Variations in the G6PC2/ABCB11 genomic region are associated with fasting glucose levels

Wei-Min Chen,1,2 Michael R. Erdos,3 Anne U. Jackson,4 Richa Saxena,5 Serena Sanna,4,6 Kristi D. Silver,7 Nicholas J. Timpson,8 Torben Hansen,9 Marco Orrù,6 Maria Grazia Piras,6 Lori L. Bonnycastle,3 Cristen J. Willer,4 Valeriya Lyssenko,10 Haiqing Shen,7 Johanna Kuusisto,11 Shah Ebrahim,12 Natascia Sestu,13 William L. Duren,4 Maria Cristina Spada,6 Heather M. Stringham,4 Laura J. Scott,4 Nazario Olla,6 Amy J. Swift,8 Samer Nijjar,13 Braxton D. Mitchell,7 Debbie A. Lawlor,9 George Davey Smith,8 Yoav Ben-Shlomo,14 Gitte Andersen,9 Knut Borch-Johnsen,9,15,16 Torben Jørgensen,15 Jouko Saramies,17 Timo T. Valle,18 Thomas A. Buchanan,19,20 Alan R. Shuldiner,7 Edward Lakatta,13 Richard N. Bergman,20 Manuela Uda,6 Jaakko Tuomilehto,18,21 Oluf Pedersen,9,16 Antonio Cao,6 Leif Groop,10 Karen L. Mohlke,22 Markku Laakso,11 David Schlessinger,13 Francis S. Collins,3 David Altshuler,5 Gonçalo R. Abecasis,4 Michael Boehnke,4 Angelo Scuteri,23,24 and Richard M. Watanabe20,25

1Department of Public Health Sciences and 2Center for Public Health Genomics, University of Virginia, Charlottesville, Virginia, USA.
3Genome Technology Branch, National Human Genome Research Institute, Bethesda, Maryland, USA.
4Center for Statistical Genetics and Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, USA.
5Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA.
6Istituto di Neurogenetica e Neurofarmacologia, Consiglio Nazionale delle Ricerche, Cagliari, Italy.
7Division of Endocrinology, Diabetes and Nutrition, University of Maryland School of Medicine, Baltimore, Maryland, USA.
8MRC Centre for Causal Analyses in Translational Epidemiology, Department of Social Medicine, University of Bristol, Bristol, United Kingdom.
9Steno Diabetes Center, Gentofte, Denmark.
10Department of Clinical Sciences, Diabetes and Endocrinology, Lund University, University Hospital Malmö, Malmö, Sweden.
11Department of Medicine, University of Kuopio and Kuopio University Hospital, Kuopio, Finland.
12Department of Epidemiology and Population Health, Non-communicable Disease Epidemiology Unit, London School of Hygiene and Tropical Medicine, University of London, London, United Kingdom.
13Gerontology Research Center, National Institute on Aging, Baltimore, Maryland, USA.
14Social Medicine Department, University of Bristol, Bristol, United Kingdom.
15Research Centre for Prevention and Health, Glostrup University Hospital, Glostrup, Denmark.
16Faculty of Health Sciences, University of Aarhus, Aarhus, Denmark.
17Savitaipale Health Center, Savitaipale, Finland.
18Diabetes Unit, Department of Health Promotion and Chronic Disease Prevention, National Public Health Institute, and Department of Public Health, University of Helsinki, Helsinki, Finland.
19Department of Medicine, Division of Endocrinology, and 20Department of Physiology and Biophysics, Keck School of Medicine, University of Southern California, Los Angeles, California, USA.
21South Ostrobothnia Central Hospital, Säkkipää, Finland.
22Department of Genetics, University of North Carolina, Chapel Hill, North Carolina, USA.
23Laboratory of Cardiovascular Science, National Institute on Aging, NIH, Baltimore, Maryland, USA.
24Unità Operativa Geriatria, Istituto Nazionale Ricovero E Cura Anziani, Rome, Italy.
25Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California, USA.

Identifying the genetic variants that regulate fasting glucose concentrations may further our understanding of the pathogenesis of diabetes. We therefore investigated the association of fasting glucose levels with SNPs in 2 genome-wide scans including a total of 5,088 nondiabetic individuals from Finland and Sardinia. We found a significant association between the SNP rs563694 and fasting glucose concentrations (P = 3.5 × 10⁻⁸). This association was further investigated in an additional 18,436 nondiabetic individuals of mixed European descent from 7 different studies. The combined P value for association in these follow-up samples was 6.9 × 10⁻⁻⁶, and combining results from all studies resulted in an overall P value for association of 6.4 × 10⁻¹³. Across these studies, fasting glucose concentrations increased 0.01–0.16 mM with each copy of the major allele, accounting for approximately 1% of the total variation in fasting glucose. The rs563694 SNP is located between the genes glucose-6-phosphatase catalytic subunit 2 (G6PC2) and ATP-binding cassette, subfamily B (MDR/TAP), member 11 (ABCB11). Our results in combination with data reported in the literature suggest that G6PC2, a glucose-6-phosphatase almost exclusively expressed in pancreatic islet cells, may underlie variation in fasting glucose, though it is possible that ABCB11, which is expressed primarily in liver, may also contribute to such variation.

Introduction

Glucose is the major source of energy in humans, with levels in vivo determined by a balance of glucose absorption via the gut, production primarily by the liver, and utilization by both insulin-sensitive and insulin-insensitive tissues (1, 2). Homeostatic control of glucose levels involves complex interactions between humoral and neural mechanisms that work in concert to regulate tightly the balance between production and utilization to maintain a nor-
mal fasting glucose. Elevations in blood glucose are diagnostic of diabetes. Type 2 diabetes mellitus (T2DM) afflicts more than 171 million worldwide and is a leading cause of kidney failure, blindness, and lower limb amputations (3–5). Even more modest elevations in glucose concentration (so-called prediabetes) are associated with cardiovascular disease and accelerated atherosclerosis (6). In individuals progressing toward future T2DM, the fasting glucose concentration appears to change only modestly over time until the advent of β cell dysfunction, at which point the glucose concentration increases rapidly (7, 8). Many studies have shown that the lowering of glucose levels in individuals with diabetes can prevent or delay diabetes-related complications, providing further evidence for the damaging effects of chronic glucose elevations.

Both genetic and environmental factors contribute to the pathophysiology of T2DM (9–11). The contributions of environmental exposures to T2DM risk are best illustrated by results from the Diabetes Prevention Program (11) and the Finnish Diabetes Prevention Study (12), in which T2DM incidence was significantly reduced by intensive lifestyle modification. However, the contribution of genetic factors to T2DM risk is not as well understood. Recent genome-wide association (GWA) studies have identified 16 novel T2DM susceptibility loci (13–18), generating new insights into the genetic architecture underlying T2DM. In contrast to disease status, even less is known about genetic variation that alters specific T2DM-related quantitative traits such as glucose and insulin concentrations. As seen for T2DM, identification of genetic variants associated with T2DM-related quantitative traits is likely to require large sample sizes due to relatively small gene effect sizes. Fasting glucose concentrations have been shown to be heritable, with narrow-sense heritability estimates ranging from 25% to 40% (19–24). Given the central role of glucose concentration in the pathogenesis and diagnosis of T2DM and its complications, GWA for glucose concentrations provides an excellent opportunity to identify genes underlying variation in glucose concentrations that may also represent additional T2DM susceptibility loci. An example of this comes from the studies by Weedon et al., who showed by metaanalysis and large cohorts that variation in the glucokinase gene is associated with both fasting glucose and birth weight (25).

GWA studies for T2DM and adiposity were completed by the groups undertaking the Finland–United States Investigation of Non–Insulin-Dependent Diabetes Mellitus Genetics (FUSION) (14, 26, 27) and the SardiNIA Study of Aging (24, 28), respectively. Both studies assessed fasting glucose in their respective cohorts, allowing GWAs for fasting glucose in each study and combination of these results in a metaanalysis. The strongest signals from the fasting glucose GWA metaanalysis were from variants near genes for ATP-binding cassette, subfamily B (MDR/TAP), member 11 (ABCB11) and glucose-6-phosphatase catalytic subunit 2 (G6PC2). This association was replicated in a series of 7 studies involving a total of 18,436 individuals (13, 29–35), suggesting for what we believe is the first time that variation in one of these genes may play a role in the regulation of fasting glucose concentrations in humans.

Results
Subject demographics and clinical characteristics for the FUSION and SardiNIA samples are summarized in Table 1. Because treatment for T2DM affects fasting glucose concentrations, all analyses in this report were restricted to nondiabetic subjects. Initial review of association results from both the FUSION stage 1 and SardiNIA GWA scans of a combined total of 5,088 nondiabetic individuals focused on SNPs that were genotyped in the SardiNIA study and imputed in the FUSION study. Among these, rs563694 exhibited the strongest evidence for association in both samples (SardiNIA, \( P = 7.6 \times 10^{-6} \); FUSION stage 1, \( P = 8.0 \times 10^{-4} \); Table 2), with a metaanalysis \( P \) value of \( 3.5 \times 10^{-2} \). Given the strength of this initial association, our follow-up efforts focused on rs563694. Additional independent associations from our fasting glucose GWA study are presented in Supplemental Table 1 (supplemental material available online with this article; doi:10.1172/JCI34566DS1).

Analyses were repeated once imputation was completed in both the FUSION stage 1 and SardiNIA samples. SNP rs563694 and other SNPs in strong linkage disequilibrium (LD; defined as \( r^2 > 0.8 \) in the FUSION samples) constituted the 17 strongest association results in the combined FUSION/SardiNIA GWA for fasting glucose metaanalysis (Figure 1). In fact, 22 SNPs associated with fasting plasma glucose with \( P \leq 1 \times 10^{-4} \) were located within a 63.9-kb region on chromosome 2 (Supplemental Table 2). These SNPs were located in an extended region of LD that spans 2 biologically plausible candidate genes for glucose regulation (Figure 1). The first is G6PC2, also known as islet-specific glucose-6-phosphatase-related protein (IGRP). G6PC2 is part of a larger family of enzymes involved in hydrolysis of glucose-6-phosphate in the gluconeogenic and glycolytic pathways (36, 37). The second is ABCB11, a member of the MDR/TAP subfamily of ATP-binding cassette transporters involved in multidrug resistance (38, 39).

Table 1
Subject demographics and clinical characteristics for individuals with rs563694 genotype data

<table>
<thead>
<tr>
<th>Study/Geographic origin</th>
<th>Phenotyped subjects</th>
<th>Geographic origin</th>
<th>Study age (years)</th>
<th>BMI (kg/m²)</th>
<th>Fasting glucose (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FUSION stage 1 Finland</td>
<td>1,233</td>
<td>Finland</td>
<td>63.0 (13.7)</td>
<td>26.6 (5.0)</td>
<td>5.36 (0.72)</td>
</tr>
<tr>
<td>FUSION stage 2 Finland</td>
<td>655</td>
<td>Finland</td>
<td>61.0 (12.3)</td>
<td>26.3 (4.9)</td>
<td>5.48 (0.50)</td>
</tr>
<tr>
<td>FUSION additional spouses/offspring</td>
<td>522</td>
<td>Finland, Italy</td>
<td>39.1 (12.2)</td>
<td>26.0 (6.4)</td>
<td>5.11 (0.78)</td>
</tr>
<tr>
<td>SardiNIA</td>
<td>3,855</td>
<td>Sardinia, Italy</td>
<td>41.3 (27.1)</td>
<td>24.7 (6.3)</td>
<td>4.72 (0.67)</td>
</tr>
<tr>
<td>DGI Finland and Sweden</td>
<td>1,411</td>
<td>Finland</td>
<td>58.7 (15.4)</td>
<td>26.7 (4.78)</td>
<td>5.28 (0.70)</td>
</tr>
<tr>
<td>Amish USA</td>
<td>1,655</td>
<td>USA</td>
<td>49.0 (23.7)</td>
<td>26.7 (6.6)</td>
<td>4.90 (0.58)</td>
</tr>
<tr>
<td>METSIM Finland</td>
<td>4,386</td>
<td>Finland</td>
<td>59.0 (10.0)</td>
<td>26.4 (4.5)</td>
<td>5.60 (0.70)</td>
</tr>
<tr>
<td>Caerphilly United Kingdom</td>
<td>1,063</td>
<td>United Kingdom</td>
<td>56.7 (4.4)</td>
<td>26.2 (3.5)</td>
<td>4.80 (0.86)</td>
</tr>
<tr>
<td>BWHHS United Kingdom</td>
<td>3,532</td>
<td>United Kingdom</td>
<td>68.5 (5.9)</td>
<td>27.6 (5.0)</td>
<td>5.80 (0.87)</td>
</tr>
<tr>
<td>Inter99 Denmark</td>
<td>5,734</td>
<td>Denmark</td>
<td>46.1 (7.9)</td>
<td>26.3 (4.5)</td>
<td>5.54 (0.80)</td>
</tr>
</tbody>
</table>

Values are reported as median (interquartile range).
Among all genotyped or imputed SNPs in this region, rs560887, which was genotyped in FUSION stage 1, and imputed and followed up by genotyping in SardiNIA, showed the strongest overall evidence for association (SardiNIA, $P = 4.4 \times 10^{-6}$; FUSION stage 1, $P = 1.7 \times 10^{-5}$; Supplemental Table 2), with a metaanalysis $P$ value of $2.8 \times 10^{-10}$. In addition, rs853789 and rs853787, both located in intron 19 of ABCB11 and in perfect LD with each other ($D^\prime = 1.0$, $r^2 = 1.0$), showed strong evidence for association with fasting glucose concentrations with metaanalysis $P$ values of $1.4 \times 10^{-9}$ and $1.0 \times 10^{-9}$, respectively (Supplemental Table 2). rs853789 is located 38.3 kb from rs560887 and 27.4 kb from rs563694 and is in strong LD with both SNPs ($D^\prime = 0.98$, $r^2 = 0.81$ with rs560887; and $D^\prime = 0.98$, $r^2 = 0.95$ with rs563694). rs560887 is 10.9 kb from rs563694, is in high LD with rs563694 ($D^\prime = 0.99$, $r^2 = 0.84$), and is located in intron 3 of G6PC2. In contrast, rs853789 lies between G6PC2 and ABCB11 and is in extended LD with ABCB11. In both the SardiNIA and FUSION stage 1 samples, each copy of the A allele for rs563694 was associated with small increases in fasting glucose (0.064 mM for SardiNIA and 0.051 mM for FUSION stage 1; Table 2) that are clinically insignificant and accounted for approximately 1% of the variance in fasting glucose. Similar effect sizes were observed for rs560887 (0.089 mM for SardiNIA and 0.052 mM for FUSION stage 1).

We assessed the potential contribution of population stratification by computing the genomic control parameter (40) independently for both studies. The genomic control values were 1.01 for both FUSION and SardiNIA, suggesting that population stratification and/or unmodeled relatedness did not contribute significantly to our observed association. Analyses that included BMI as a covariate did not significantly alter the association between rs563694 and fasting glucose in the FUSION stage 1 ($P = 8.1 \times 10^{-4}$ without BMI versus $9.1 \times 10^{-4}$ with BMI) and SardiNIA samples ($7.6 \times 10^{-5}$ without BMI versus $3.8 \times 10^{-5}$ with BMI) independently or jointly ($P = 3.5 \times 10^{-7}$ without BMI versus $1.8 \times 10^{-7}$ with BMI), suggesting the association was not a consequence of adiposity, which is known to induce insulin resistance and increase glucose concentrations (2). The association between rs563694 and fasting glucose also remained significant after individual adjustment for each of the 10 SNPs shown to be associated with T2DM in our recent GWA studies (Supplemental Table 3) (13–15) or when all 10 SNPs were included jointly in the model ($P = 6.8 \times 10^{-4}$ versus $P = 8.0 \times 10^{-4}$ for FUSION stage 1 samples and $5.1 \times 10^{-5}$ versus $7.6 \times 10^{-5}$ for SardiNIA samples). The 10 SNPs shown to be associated with T2DM were themselves not significantly associated with fasting glucose concentrations in the FUSION stage 1 or SardiNIA samples (Supplemental Table 4).

FUSION investigators genotyped rs563694 in 655 stage 2 samples and 522 additional spouses and offspring of T2DM patients included in stage 1; this SNP continued to show evidence for association with fasting glucose (FUSION stage 2, $P = 2.0 \times 10^{-5}$; FUSION stage 1 families, $P = 1.9 \times 10^{-5}$; see Table 2). The metaanalysis that combined results from the FUSION stage 1 and 2 and SardiNIA studies resulted in a $P$ value of $5.3 \times 10^{-9}$, surpassing standard thresholds for genome-wide significance.

We also examined the association between rs563694 and fasting glucose in 6 follow-up samples (Table 2). The characteristics of these samples are summarized in Table 1. Association between rs563694 and fasting glucose was confirmed in the Amish study ($P = 4.1 \times 10^{-6}$); the METabolic Syndrome In Men study (METSIM; $P = 1.3 \times 10^{-10}$), the Caerphilly study (2.6 $\times 10^{-7}$), the British Women’s Heart and Health Study (BWHH; $P = 1.2 \times 10^{-3}$), and Inter99 ($P = 8.2 \times 10^{-4}$; Table 2), with fasting glucose concentrations increasing with each copy of the A allele in all studies. While evidence for association in the Diabetes Genetics Initiative (DGI) study was not statistically significant ($P = 0.19$; Table 2), the results show a trend in the same direction as observed in the other samples. When the results from all follow-up studies were combined in a metaanalysis of 24,046 samples, there was strong evidence for association between rs563694 and fasting glucose in both the follow-up samples ($n = 18,435$, $P = 6.9 \times 10^{-26}$; Table 2) and in all GWA and follow-up samples combined ($n = 24,046$, $P = 6.4 \times 10^{-33}$; Table 2). In contrast, rs563694 did not show evidence for association with T2DM in the FUSION stage 1 study ($P = 0.22$), the DGI GWA samples ($P = 0.78$), or the METSIM study ($P = 0.09$).

### Table 2

Association between rs563694 and fasting glucose in nondiabetic individuals

<table>
<thead>
<tr>
<th>Study</th>
<th>Frequency</th>
<th>Mean fasting glucose (mM) (SD)</th>
<th>Effect (SE)</th>
<th>Effect (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
<td>C allele</td>
<td>CC</td>
<td>AC</td>
</tr>
<tr>
<td><strong>GWA Samples</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FUSION stage 1</td>
<td>1,233</td>
<td>0.34</td>
<td>5.26 (0.48)</td>
<td>5.31 (0.48)</td>
</tr>
<tr>
<td>SardiNIA</td>
<td>3,855</td>
<td>0.46</td>
<td>4.88 (0.67)</td>
<td>4.95 (0.62)</td>
</tr>
<tr>
<td><strong>FUSION 1 families</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1,755</td>
<td>0.34</td>
<td>5.20 (0.49)</td>
<td>5.28 (0.50)</td>
</tr>
<tr>
<td>Follow-up samples</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FUSION stage 2</td>
<td>655</td>
<td>0.36</td>
<td>5.28 (0.43)</td>
<td>5.44 (0.35)</td>
</tr>
<tr>
<td>DGI</td>
<td>1,411</td>
<td>0.34</td>
<td>5.24 (0.50)</td>
<td>5.28 (0.51)</td>
</tr>
<tr>
<td>Amish</td>
<td>1,655</td>
<td>0.24</td>
<td>4.90 (0.47)</td>
<td>4.89 (0.51)</td>
</tr>
<tr>
<td>METSIM</td>
<td>4,386</td>
<td>0.32</td>
<td>5.55 (0.49)</td>
<td>5.84 (0.50)</td>
</tr>
<tr>
<td>Caerphilly</td>
<td>1,063</td>
<td>0.36</td>
<td>4.69 (0.91)</td>
<td>4.87 (0.99)</td>
</tr>
<tr>
<td>BWHS</td>
<td>3,532</td>
<td>0.34</td>
<td>6.01 (1.69)</td>
<td>6.09 (1.81)</td>
</tr>
<tr>
<td>Inter99</td>
<td>5,734</td>
<td>0.36</td>
<td>5.46 (0.85)</td>
<td>5.52 (0.87)</td>
</tr>
</tbody>
</table>

Glucose values are reported unadjusted for covariates. <sup>a</sup>GWA metaanalysis $P$ value. <sup>b</sup>These data represent the 522 additional spouses and offspring combined with their FUSION stage 1 family members. <sup>c</sup>Follow-up metaanalysis $P$ value. <sup>d</sup>Overall metaanalysis $P$ value.
Figure 2 shows the results of a metaanalysis based upon the effect size observed in each of the 8 studies. Overall, fasting glucose concentrations increased 0.065 mM (95% CI: 0.053–0.077 mM) with each copy of the major allele.

**Discussion**

We took advantage of GWA studies originally performed to identify susceptibility genes for T2DM (FUSION) and aging-related traits (SardiNIA) to also identify genes underlying variation in fasting glucose concentration. Both FUSION and SardiNIA initially identified rs563694 as being associated with fasting glucose levels. Given that both studies were performed in relatively homogeneous populations of mixed European descent, it is unlikely that population stratification accounted for the initial association. The estimated genomic control (40) values for FUSION stage 1 and SardiNIA were both 1.01, providing further evidence against the contribution of population stratification to the observed association.

In the SardiNIA sample, we genotyped SNPs rs560887 and rs853789 to validate the results based on imputation. The discrepancy rate per allele between the imputed and typed genotypes at these 2 SNPs was 1.4% and 2.4%, respectively, and the association result with the actual genotypes was stronger than with the imputed genotypes: \( \log(P) = 9.0 \times 10^{-10} \) and \( 2.6 \times 10^{-8} \), respectively.

Adiposity may induce insulin resistance and thus alter glucose concentrations (2) independent of the effects of the SNP on glucose concentrations per se. However, the association remained significant even when we included BMI as a covariate in the analysis, suggesting adiposity is not a major contributor to the observed association. Similarly, in the follow-up studies, the results did not change whether BMI was included or excluded as a covariate. Some known sex-specific effects, such as differences in fat distribution, could also confound our results. We found no sex-specific effect modification in the FUSION and SardiNIA samples. Also, it should be noted that we observed evidence for association between rs563694 in the METSIM and Caerphilly samples that only included men and in the BWHHS that comprised women only. Thus, the lack of a sex-specific effect in FUSION and SardiNIA is supported by the independent associations observed in these samples. Subsequent analyses of the GWA data revealed rs560887 as having the strongest evidence for association with fasting glucose in this region and suggested, based on the SNP location, that G6PC2 plays a role in glucoregulation. However, 2 additional SNPs in strong LD with rs560887 located in the adjacent ABCB11 also showed similar evidence for association with fasting glucose.

In the 7 follow-up studies, rs563694 continued to show association with fasting glucose, although marginal evidence for heterogeneity among studies was noted (\( Q = 14.6; \ P = 0.02; P = 59.0\% \); 95% CI: 5.6–82.2%) (37). For example, the DGI samples did not exhibit a significant association and the BWHHS samples, despite being among the largest follow-up samples, showed only modest evidence for association (Table 2). These 2 studies yielded similar effect size estimates (0.053 for DGI and 0.079 for BWHHS) that were smaller than in the other studies (Table 2). Differences in effect size, the direction of the effect was the same in all studies.

There are 2 biologically plausible candidate genes in the region identified by our association analyses that may affect glucose levels. Although rs560887, which is located in intron 3 of G6PC2 just 26 bp proximal to exon 4, showed the strongest evidence for association in the GWA studies, SNPs in LD with rs560887 and rs563694 that show similar levels of association with fasting glucose concentrations were located in intron 19 of ABCB11. ABCB11 is involved in ATP-dependent secretion of bile salts and is almost exclusively expressed in the liver. Mutations in ABCB11 have been shown to be associated with intrahepatic cholestasis (OMIM 603201) (38) and drug-induced hepatotoxicity (39, 41). In antilipid drug trials, bile acid sequestrants have been shown to lower glucose concentrations and improve insulin sensitivity, presumably through reduction of triglyceride levels (42). Based on these observations, if ABCB11 were contributing significantly to variation in fasting glucose, one might expect to also see associations...
with lipids or insulin sensitivity. However, rs560887, rs563694, rs853789, and rs853787 were not associated with lipid measurements in a metaanalysis of FUSION stage 1 and SardiNIA samples ($P > 0.16$). Also, none of these SNPs were associated with minimal model-derived insulin sensitivity in FUSION samples ($P > 0.30$). Thus, our data do not support a role for ABCB11 in glucoregulation, and other evidence directly linking ABCB11 to regulation of glucose concentrations is scarce.

In contrast, G6PC2, the β cell–specific isoform of glucose-6-phosphatase is a highly relevant candidate gene for glucoregulation. The mouse homolog G6pc2 has been previously implicated as an autoantigen in the NOD mouse model of type 1 diabetes (43). Wang et al. recently generated G6pc2-null mice and noted that at 16 weeks of age, fasting glucose concentrations had decreased approximately 13% in both male and female G6pc2-null mice when compared with wild-type mice (44). This modest decrease in glucose concentration was observed despite the absence of any differences in body weight, fasting insulin, or fasting glucagon concentrations. The characteristics of these G6pc2-null mice closely paralleled our observations that rs560887 and rs563694 were associated with modest changes in fasting glucose but not in BMI or fasting insulin, which are consistent with the hypothesis that presence of a C allele results in lower G6PC2 expression and therefore lower glucose concentrations. Interestingly, G6pc2 mRNA levels appear to increase with increasing glucose concentration in isolated mouse islets (36).

Molecular cloning of G6pc2 identified 2 splice forms that differ by the presence or absence of exon 4 in BALB/C and ob/ob mice and in insulinoma tissue (45). The longer cDNA including exon 4 has approximately 50% homology with glucose-6-phosphatase catalytic subunit (G6pc) across a variety of species including humans and is membrane bound in the endoplasmic reticulum (46). The corresponding G6PC2 splice forms have been observed in human pancreas (47). rs560887 is located in intron 3, just 26 bp proximal to exon 4, raising the possibility that this variant may play a role in whether the full-length transcript is formed.

G6PC hydrolyzes glucose-6-phosphate to form glucose and release a phosphate group. Despite its similarity to G6PC, G6PC2 is reported to have little to no hydrolase activity in humans (36, 37, 45, 46). In normal and genetically obese mice, the splice form lacking exon 4 appears to be the most predominant observed in islets (45) and lacks sequences that may be critical for hydrolytic activity (45, 48), suggesting the full-length form of G6pc2 may have implications for activity of G6pc2 and its potential role in glucoregulation. Greater hydrolase activity has been reported in cell lines over-expressing the full-length form of G6PC2 (36). Also, in islets from streptozotocin-treated mice, glucose cycling, an indicator of G6pc2 activity, was approximately 3-fold higher compared with islets from untreated mice (49), and even greater increases were observed in islets from ob/ob mice (50, 51). The conversion of glucose to glucose-6-phosphate is the critical step in stimulus-secretion coupling for insulin secretion. Variation in G6PC2 may increase glucose cycling in β cells, resulting in altered generation of ATP, which would have implications for insulin secretion. In addition, G6PC2-induced alterations in β cell glucose metabolism would also have downstream effects on phosphoinositide 3-kinase activity, which regulates pancreas duodenum homeobox-1 (PDX1) binding to the insulin gene and subsequent insulin gene transcription (52).

The possible role for G6PC2 in altering glucose concentrations raises the question of whether this gene also confers susceptibility to T2DM. We observed no association between fasting glucose and rs563694 and rs560887 in individuals with T2DM from the FUSION, DGI, and METSIM studies ($P > 0.50$). However, the analysis of fasting glucose concentration in individuals with T2DM is confounded by diabetes pathology, treatment, and differential response to therapy. Therefore the lack of association with fasting glucose in individuals with T2DM does not preclude G6PC2 as contributing to susceptibility to T2DM. Similarly, when we tested these SNPs for association with T2DM in the FUSION, DGI, and METSIM samples, we observed no evidence for association ($P > 0.08$). Further, the modest effect on glucose concentrations observed in our analysis of nondiabetic individuals suggests we may lack sufficient power to detect association with T2DM. Whereas the cumulative evidence would suggest that G6PC2 may regulate fasting glucose concentrations and does not contribute significantly to susceptibility to T2DM, larger studies may be required to elucidate the role of this gene in T2DM susceptibility.

Variation in the promoter region of glucokinase (GCK, rs1799884) has been shown to be associated with fasting glucose and impaired insulin secretion (53–55) and may play a role in altering birth weight (36). These initial findings were confirmed in a comprehensive metaanalysis performed by Weedon et al., demonstrating that rs1799884 was associated with fasting glucose (meta $P = 1.0 \times 10^{-5}$) and that the presence of a maternal A allele for rs1799884 was associated with increased birth weight of the child ($P = 0.02$) (25). GCK, an enzyme that works counter to G6PC2, converts glucose to glucose-6-phosphate, forming the critical step in secretion-stimulus coupling in pancreatic β cells. In addition, the recent GWA study from the DGI identified variation in glucokinase regulatory protein (GCKR) (rs780094) to be associated with triglyceride levels (13). GCKR is an allosteric regulator of GCK in both liver and pancreatic islets whose inhibitory effect is enhanced by fructose-6-phosphate and suppressed by fructose-1-phosphate (57). We found modest evidence...
for association between fasting glucose and rs1799884 (FUSION stage 1, \( P = 1.6 \times 10^{-5} \); SardiNIA, \( P = 2.0 \times 10^{-3} \); meta \( P = 1.1 \times 10^{-4} \)) and no evidence for association between fasting glucose and rs780094 (FUSION stage 1, \( P = 0.44 \); SardiNIA, \( P = 0.11 \); meta \( P = 0.077 \)). While these results provide evidence for association between variation in GCK and fasting glucose but not between GCKR and fasting glucose in our studies, we cannot exclude the possibility that a complex interaction among GCK, GCKR, and G6PC2 may regulate fasting glucose levels. This will require further study.

In conclusion, we used GWA to identify variation in both ABCB11 and G6PC2 as genes that potentially contribute to variation in fasting glucose concentrations in nondiabetic subjects of mixed European descent. There is more literature with data supporting a role for G6PC2, but in the absence of functional data, we cannot discount the possibility that ABCB11 may also contribute significantly to variation in fasting glucose concentration. Heritability for fasting glucose has been estimated to be 25%–40% (19–24), yet the variants we identified account for approximately 1% of the variance in fasting glucose, indicating that the majority of the variability in fasting glucose remains unexplained. The remaining variability is likely due to the effects of additional common genetic variants of modest effect, less common genetic variants of moderate effect, and a variety of gene-gene and gene-environmental interaction effects. It should also be noted that the magnitude of the effect observed in our study is consistent with other reports of quantitative trait associations (58–60).

Additional studies, likely with larger sample sizes, will be required to identify additional genetic variants contributing to variation in fasting glucose. The variants identified in our study are not likely to be functional, but in LD with the functional variant(s). Additional fine mapping, sequencing, and functional studies will be required to define the molecular mechanisms underlying our observed association.

**Methods**

The FUSION and SardiNIA study samples and GWA genotyping have been described in detail (14, 24, 26–28). Here, we briefly review the study cohorts and genotyping methods. We also describe briefly each of the 7 follow-up samples. Subject demographics and basic clinical characteristics for individuals genotyped for rs563694 for each sample are described below and summarized in Table 1. All protocols were approved by the institutional review boards or research ethics committees at the respective institutions, and informed consent was obtained from all subjects.

**FUSION GWA study.** The goal of the FUSION study is to identify genetic variants that predispose to T2DM or that determine the variability in T2DM-related quantitative traits. The study began as an affected sibling-pair family study (26, 27), later augmented by large numbers of cases and controls for association analysis (14). The FUSION GWA study was performed using a 2-stage case-control design (14). Cases and controls were approximately frequency matched on 5-year age category, sex, and birth province. All stage 1 DNA samples were genotyped using the Illumina HumanHap300 BeadChip version 1.0, resulting in data on 315,635 SNPs that passed quality control filters (14). Genotype data for an additional 2.09 million SNPs were estimated using an imputation procedure (61). The genotype imputation method uses stretches of chromosome shared between individuals genotyped at relatively low density in our studies and individuals genotyped in greater density by the International HapMap Consortium (61) to estimate the missing genotypes. Comparison of imputed and measured genotypes yielded estimated error rates of 1.46% (Illumina) to 2.14% (Affymetrix) per allele with an average concordance of 98.5%, consistent with expectations from HapMap data (61). SNPs showing promising association with fasting plasma glucose in the stage 1 samples were genotyped in the stage 2 DNA samples by homogeneous MassEXTEND reaction using the MassARRAY System (Sequenom) (14). Because treatment for T2DM affects fasting glucose concentrations, all analyses in this report were restricted to nondiabetic subjects. Diabetes status was confirmed by WHO criteria (62) or confirmation of treatment for diabetes by medical record review. Fasting plasma glucose concentrations were available for 1,233 stage 1 and 655 stage 2 samples. Additional FUSION stage 1 samples included nondiabetic spouses or offspring from FUSION stage 1 families; fasting plasma glucose data were available for 578 individuals. These 578 samples were genotyped using the Applied Biosystems TaqMan allelic discrimination assays (63) and yielded 522 samples with both genotype and fasting glucose data. These samples were integrated into the FUSION stage 1 samples and independently analyzed to assess whether the additional family members improved the evidence for association. We have denoted this analysis FUSION 1 families.

SardiNIA GWA study. The SardiNIA study is a longitudinal study of aging-related quantitative traits and comprises a cohort of 6,148 individuals 14 years or older recruited from 4 towns in the Lanusei Valley in Sardinia. Data from 4,350 individuals with fasting serum glucose measurements from this cohort were used for the GWA study; 3,331 were genotyped using the Affymetrix 10K SNP Mapping Array, and an additional 1,412 were genotyped using the Affymetrix 500K SNP Mapping Array (28). 356,359 SNPs passed quality control and were tested for association with fasting serum glucose. We first used the genotyped SNPs in the 1,412 individuals to estimate genotypes for all the polymorphic SNPs genotyped by the HapMap Consortium. Taking advantage of the relatedness among individuals in the SardiNIA sample, we then conducted a second round of computational analysis to impute genotypes for analysis in the 2,938 individuals not genotyped with the 500K SNP Array. In this second round, we identified large stretches of chromosome shared within each family and probabilistically “filled-in” genotypes within each stretch whenever 1 or more of its carriers was genotyped with the 500K Array Set (64, 65). For these analyses, 37 non-Sardinians and 281 of their family members (n = 318) and 177 individuals with known diabetes were excluded from the analysis, resulting in a final sample size of 3,855.

**Follow-up samples.** The initial association identified in the metaanalysis of the FUSION and SardiNIA GWA studies was also tested in a series of follow-up samples (Table 1). 1 from FUSION described above and 6 others, which are described briefly below.

**DGI.** The DGI case-control GWA sample consists of 1,464 cases with T2DM and 1,467 normoglycemic controls from Finland and Sweden and has been previously described in detail (13). Fasting glucose measurements were available for 1,455 nondiabetic control subjects (1,305 unrelated subjects and 150 siblings). Among these, fasting plasma glucose was measured in 537 subjects and fasting whole blood glucose was measured in 918 subjects. Whole-blood glucose concentrations were converted to equivalent plasma values using a conversion factor of 1.13 (66). All samples were genotyped using the Affymetrix GeneChip Human Mapping 500K Array set; results of GWA of 389,878 SNPs with fasting glucose levels (including SNP rs563694) are publicly available at www.broad.mit.edu/diabetes/scandinavs/index.html. 1,411 individuals were available with both rs563694 genotype and fasting glucose data.

**Old Order Amish subjects.** The Old Order Amish study participants were included here were 1,655 nondiabetic subjects from Lancaster, Pennsylvania, USA, for whom fasting plasma glucose measurements were available. These subjects were enrolled in ongoing family studies of complex diseases and traits (29–31). Genotyping for rs563694 was performed using the TaqMan allelic discrimination assay (63).
Subjects were selected from the ongoing METSIM study, which includes 7,000 men, aged 50 to 70 years, randomly selected from the population of the town of Kuopio, Eastern Finland, Finland (population 95,000). The present analysis is based on the first 4,386 non-diabetic subjects examined for METSIM with available fasting plasma glucose values. Genotyping was performed using the TaqMan allelic discrimination assay (63).

Caerphilly study. The Caerphilly study is a cohort study of white, European men (n = 1,069; 97.4% born in the United Kingdom), aged 45–59 years at entry in 1979–1983 (32), recruited from the town of Caerphilly, United Kingdom, and 5 adjacent villages. Men were selected using the electoral roll and general practitioner records. DNA and fasting plasma glucose measurements were used in this study to relate to the first phase of data collection.

BWHHS. The BWHHS consists of female participants, aged 60 to 70 years and recruited between April 1999 and March 2001. Initially, 4,286 women were randomly selected from 23 British towns and were interviewed and clinically examined. They also completed medical questionnaires (33).

Genotyping for the Caerphilly study and BWHHS was performed by KBioscience using their fluorescence-based competitive allele-specific PCR (KASPar) technology.

The Inter99 Study. rs563694 was genotyped in 5,734 Danes for whom fasting plasma glucose values were available. This sample comprises part of the population-based Inter99 sample of middle-aged people sampled at Research Centre for Prevention and Health (Glostrup, Denmark; refs. 34, 35). Genotyping was performed using TaqMan allelic discrimination (KBioscience).

Statistics. Association between fasting glucose and genotypes in the FUSION and SardiNIA studies was carried out using a regression framework in which regression coefficients were estimated in the context of a variance component model to account for relatedness among individuals (65). For FUSION samples, plasma glucose concentration was adjusted for sex, age, and age2, birth province, and study group. Analyses were carried out in non-diabetic individuals excluding those known to be taking medications that directly affect glucose concentration. Similarly, SardiNIA serum glucose values were adjusted for sex, age, and age2. Because diabetes-based exclusions were based only on medical records and SardiNIA only measured fasting serum glucose, a small number of undiagnosed new-onset diabetes cases may have been included in the analysis. For both studies, analyses were repeated including BMI as an additional covariate to assess whether adiposity significantly contributed to the evidence for association. Covariate-adjusted trait values were transformed to approximate univariate normality by applying an inverse normal scores transformation; the scores were ranked, ranks were transformed into quantiles, and quantiles were converted to normal deviates.

A weighted z score–based fixed effects metaanalysis method was used to combine results from the FUSION and SardiNIA studies. In brief, for each SNP, a reference allele was identified and a z statistic summarizing the magnitude of the P value for association and direction of effect was generated for each study. An overall z statistic was then computed as a weighted average of the individual statistics, and a corresponding P value for that statistic was computed. The weights were proportional to the square root of the number of individuals in each study and scaled such that the squared weights summed to 1. For the metaanalysis of the effect size, the inverse variance was used as weights for each study. For the FUSION 1 families (FUSION stage 1 plus additional FUSION spouses and offspring) a regression-based analysis under a variance components framework was used to appropriately account for relationships among individuals (65). Because we did not have birth province information for the additional spouses and offspring, these analyses were carried out adjusting for age, age2, sex, and study group only.

Given the different sampling schemes, statistical analyses for the follow-up samples varied by study. The Old Order Amish samples consisted of large Amish pedigrees, so the evidence for association between genotype and fasting plasma glucose was evaluated using variance components analysis implemented in SOLAR to adjust for the relatedness of study subjects (67, 68). Plasma glucose levels were natural logarithm transformed for analysis, and covariates included sex, age, and age2. For the DGI study, glucose values were converted to z scores separately by sex, and tests for association were carried out using a regression framework with age and log(BMI) included as covariates; genomic control was applied to account for relatedness (13). For the METSIM study, analyses were carried out identically as in FUSION, with the exception that birth province was not included as a covariate. For the Caerphilly and BWHHS studies, association was assessed using a regression framework with age and sex as covariates. Individuals with known diabetes at the time of examination were excluded from the analyses. Results from all follow-up studies were combined in a metaanalysis as described above. Finally, a metaanalysis that combined results from all GWA and follow-up studies was performed as described above.

Acknowledgments

We would like to thank the many research volunteers who generously participated in the various studies represented in this study. For the FUSION study, we also thank Peter S. Chines, Narisu Narisu, Andrew G. Sprau, and Li Qin for informatics and genotyping support and the Center for Inherited Disease Research for the FUSION GWA genotyping. For the SardiNIA study, we thank the mayors of Lanusei, Ilbono, Arzana, and Elini, the head of local Public Health Unit ASL4, and the residents of the towns for their volunteerism and cooperation. In addition, we are grateful to the mayor and the administration in Lanusei for providing and furnishing the clinic site. We thank the team of physicians — Maria Grazia Pilia, Danilo Fois, Liana Ferrelli, Marcello Argiolas, Francesco Loi, and Pietro Figus — and the nurses Paola Loi, Monica Lai, and Anna Cau, who carried out the physical examinations and made the observations.

We thank the former Medical Research Council (MRC) Epidemiology Unit (South Wales) who undertook the Caerphilly study. The Department of Social Medicine, University of Bristol, now acts as custodian for the Caerphilly database. We are grateful to all of the men who participated in this study. For the BWHHS, we thank all of the general practitioners and their staff who supported data collection and the women who participated in the study.

For the Amish studies, we thank members of the Amish community for the generous donation of time to participate in these studies and our field nurses, Amish liaisons, and clinic staff for their extraordinary efforts. We also acknowledge Sandy Ott and John Shelton for genotyping of Amish DNA samples.

Support for this study was provided by the following: American Diabetes Association (ADA) (1-05-RA-140 to R.M. Watanabe; 7-04-RA-111 to A.R. Shuldiner; and postdoctoral fellowships to C.J. Willer and H.M. Stringham); and NIH grants (DK069922 and U54 DA021519 to R.M. Watanabe; DK062370 to M. Boehnke; DK072193 to K.L. Mohlke; DK062418 to W-M. Chen; R01 DK54361, U01 HL72515, and R01 AG18728 to A.R. Shuldiner; R01 HL69313 to B.D. Mitchell; and R01 DK068495 to K.D. Silver). D.A. Lawlor is funded by a UK Department of Health career scientist award, and N. Timpson is funded by a studentship from the MRC of the United Kingdom. The Inter99 Study was supported by the European Union (EUGENE2, LSHM-CT-2004-512013); the Lundbeck Foundation Centre of Applied Medical Genomics in Personalized Disease Prediction, Prevention and Care; the FOOD Study Group/
the Danish Ministry of Food, Agriculture and Fisheries and the Ministry of Family and Consumer Affairs (2101-05-0044); and the Danish Medical Research Council.

This research was supported in part by the intramural Research Program of the NIH, National Institute on Aging, and the NIDDK. Additional support came from contract N01-AG-1-2109 from the NIA intramural research program for the SardiNIA (ProgeNIA) team; National Human Genome Research Institute intramural project number I Z01 HG000024 (to F.S. Collins); University of Maryland General Clinical Research Center (M01 RR 16500); Johns Hopkins University General Clinical Research Center (M01 RR 00052); the NIDDK Clinical Nutrition Research Unit of Maryland (P30 DK072488); and the Department of Veterans Affairs and Veterans Affairs Medical Center Baltimore Geriatric Research, Education and Clinical Center (GRECC). The BWHHS receives core funding from the British Heart Foundation. The Caerphilly study was funded by the MRC of the United Kingdom. Funding for the Caerphilly DNA Bank was from an MRC grant (G9824960). The United Kingdom MRC supports work undertaken in the Centre for Causal Analyses in Translational Epidemiology.

The views expressed in this paper are those of the authors and not necessarily those of any funding body or others whose support is acknowledged. Those providing funding had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Received for publication November 26, 2007, and accepted in revised form April 23, 2008.

Address correspondence to: Angelo Scuteri, Unità Operativa Geriatria, Istituto Nazionale Ricoavo E. Cura Anziani, Rome, Italy. Phone: 39-33334564136; Fax: 39-06-30362896; E-mail: angelo.olefante@interfree.it. Or to: Richard M. Watanabe, Keck School of Medicine of USC, Department of Preventive Medicine, 1540 Alcator St., CHP-220, Los Angeles, California 90089-9011, USA. Phone: (323) 422-2053; Fax: (323) 442-2349; E-mail: rwatanab@usc.edu.

Wei-Min Chen and Michael R. Erdos are co-first authors.


