipoprotein B	0.29 (0.16)	1.34 (0.98–1.83)	0.06
Apolipoprotein A	−0.26 (0.14)	0.77 (0.58–1.02)	0.07
HTN with Rx	0.16 (0.08)	1.18 (1.00–1.39)	0.05

DM, diabetes mellitus; HTN, hypertension; Rx, medication; WC, waist circumference. *Intercept: −5.63 (0.50); $P = 9.7 \times 10^{-25}$.

addition, after adjusting for multiple testing, no significant interactions between any pairs of SNPs on incident T2DM were observed (data not shown).

CONCLUSIONS—In a high-risk multiethnic population, a gene score was significantly associated with T2DM. However, the clinical factors were far more predictive and the addition of the gene score did not greatly improve the prediction of T2DM. This finding was consistent across ethnic groups.

For most SNPs, the association with incident T2DM is small to moderate (odds ratio range 1.10–1.40), yet the collective genetic load as reflected by the gene score is an efficient way to investigate the contribution of genetic predisposition to population variation of T2DM (15,16). We observed significant variation in the mean gene score between ethnicities, by categories of dyglycemia, and among individuals who reported a family history of T2DM. First, our data suggest that South Asians may have a greater genetic load for T2DM compared with Europeans and Latinos. This would be consistent with the lower risk factor profile for South Asians and with the work of Chen et al. (5). Second, the gene score increased progressively with increasing severity of dysglycemia and was lowest in those normoglycemic at baseline, intermediate in those with IGT or IFG, and highest in those with established T2DM. Third, individuals with a family history of T2DM had progressively higher gene scores, from individuals with no family history to a single parent to two parents with T2DM, reinforcing previous observations that family history may be a crude marker of genetic load for conditions such as T2DM and myocardial infarction (17). However, the fact that the gene score remains significant for incident T2DM even after adjusting for the family history information suggests that adding specific genetic information is more informative than the T2DM family history information alone (7). Although these observations suggest that a gene score may stratify risk groups for T2DM, it is currently far less expensive and more efficient to record ethnicity, to measure fasting glucose, and to take a family history.

Our results are consistent with other studies that have evaluated the change in prediction of T2DM with the addition of a gene score, although the majority of these studies have been conducted in Europeans only (3,6,7). To our knowledge, the gene score approach has not been evaluated in South Asians or Latinos or in multiple populations in which the methods of phenotypic and genetic assessment were standardized across ethnic groups. In EpiDREAM, the gene score was significantly associated with incident T2DM (a 1-point increase is associated with an 8% increase in risk), and individuals with a gene score ≥20 have an odds ratio of 4.66 (95% CI 1.87–11.62) for incident T2DM compared with those with a gene score ≤10. However, the C statistic only increased marginally when the gene score was added to clinical factors. Furthermore, a gene score ≥20 only added 2.8% to the PAR when combined with clinical factors. Thus, the addition of genome-wide association study–identified genetic variants does not greatly improve the prediction of future cases of T2DM. Recent theoretical work has suggested that numerous T2DM-associated common variants remain to be identified (18), and it remains to be seen if, when summed together, they will improve discrimination of future cases.

We note that the median length of follow-up in this study was 3.3 years. Interestingly, the mean gene score of 16.45 among T2DM cases at baseline (prevalent cases) and the gene score among subjects who developed incident T2DM (16.47) are remarkably similar. Thus, by confining our predictive analysis to incident cases and using a relatively short follow-up, we have likely underestimated the life-long contribution of the genetic risk relative to the contribution of clinical variables.

EpiDREAM enrolled individuals at risk for dysglycemia defined by family history, ethnicity, or abdominal obesity, representing the breadth of at-risk individuals. This is consistent with the recruitment heterogeneity seen in recent large-scale meta-analyses confirming the SNPs that were used in our analysis (12,19). However, the genetic contribution to T2DM in each of these categories could be different and could contribute to heterogeneous results for specific risk factor classes.

We did not observe a significant interaction between the gene score and any subgroup, including ethnicity, on the risk of incident T2DM. Although in our study the gene score was higher in South Asians, the gene score performed similarly between ethnic groups, as well as between men and women, between old and young, and among different T2DM family histories (Supplementary Fig. 2). Other investigators have evaluated T2DM genetic associations across multiple ethnicities with conflicting results, possibly attributable to limited statistical power, allelic heterogeneity, or linkage disequilibrium differences (20,21). In our study, we observed minimal heterogeneity in the direction of effect between SNPs and T2DM, although for some SNPs our power within certain ethnicities (South Asian and Latinos) was low. Importantly, the gene score appeared to perform similarly across the three ethnic groups.

In a recent article, Chen et al. (5) showed that T2DM-associated SNPs

demonstrated an extreme level of interethnic risk allele differentiation in comparison with other diseases. Our study, in three ethnic groups, showed significant interethnic heterogeneity in allele frequency for 16 SNPs and confirmed this observation (5). Chen et al. (5) also demonstrated that the cumulative risk allele frequencies of T2DM SNPs may be higher in Africans, intermediate in Europeans, and lower in East Asians. However, we observed that in comparison with Europeans and Latinos, South Asians had a higher risk allele frequency for eight out of 16 SNPs (Supplementary Table 1) and had a higher overall T2DM gene score. This suggests that complex, regional, and SNP-specific patterns of evolution of T2DM variants have occurred over time.

Strengths and limitations

Our study had several strengths, including a large multiethnic sample, detailed oral glucose tolerance test information at baseline, and a clinically relevant cohort at risk for dysglycemia. Furthermore, the cohort included phenotypes recorded in a standardized way and subjects who were prospectively followed up. Limitations of our study include limited power to explore ethnic-specific associations for some SNPs and the relatively low number of T2DM-associated SNPs included in our gene score, because now >60 exist (10–12).

Conclusion

Genetic associations for T2DM are generally consistent across ethnic groups, and a genetic risk score adds marginal information to clinical factors in the prediction of T2DM.

Acknowledgments—This study was funded by a grant from the Canadian Institutes of Health Research University Industry competition with partner funding from the GlaxoSmithKline and Sanofi Aventis Global, Sanofi Aventis Canada, Genome Quebec Innovation Centre, Heart and Stroke Foundation of Canada.

S.S.A. holds a Canada Research Chair in Ethnic Diversity and Cardiovascular Disease, and is the Michael G. DeGroote and Heart and Stroke Foundation of Ontario Chair in Population Health, McMaster University. D.M. holds a Canada Research Chair in Genetics of Diabetes and Obesity at McMaster University. G.P. holds the Canada Research Chair in Molecular Epidemiology and Genetics at McMaster University. H.G. holds the Population Health Research Institute Chair in Diabetes Research sponsored by Aventis. D.G. holds the Canada Research Chair in Preventive Genetics and Community Genomics. S.Y. holds a Heart and Stroke Foundation Chair in Cardiovascular Research, McMaster University. S.S.A. has received lecture fees from GlaxoSmithKline and Sanofi Aventis. V.M. has received honoraria/research grants from Novo Nordisk, LifeScan, J & J, MSD Pharmaceuticals, USV, Sanofi Aventis, Roche Diagnotics, and Abbott Nutrition International. No other potential conflicts of interest relevant to this article were reported.

S.S.A. participated in data collection and analysis and wrote the manuscript. D.M. participated in data analysis and wrote the manuscript. G.P. reviewed and edited the manuscript. S.D.B. contributed to data analysis and reviewed and edited the manuscript. C.X., X.Z., A.M., and D.D. contributed to data and laboratory analysis. J.B., V.M., H.G., S.Y., and R.D. participated in data collection and reviewed and edited the manuscript. M.J.M. and H.J.C. contributed to data and laboratory analysis and reviewed and edited the manuscript. B.K. and D.G. reviewed and edited the manuscript. J.C.E. participated in analysis and wrote the manuscript. S.S.A. 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.

References

1. Shaw JE, Sicree RA, Zimmet PZ. Global estimates of the prevalence of diabetes for 2010 and 2030. Diabetes Res Clin Pract 2010;87:4–14
2. Wareham NJ, Griffin SJ. Risk scores for predicting type 2 diabetes: comparing axes and spades. Diabetologia 2011;54:994–995
3. Noble D, Mathur R, Dent T, Meads C, Greenhalgh T. Risk models and scores for type 2 diabetes: systematic review. BMJ 2011;343:d7163
4. Ntzani EE, Kavvoura FK. Genetic risk factors for type 2 diabetes: insights from the emerging genomic evidence. Curr Vasc Pharmacol 2012;10:147–155
5. Chen R, Corona E, Sikora M, et al. Type 2 diabetes risk alleles demonstrate extreme directional differentiation among human populations, compared to other diseases. PLoS Genet 2012;8:e1002621
6. Meigs JB, Shrader P, Sullivan LM, et al. Genotype score in addition to common risk factors for prediction of type 2 diabetes. N Engl J Med 2008;359:2208–2219
7. Lyssenko V, Jonsson A, Almgren P, et al. Clinical risk factors, DNA variants, and the development of type 2 diabetes. N Engl J Med 2008;359:2220–2232
8. Anand SS, Dagenais GR, Mohan V, et al. Glucose levels are associated with cardiovascular disease and death in an international cohort of normal glycaemic and dysglycaemic men and women: the EpiDREAM cohort study. Eur J Cardiovasc Prev Rehabil 2012;19:755-764
9. Gerstein HC, Yusuf S, Bosch J, et al.; DREAM (Diabetes REduction Assessment with ramipril and rosiglitazone Medication) Trial Investigators. Effect of rosiglitazone on the frequency of diabetes in patients with impaired glucose tolerance or impaired fasting glucose: a randomised controlled trial. Lancet 2006;368:1096–1105
10. Keating BJ, Tischfield S, Murray SS, et al. Concept, design and implementation of a cardiovascular gene-centric 50 k SNP array for large-scale genomic association studies. PLoS ONE 2008;3:e3583
11. Voight BF, Scott LJ, Steinthorsdottir V, et al.; MAGIC investigators. GIANT Consortium. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet 2010;42:579–589
12. Saxena R, Elbers CC, Guo Y, et al.; Look AHEAD Research Group; DIAGRAM consortium. Large-scale gene-centric meta-analysis across 39 studies identifies type 2 diabetes loci [corrected in Am J Hum Genet 2012;90;753]. Am J Hum Genet 2012;90:410–425
13. Genuth S, Alberti KG, Bennett P, et al.; Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Follow-up report on the diagnosis of diabetes mellitus. Diabetes Care 2003;26:3160–3167
14. Benichou J, Gail MH. Variance calculations and confidence intervals for estimates of the attributable risk based on logistic models. Biometrics 1990;46:991–1003
15. Mihaescu R, van Zitteren M, van Hoek M, et al. Improvement of risk prediction by genomic profiling: reclassification measures versus the area under the receiver operating characteristic curve. Am J Epidemiol 2010;172:353–361
16. Travers ME, McCarthy MI. Type 2 diabetes and obesity: genomics and the clinic. Hum Genet 2011;130:41–58
17. Chow CK, Islam S, Bautista L, et al. Parental history and myocardial infarction risk across the world: the INTERHEART Study. J Am Coll Cardiol 2011;57:619–627
18. Stahl EA, Wegmann D, Trynka G, et al.; Diabetes Genetics Replication and Meta-analysis Consortium; Myocardial Infarction Genetics Consortium. Bayesian inference analyses of the polygenic architecture of rheumatoid arthritis. Nat Genet 2012;44:483–489
19. Morris AP, Voight BF, Teslovich TM, et al.; Wellcome Trust Case Control

Consortium; Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC) Investigators; Genetic Investigation of ANthropometric Traits (GIANT) Consortium; Asian Genetic Epidemiology Network–Type 2 Diabetes (AGEN-T2D) Consortium; South Asian Type 2 Diabetes (SAT2D) Consortium; DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat Genet 2012;44:981–990

20. Waters KM, Stram DO, Hassanein MT, Le Marchand L, Wilkens LR, Maskarinec G, Monroe KR, Kolonel LN, Altshuler D, Henderson BE, Haiman CA. Consistent association of type 2 diabetes risk variants found in Europeans in diverse racial and ethnic groups. PLoS Genet. 2010;6:e1001078.
21. Sim X, Ong RT, Suo C, et al. Transferability of type 2 diabetes implicated loci in multi-ethnic cohorts from Southeast Asia. PLoS Genet 2011;7:e1001363