



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

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

<sup>1</sup>Department of Public Health Sciences and <sup>2</sup>Center for Public Health Genomics, University of Virginia, Charlottesville, Virginia, USA.

<sup>3</sup>Genome Technology Branch, National Human Genome Research Institute, Bethesda, Maryland, USA.

<sup>4</sup>Center for Statistical Genetics and Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, USA.

<sup>5</sup>Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA.

<sup>6</sup>Istituto di Neurogenetica e Neurofarmacologia, Consiglio Nazionale delle Ricerche, Cagliari, Italy. <sup>7</sup>Division of Endocrinology, Diabetes and Nutrition, University of Maryland School of Medicine, Baltimore, Maryland, USA. <sup>8</sup>MRC Centre for Causal Analyses in Translational Epidemiology, Department of Social Medicine, University of Bristol, Bristol, United Kingdom. <sup>9</sup>Steno Diabetes Center, Gentofte, Denmark.

<sup>10</sup>Department of Clinical Sciences, Diabetes and Endocrinology, Lund University, University Hospital Malmö, Malmö, Sweden.

<sup>11</sup>Department of Medicine, University of Kuopio and Kuopio University Hospital, Kuopio, Finland. <sup>12</sup>Department of Epidemiology and Population Health, Non-communicable Disease Epidemiology Unit, London School of Hygiene and Tropical Medicine, University of London, London, United Kingdom.

<sup>13</sup>Gerontology Research Center, National Institute on Aging, Baltimore, Maryland, USA. <sup>14</sup>Social Medicine Department, University of Bristol, Bristol, United Kingdom. <sup>15</sup>Research Centre for Prevention and Health, Glostrup University Hospital, Glostrup, Denmark. <sup>16</sup>Faculty of Health Sciences, University of Aarhus, Aarhus, Denmark. <sup>17</sup>Savitaipale Health Center, Savitaipale, Finland. <sup>18</sup>Diabetes Unit,

Department of Health Promotion and Chronic Disease Prevention, National Public Health Institute, and Department of Public Health, University of Helsinki, Helsinki, Finland. <sup>19</sup>Department of Medicine, Division of Endocrinology, and <sup>20</sup>Department of Physiology and Biophysics, Keck School of Medicine, University of Southern California, Los Angeles, California, USA. <sup>21</sup>South Ostrobothnia Central Hospital, Seinäjoki, Finland. <sup>22</sup>Department of Genetics, University of North Carolina, Chapel Hill, North Carolina, USA. <sup>23</sup>Laboratory of Cardiovascular Science, National Institute on Aging, NIH, Baltimore, Maryland, USA. <sup>24</sup>Unità Operativa Geriatria, Istituto Nazionale Ricovero E Cura Anziari, Rome, Italy.

<sup>25</sup>Department 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 ( $P = 3.5 \times 10^{-7}$ ). 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  $6.9 \times 10^{-26}$ , and combining results from all studies resulted in an overall  $P$  value for association of  $6.4 \times 10^{-33}$ . 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.**

**Nonstandard abbreviations used:** *ABCB11*, ATP-binding cassette, subfamily B (*MDR/TAP*), member 11; BWHHS, British Women's Heart and Health Study; DGI, Diabetes Genetics Initiative; FUSION, Finland–United States Investigation of Non-Insulin-Dependent Diabetes Mellitus Genetics; *G6PC2*, glucose-6-phosphatase catalytic subunit 2; GWA, genome-wide association; LD, linkage disequilibrium; METSIM, METabolic Syndrome in Men; T2DM, type 2 diabetes mellitus.

**Conflict of interest:** The authors have declared that no conflict of interest exists.

**Citation for this article:** *J. Clin. Invest.* 118:2620–2628 (2008). doi:10.1172/JCI34566.

## 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  $\beta$  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 was 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,  $P = 7.6 \times 10^{-5}$ ; FUSION stage 1,  $P = 8.0 \times 10^{-4}$ ; Table 2), with a metaanalysis  $P$  value of  $3.5 \times 10^{-7}$ . 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 glucoregulation (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 glycogenolytic 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

Study	Phenotyped subjects	Geographic origin	Study age (years)	BMI (kg/m <sup>2</sup> )	Fasting glucose (mM)
FUSION stage 1	1,233	Finland	63.0 (13.7)	26.6 (5.0)	5.36 (0.72)
FUSION stage 2	655	Finland	61.0 (12.3)	26.3 (4.9)	5.48 (0.50)
FUSION additional spouses/offspring	522	Finland	39.1 (12.2)	26.0 (6.4)	5.11 (0.78)
SardinIA	3,855	Sardinia, Italy	41.3 (27.1)	24.7 (6.3)	4.72 (0.77)
DGI	1,411	Finland and Sweden	58.7 (15.4)	26.7 (4.78)	5.28 (0.70)
Amish	1,655	USA	49.0 (23.7)	26.7 (6.6)	4.90 (0.58)
METSIM	4,386	Finland	59.0 (10.0)	26.4 (4.5)	5.60 (0.70)
Caerphilly	1,063	United Kingdom	56.7 (4.4)	26.2 (3.5)	4.80 (0.86)
BWHHS	3,532	United Kingdom	68.5 (5.9)	27.6 (5.0)	5.80 (0.87)
Inter99	5,734	Denmark	46.1 (7.9)	26.3 (4.5)	5.54 (0.80)

Values are reported as median (interquartile range).



**Table 2**  
Association between rs563694 and fasting glucose in nondiabetic individuals

Study	n	Frequency	Mean fasting glucose (mM) (SD)			Effect (SE)	Effect (SE)	P value
		C allele	CC	AC	AA	mM	Standardized	
GWA Samples								
FUSION stage 1	1,233	0.34	5.26 (0.48)	5.31 (0.48)	5.33 (0.47)	0.051 (0.019)	0.143 (0.043)	$8.0 \times 10^{-4}$
SardinIA	3,855	0.46	4.88 (0.67)	4.95 (0.62)	5.00 (0.59)	0.064 (0.018)	0.118 (0.030)	$7.6 \times 10^{-5}$
								$3.5 \times 10^{-7A}$
FUSION 1 families <sup>B</sup>	1,755	0.34	5.20 (0.49)	5.28 (0.50)	5.31 (0.54)	0.065 (0.018)	0.155 (0.036)	$1.9 \times 10^{-5}$
Follow-up samples								
FUSION stage 2	655	0.36	5.28 (0.43)	5.44 (0.35)	5.46 (0.36)	0.068 (0.021)	0.180 (0.058)	$2.0 \times 10^{-3}$
DGI	1,411	0.34	5.24 (0.50)	5.28 (0.51)	5.29 (0.49)	0.022 (0.021)	0.053 (0.039)	0.19
Amish	1,655	0.24	4.90 (0.47)	4.89 (0.51)	5.03 (0.53)	0.090 (0.022)	0.175 (0.042)	$4.1 \times 10^{-5}$
METSIM	4,386	0.32	5.55 (0.49)	5.64 (0.50)	5.71 (0.49)	0.074 (0.011)	0.145 (0.023)	$1.3 \times 10^{-10}$
Caerphilly	1,063	0.36	4.69 (0.91)	4.87 (0.99)	5.00 (1.19)	0.155 (0.047)	0.214 (0.041)	$2.6 \times 10^{-7}$
BWHHS	3,532	0.34	6.01 (1.69)	6.09 (1.81)	6.06 (1.49)	0.006 (0.042)	0.079 (0.025)	$1.2 \times 10^{-3}$
Inter99	5,734	0.36	5.46 (0.85)	5.52 (0.87)	5.58 (0.70)	0.057 (0.015)	0.135 (0.019)	$8.2 \times 10^{-8}$
								$6.3 \times 10^{-28C}$
								$6.1 \times 10^{-35D}$

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.

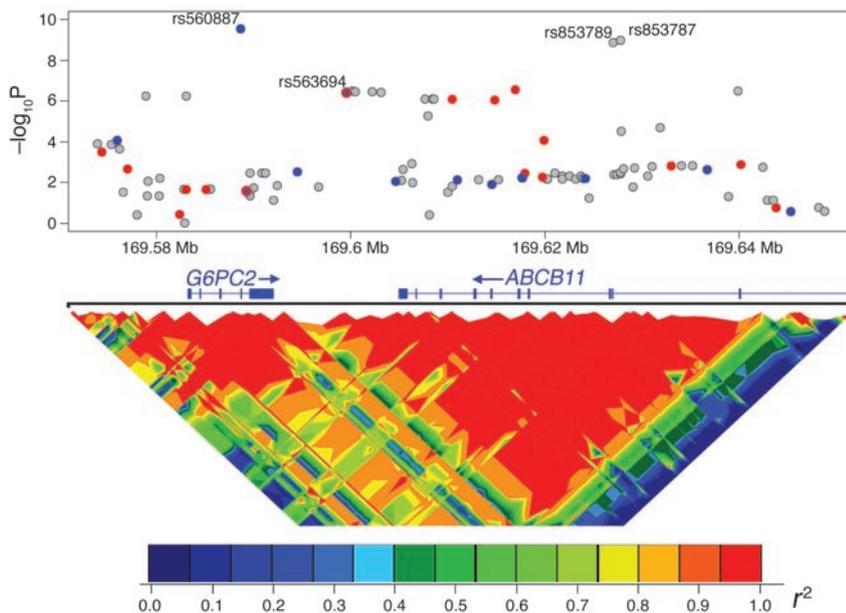
Among all genotyped or imputed SNPs in this region, rs560887, which was genotyped in FUSION stage 1, and imputed and followed up by genotyping in SardinIA, showed the strongest overall evidence for association (SardinIA,  $P = 4.4 \times 10^{-8}$ ; FUSION stage 1,  $P = 1.7 \times 10^{-3}$ ; Supplemental 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 ( $D' = 1.0$ ,  $r^2 = 1.0$ ), 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 ( $D' = 0.98$ ,  $r^2 = 0.81$  with rs560887; and  $D' = 0.98$ ,  $r^2 = 0.95$  with rs563694). rs560887 is 10.9 kb from rs563694, is in high LD with rs563694 ( $D' = 0.99$ ,  $r^2 = 0.84$ ), and is located in intron 3 of *G6PC2*. In contrast, rs563694 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,  $P = 2.0 \times 10^{-3}$ ; FUSION stage 1 families,  $P = 1.9 \times 10^{-5}$ ; see Table 2). The meta-analysis 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 ( $P = 4.1 \times 10^{-5}$ ); 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 (BWHHS;  $P = 1.2 \times 10^{-3}$ ), and Inter99 ( $P = 8.2 \times 10^{-8}$ ; 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 meta-analysis 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 ( $P = 0.09$ ).

**Figure 1**

Fasting glucose association in the FUSION and SardiNIA GWA metaanalysis. Top panel shows evidence for association with fasting glucose under an additive genetic model for the combined FUSION stage 1 and SardiNIA metaanalysis. The  $-\log(P)$  value for the test of association is plotted against genomic position (NCBI build 35) for all genotyped (red circles, SardiNIA; blue circles, FUSION) and imputed (gray circles) SNPs in the SardiNIA data. SNPs typed in both samples are indicated by red circles with blue centers. SNPs rs560887 and rs853789 were later typed in the SardiNIA samples, and the actual genotypes resulted in even stronger association than shown here. Bottom panel shows the LD pattern ( $r^2$ ) around *G6PC2* and *ABCB11* for the CEPH (Utah residents with ancestry from northern and western Europe) sample from the HapMap. The scale at the bottom shows the magnitude of LD in 15 colors, ranging from blue for low LD to red for high LD.

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.

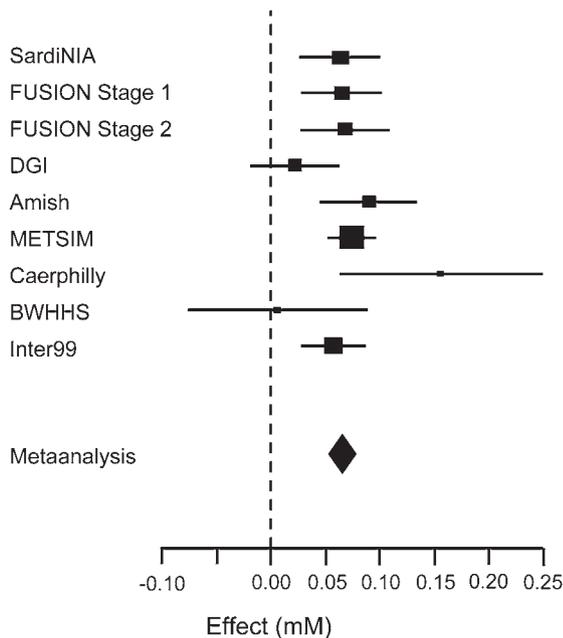
In the SardiNIA sample, we genotyped SNPs rs560887 and rs853789 to validate the results based on imputation. The discrepancy rate per allele between the imputed and typed genotypes at these 2 SNPs was 1.4% and 2.4%, respectively, and the association result with the actual genotypes was stronger than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

ported 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 ( $Q = 14.6$ ;  $P = 0.02$ ;  $I^2 = 59.0\%$ ; 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 both populations and sample ascertainment could be contributing to the observed heterogeneity. When these 2 studies are not considered, the heterogeneity estimate is reduced ( $Q = 3.7$ ;  $P = 0.45$ ;  $I^2 = 0\%$ ; 95% CI = 0–77.6%) However, despite the variability 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 anti-lipid drug trials, bile acid sequestrants have been shown to lower glucose concentrations and improve insulin sensitivity, presumably through reduction of triglyceride levels (42). Based upon these observations, if *ABCB11* were contributing significantly to variation in fasting glucose, one might expect to also see associations

**Figure 2**

Effect size and 95% CI for rs563694 are shown for the 8 studies. The overall metaanalysis across these studies yielded an effect size of 0.065 mM (95% CI: 0.053, 0.077 mM).

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  $\beta$  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 overexpressing 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  $\beta$  cells, resulting in altered generation of ATP, which would have implications for insulin secretion. In addition, *G6PC2*-induced alterations in  $\beta$  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 (56). These initial findings were confirmed in a comprehensive meta-analysis performed by Weedon et al., demonstrating that rs1799884 was associated with fasting glucose (meta  $P = 1.0 \times 10^{-9}$ ) 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  $\beta$  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,  $P = 1.6 \times 10^{-2}$ ; 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 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](http://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 reported 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).



**METSIM study.** 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 nondiabetic 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 role and general practitioner records. DNA and fasting plasma glucose measurements used in this study relate to the first phase of data collection.

**BWHHS.** The BWHHS consists of female participants, aged 60 to 79 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, age<sup>2</sup>, birth province, and study group. Analyses were carried out in nondiabetic individuals excluding those known to be taking medications that directly affect glucose concentration. Similarly, SardiNIA serum glucose values were adjusted for sex, age, and age<sup>2</sup>. 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, age<sup>2</sup>, 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 age<sup>2</sup>. 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, age<sup>2</sup>, and BMI as covariates. For the Inter99 study, 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 Ferreli, 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 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 1 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 000052); 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 United Kingdom Department of Health policy research program. The DNA extraction and genotyping for BWHHS were funded by 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 Ricovero E Cura Anziari, Rome, Italy. Phone: 39-3334564136; Fax: 39-06-30362896; E-mail: angeloedefante@interfree.it. Or to: Richard M. Watanabe, Keck School of Medicine of USC, Department of Preventive Medicine, 1540 Alcazar St., CHP-220, Los Angeles, California 90089-9011, USA. Phone: (323) 442-2053; Fax: (323) 442-2349; E-mail: rwatanab@usc.edu.

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

1. Reaven, G.M. 1988. Role of insulin resistance in human disease. *Diabetes*. **37**:1595–1607.
2. DeFronzo, R.A. 1987. The triumvirate: B-cell, muscle, liver. A collusion responsible for NIDDM. *Diabetes*. **37**:667–687.
3. National Diabetes Data Group. 1979. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. *Diabetes*. **28**:1039–1057.
4. [No authors listed]. 1985. Diabetes Mellitus: Report of a WHO Study Group. *World Health Organ. Tech. Rep. Ser.* **727**:1–113.
5. The Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. 1997. Report of the expert committee on the diagnosis and classification of diabetes mellitus. *Diabetes Care*. **20**:1183–1197.
6. DeFronzo, R.A., and Ferrannini, E. 1991. Insulin resistance: A multifaceted syndrome responsible for NIDDM, obesity, hypertension, dyslipidemia, and atherosclerotic cardiovascular disease. *Diabetes Care*. **14**:173–194.
7. Xiang, A.H., et al. 2006. Coordinate changes in plasma glucose and pancreatic  $\beta$ -cell function in Latino women at high risk for type 2 diabetes. *Diabetes*. **55**:1074–1079.
8. Mason, C.C., Hanson, R.L., and Knowler, W.C. 2007. Progression to type 2 diabetes characterized by moderate then rapid glucose increases. *Diabetes*. **56**:2054–2061.
9. Rich, S.S. 1990. Mapping genes in diabetes. *Diabetes*. **39**:1315–1319.
10. Ghosh, S., and Schork, N.J. 1996. Genetic analysis of NIDDM: the study of quantitative traits. *Diabetes*. **45**:1–14.
11. Diabetes Prevention Program Research Group. 2002. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N. Engl. J. Med.* **346**:393–403.
12. Tuomilehto, J., et al. 2001. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. *N. Engl. J. Med.* **344**:1343–1350.
13. Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, et al. 2007. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science*. **316**:1331–1336.
14. Scott, L.J., et al. 2007. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science*. **316**:1341–1345.
15. Zeggini, E., et al. 2007. Replication of genome-wide association signals in U.K. samples reveals risk loci for type 2 diabetes. *Science*. **316**:1336–1341.
16. Sladek, R., et al. 2007. A genome-wide association study identified novel risk loci for type 2 diabetes. *Nature*. **445**:881–885.
17. Steinthorsdottir, V., et al. 2007. A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat. Genet.* **39**:770–775.
18. Zeggini, E., et al. 2008. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat. Genet.* **40**:638–645.
19. Beatty, T.H., and Fajans, S.S. 1982. Estimating genetic and non-genetic components of variance for fasting glucose levels in pedigrees ascertained through non-insulin dependent diabetes. *Ann. Hum. Genet.* **46**:355–362.
20. Boehnke, M., Moll, P.P., Kottke, B.A., and Weidman, W.H. 1987. Partitioning the variability of fasting plasma glucose levels in pedigrees. *Am. J. Epidemiol.* **125**:679–689.
21. Sakul, H., et al. 1997. Familiality of physical and metabolic characteristics that predict the development of non-insulin-dependent diabetes mellitus in Pima Indians. *Am. J. Hum. Genet.* **60**:651–656.
22. Watanabe, R.M., et al. 1999. Familiality of quantitative metabolic traits in Finnish families with non-insulin-dependent diabetes mellitus. *Hum. Hered.* **49**:159–168.
23. Henkin, L., et al. 2003. Genetic epidemiology of insulin resistance and visceral adiposity. The IRAS family study design and methods. *Ann. Epidemiol.* **13**:211–217.
24. Pilia, G., et al. 2006. Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genet.* **2**:e132.
25. Weedon, M.N., et al. 2006. A common haplotype of the glucokinase gene alters fasting glucose and birth weight: Association in six studies and population-genetics analyses. *Am. J. Hum. Genet.* **79**:991–1001.
26. Valle, T., et al. 1998. Mapping genes for NIDDM. Design of the Finland-United States Investigation of NIDDM Genetics (FUSION) Study. *Diabetes Care*. **21**:949–958.
27. Silander, K., et al. 2004. A large set of Finnish affected sibling pair families with type 2 diabetes suggests susceptibility loci on chromosomes 6, 11, and 14. *Diabetes*. **53**:821–829.
28. Scuteri, A., et al. 2007. Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genet.* **3**:1200–1210.
29. Hsueh, W.-C., et al. 2001. Genome-wide scan of obesity in the Old Order Amish. *J. Clin. Endocrinol. Metab.* **86**:1199–1205.
30. Sorkin, J., et al. 2005. Exploring the genetics of longevity in the Old Order Amish. *Mech. Ageing Dev.* **126**:347–350.
31. Post, W., et al. 2007. Associations between genetic variants in the NOS1AP (CAPON) gene and cardiac repolarization in the Old Order Amish. *Hum. Hered.* **64**:214–219.
32. The Caerphilly and Speedwell Collaborative Group. 1984. Caerphilly and Speedwell collaborative heart disease studies. *J. Epidemiol. Community Health*. **38**:259–262.
33. Lawlor, D., Bedford, C., Taylor, M., and Ebrahim, S. 2003. Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study. *J. Epidemiol. Community Health*. **57**:134–140.
34. Jorgensen, M.E., et al. 2003. Obesity and central fat pattern among Greenland Inuit and a general population of Denmark (Inter99): relationship to metabolic risk factors. *Int. J. Obes. Relat. Metab. Disord.* **27**:1507–1515.
35. Glümer, C., Jorgensen, T., Borch-Johnsen, K., and Inter99 study. 2003. Prevalences of diabetes and impaired glucose regulation in a Danish population: the Inter99 study. *Diabetes Care*. **26**:2335–2340.
36. Petrolonis, A.J., et al. 2004. Enzymatic characterization of the pancreatic islet-specific glucose-6-phosphatase-related protein (IGRP). *J. Biol. Chem.* **279**:13976–13983.
37. Shieh, J.-J., Pan, C.-J., Mansfield, B.C., and Chou, J.Y. 2005. In islet-specific glucose-6-phosphatase-related protein, the beta cell antigenic sequence that is targeted in diabetes is not responsible for the loss of phosphohydrolase activity. *Diabetologia*. **48**:1851–1859.
38. van Mil, S.W.C., et al. 2004. Benign recurrent intrahepatic cholestasis type 2 is caused by mutations in ABCB11. *Gastroenterology*. **127**:379–384.
39. Lang, C., et al. 2007. Mutations and polymorphisms in the bile salt export pump and the multidrug resistance protein 3 associated with drug-induced liver injury. *Pharmacogenet. Genomics*. **17**:47–60.
40. Devlin, B., and Roeder, K. 1999. Genomic control for association studies. *Biometrics*. **55**:997–1004.
41. Funk, C., Ponelle, C., Scheuerermann, G., and Pantze, M. 2001. Cholestatic potential of troglitazone as a possible factor contributing to troglitazone-induced hepatotoxicity: in vivo and in vitro interac-



- tion at the canalicular bile salt export pump (Bsep) in the rat. *Mol. Pharmacol.* **59**:627–635.
42. Staels, B., and Kuipers, F. 2007. Bile acid sequestrants and the treatment of type 2 diabetes mellitus. *Drugs.* **67**:1383–1392.
43. Mukherjee, R., Wagar, D., Stephens, T.A., Lee-Chan, E., and Singh, B. 2005. Identification of CD4+ T cell-specific epitopes of islet-specific glucose-6-phosphatase catalytic subunit-related protein: a novel beta cell autoantigen in type 1 diabetes. *J. Immunol.* **174**:5306–5315.
44. Wang, Y., et al. 2007. Deletion of the gene encoding the islet-specific glucose-6-phosphatase catalytic subunit-related protein autoantigen results in a mild metabolic phenotype. *Diabetologia.* **50**:774–778.
45. Arden, S.D., et al. 1999. Molecular cloning of a pancreatic islet-specific glucose-6-phosphatase catalytic subunit-related protein. *Diabetes.* **48**:531–542.
46. Shieh, J.-J., Pan, C.-J., Mansfield, B.C., and Chou, J.Y. 2004. The islet-specific glucose-6-phosphatase-related protein, implicated in diabetes, is a glycoprotein embedded in the endoplasmic reticulum membrane. *FEBS Lett.* **562**:160–164.
47. Dogra, R.S., et al. 2006. Alternative splicing of G6PC2, the gene coding for the islet-specific glucose-6-phosphatase catalytic subunit-related protein (IGRP), results in differential expression in human thymus and spleen compared with pancreas. *Diabetologia.* **49**:953–957.
48. Pan, C.-J., Lei, K.-J., Annabi, B., Hemrika, W., and Chou, J.Y. 1998. Transmembrane topology of glucose-6-phosphatase. *J. Biol. Chem.* **273**:6144–6148.
49. Khan, A., et al. 1990. Glucose cycling in islets from healthy and diabetic rats. *Diabetes.* **39**:456–459.
50. Khan, A., et al. 1989. Evidence for the presence of glucose cycling in pancreatic islets of the ob/ob mouse. *J. Biol. Chem.* **264**:9732–9733.
51. Khan, A., et al. 1990. Glucose cycling is markedly enhanced in pancreatic islets of obese hyperglycemic mice. *Endocrinology.* **126**:2413–2416.
52. Vaulont, S., Vasseur-Cognet, M., and Kahn, A. 2000. Glucose regulation of gene transcription. *J. Biol. Chem.* **275**:31555–31558.
53. Stone, L.M., Kahn, S.E., Deeb, S.S., Fujimoto, W.Y., and Porte, D., Jr. 1994. Glucokinase gene variations in Japanese-Americans with a family history of NIDDM. *Diabetes Care.* **17**:1480–1483.
54. Stone, L.M., Kahn, S.E., Fujimoto, W.Y., Deeb, S.S., and Porte, D., Jr. 1996. A variation at position -30 of the  $\beta$ -cell glucokinase gene promoter is associated with reduced  $\beta$ -cell function in middle-aged Japanese-American men. *Diabetes.* **45**:422–428.
55. Rose, C.S., et al. 2005. A -30G>A polymorphism of the beta-cell-specific glucokinase promoter associates with hyperglycemia in the general population of whites. *Diabetes.* **54**:3026–3031.
56. Weedon, M.N., et al. 2005. Genetic regulation of birth weight and fasting glucose by a common polymorphism in the islet promoter of the glucokinase gene. *Diabetes.* **54**:576–581.
57. Malaisse, W.J., Malaisse-Lagae, F., Davies, D.R., Vandercammen, A., and Van Schaftingen, E. 1990. Regulation of glucokinase by a fructose-1-phosphate-sensitive protein in pancreatic islets. *Eur. J. Biochem.* **190**:539–545.
58. Frayling, T.M., et al. 2007. A common variant in the *FTO* gene is associated with body mass index and predisposes to childhood and adult obesity. *Science.* **316**:889–894.
59. Sanna, S., et al. 2008. Common variants in the GDF5-UQCC region are associated with variation in human height. *Nat. Genet.* **40**:198–203.
60. Willer, C.J., et al. 2008. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat. Genet.* **40**:161–169.
61. Li, Y., Willer, C.J., Ding, J., Scheet, P., and Abecasis, G.R. 2008. Markov model for rapid haplotyping and genotype imputation in genome wide studies. *Nat. Genet.* In press.
62. [Anonymous]. 1999. Definition, diagnosis and classification of diabetes mellitus and its complications. Report of a WHO Consultation. WHO. Geneva, Switzerland. [www.diabetes.com.au/pdf/who\\_report.pdf](http://www.diabetes.com.au/pdf/who_report.pdf).
63. Livak, K.J. 1999. Allelic discrimination using fluorescent probes and the 5' nuclease assay. *Genet. Anal.* **14**:143–149.
64. Burdick, J.T., Chen, W.M., Abecasis, G.R., and Cheung, V.G. 2006. In silico method for inferring genotypes in pedigrees. *Nat. Genet.* **38**:1002–1004.
65. Chen, W.-M., and Abecasis, G.R. 2007. Family based association tests for genome wide association scans. *Am. J. Hum. Genet.* **81**:913–926.
66. D'Orazio, P., et al. 2005. Approved IFCC recommendations on reporting results for blood glucose (abbreviated). *Clin. Chem.* **51**:1573–1576.
67. Blangero, J., and Almasy, L. 1997. Multipoint oligogenic linkage analysis of quantitative traits. *Genet. Epidemiol.* **14**:959–964.
68. Almasy, L., and Blangero, J. 1998. Multipoint quantitative-trait linkage analysis in general pedigrees. *Am. J. Hum. Genet.* **62**:1198–1211.

## **SUPPLEMENTARY MATERIAL**

Supplemental Table I. Top 10 Independent\* Genome-wide Associations with Fasting Glucose.

SNP	Chromosome	Position	FUSION Stage 1	SardiNIA	GWA Meta-analysis
			p-value	p-value	p-value†
rs560887	2	169588655	$1.7 \times 10^{-3}$	$4.4 \times 10^{-8}$	$2.8 \times 10^{-10}$
rs9981885	21	19252508	$2.7 \times 10^{-2}$	$1.2 \times 10^{-6}$	$3.2 \times 10^{-7}$
rs1387153	11	92313476	$8.5 \times 10^{-3}$	$3.6 \times 10^{-5}$	$1.0 \times 10^{-6}$
rs2027281	10	29424184	$3.0 \times 10^{-2}$	$2.1 \times 10^{-5}$	$1.8 \times 10^{-6}$
rs420510	21	19226244	$3.2 \times 10^{-2}$	$2.5 \times 10^{-5}$	$2.3 \times 10^{-6}$
rs693793	6	124830968	$8.3 \times 10^{-3}$	$1.1 \times 10^{-4}$	$3.1 \times 10^{-6}$
rs2214108	7	11428183	$4.9 \times 10^{-2}$	$2.8 \times 10^{-5}$	$3.9 \times 10^{-6}$
rs7251204	19	20558498	$3.9 \times 10^{-2}$	$3.7 \times 10^{-5}$	$4.0 \times 10^{-6}$
rs11122355	1	228319928	0.78	$3.9 \times 10^{-7}$	$5.3 \times 10^{-6}$
rs6462079	7	28189067	$8.6 \times 10^{-6}$	$6.8 \times 10^{-3}$	$5.5 \times 10^{-6}$

\* Defined as a pair-wise  $D' < 0.8$

Supplemental Table II. GWA for Fasting Glucose in Non-diabetic Individuals With Meta-analysis  $p$ -value  $\leq 1.0 \times 10^{-4}$  on Chromosome 2.

SNP	Position**	FUSION Stage 1			SardiNIA			GWA Meta-analysis $p$ -value†
		Minor Allele	MAF	$p$ -value	Minor Allele	MAF	$p$ -value	
rs477224	169575990	C	0.184*	$8.9 \times 10^{-2}$	C	0.361	$3.4 \times 10^{-4}$	$7.7 \times 10^{-5}$
rs13431652	169578922	C	0.301	$6.9 \times 10^{-3}$	C	0.412	$2.4 \times 10^{-5}$	$5.6 \times 10^{-7}$
rs573225	169583048	G	0.307	$2.9 \times 10^{-3}$	G	0.436	$4.9 \times 10^{-5}$	$5.7 \times 10^{-7}$
rs560887	169588655	T	0.305*	$1.7 \times 10^{-3}$	T	0.372*	$4.4 \times 10^{-8}$	$2.8 \times 10^{-10}$
rs563694	169599578	C	0.339*	$8.1 \times 10^{-4}$	C	0.455*	$7.6 \times 10^{-5}$	$3.5 \times 10^{-7}$
rs537183	169600153	C	0.340	$8.6 \times 10^{-4}$	C	0.455	$6.5 \times 10^{-5}$	$3.1 \times 10^{-7}$
rs502570	169600466	A	0.340	$9.3 \times 10^{-4}$	A	0.455	$6.5 \times 10^{-5}$	$3.3 \times 10^{-7}$
rs475612	169602253	T	0.343	$1.1 \times 10^{-3}$	T	0.451	$5.7 \times 10^{-5}$	$3.3 \times 10^{-7}$
rs557462	169603102	T	0.343	$1.1 \times 10^{-3}$	C	0.455	$6.4 \times 10^{-5}$	$3.6 \times 10^{-7}$
rs486981	169607656	A	0.343	$1.6 \times 10^{-3}$	A	0.469	$1.0 \times 10^{-4}$	$7.9 \times 10^{-7}$
rs484066	169607988	A	0.379	$1.9 \times 10^{-2}$	A	0.370	$9.0 \times 10^{-5}$	$9.0 \times 10^{-7}$
rs569805	169608387	A	0.342	$1.6 \times 10^{-3}$	A	0.469	$1.0 \times 10^{-4}$	$7.9 \times 10^{-7}$
rs579060	169608546	G	0.342	$1.6 \times 10^{-3}$	G	0.469	$1.0 \times 10^{-4}$	$7.9 \times 10^{-7}$
rs508506	169610462	A	0.342	$1.6 \times 10^{-3}$	A	0.469*	$1.0 \times 10^{-4}$	$7.8 \times 10^{-7}$
rs494874	169614813	T	0.342	$1.6 \times 10^{-3}$	T	0.470*	$1.1 \times 10^{-4}$	$8.7 \times 10^{-7}$
rs552976	169616945	A	0.342	$1.6 \times 10^{-3}$	A	0.472*	$3.5 \times 10^{-5}$	$2.5 \times 10^{-7}$
rs567074	169619938	T	0.442	$8.0 \times 10^{-2}$	C	0.446*	$4.1 \times 10^{-4}$	$8.2 \times 10^{-5}$

rs853789	169626995	A	0.350	$1.4 \times 10^{-3}$	A	0.412*	$2.5 \times 10^{-7}$	$1.4 \times 10^{-9}$
rs853787	169627759	G	0.352	$1.1 \times 10^{-3}$	G	0.412	$2.3 \times 10^{-7}$	$1.0 \times 10^{-9}$
rs862662	169627836	C	0.438	$4.5 \times 10^{-2}$	A	0.449	$2.4 \times 10^{-4}$	$2.9 \times 10^{-5}$
rs853781	169631828	A	0.459	$2.9 \times 10^{-2}$	G	0.449	$2.4 \times 10^{-4}$	$1.9 \times 10^{-5}$
rs853773	169639854	A	0.482	$2.7 \times 10^{-2}$	G	0.480	$3.8 \times 10^{-6}$	$3.2 \times 10^{-7}$

---

\* Based on genotyped data

\*\* Based on NCBI build 35

† Meta-analysis for FUSION stage 1 and SardinIA

Supplemental Table III. Association Between rs563694 and Fasting Glucose in Non-diabetic Individuals Adjusting for SNPs Associated with Type 2 Diabetes from Previous GWA Studies\*.

	Covariate		FUSION	SardiNIA	Meta
Chr	SNP	Gene	p-value	p-value	p-value
	None		$8.0 \times 10^{-4}$	$7.6 \times 10^{-5}$	$3.5 \times 10^{-7}$
3	rs4402960	<i>IGF2BP2</i>	$6.7 \times 10^{-4}$	$8.1 \times 10^{-5}$	$3.3 \times 10^{-7}$
3	rs1801282	<i>PPARG</i>	$8.2 \times 10^{-4}$	$7.5 \times 10^{-5}$	$3.5 \times 10^{-7}$
6	rs7754840	<i>CDKAL1</i>	$8.1 \times 10^{-4}$	$7.8 \times 10^{-5}$	$3.6 \times 10^{-7}$
8	rs13266634	<i>SLC30A8</i>	$1.1 \times 10^{-3}$	$6.3 \times 10^{-5}$	$3.6 \times 10^{-7}$
9	rs10811661	<i>CDKN2A/2B</i>	$1.1 \times 10^{-3}$	$6.4 \times 10^{-5}$	$3.7 \times 10^{-7}$
10	rs1111875	<i>HHEX</i>	$1.1 \times 10^{-3}$	$7.1 \times 10^{-5}$	$4.1 \times 10^{-7}$
10	rs7903146	<i>TCF7L2</i>	$1.1 \times 10^{-3}$	$6.8 \times 10^{-5}$	$3.9 \times 10^{-7}$
11	rs9300039	Chr 11 intragenic	$6.1 \times 10^{-4}$	$7.5 \times 10^{-5}$	$2.8 \times 10^{-7}$
11	rs5215	<i>KCNJ11</i>	$7.1 \times 10^{-4}$	$7.1 \times 10^{-5}$	$3.0 \times 10^{-7}$
16	rs8050136	<i>FTO</i>	$1.1 \times 10^{-3}$	$7.3 \times 10^{-5}$	$4.2 \times 10^{-7}$

\* The direction of the effect of rs563694 on fasting glucose was not altered by the addition of these SNPs as covariates.

Supplemental Table IV. Association Between Fasting Glucose in Non-diabetic Individuals and SNPs Associated with Type 2 Diabetes from Previous GWA Studies.

Chr	SNP	Gene	FUSION	SardiNIA	Meta
			p-value	p-value	p-value
3	rs4402960	<i>IGF2BP2</i>	0.453	0.690	0.751
3	rs1801282	<i>PPARG</i>	0.713	0.983	0.160
6	rs7754840	<i>CDKAL1</i>	0.780	0.666	0.805
8	rs13266634	<i>SLC30A8</i>	0.728	0.738	0.029
9	rs10811661	<i>CDKN2A/2B</i>	0.180	0.794	0.060
10	rs1111875	<i>HHEX</i>	0.360	0.538	0.582
10	rs7903146	<i>TCF7L2</i>	0.816	0.530	0.597
11	rs9300039	Chr 11 intragenic	0.298	0.901	0.851
11	rs5215	<i>KCNJ11</i>	0.187	0.596	0.148
16	rs8050136	<i>FTO</i>	0.810	0.541	0.582