OBJECTIVE—Using the genome-wide association approach, we recently identified the glucokinase regulatory protein gene (GCKR, rs780094) region as a novel quantitative trait locus for plasma triglyceride concentration in Europeans. Here, we sought to study the association of GCKR variants with metabolic phenotype, including measures of glucose homeostasis, to evaluate the GCKR locus in samples of non-European ancestry and to fine-map across the associated genomic interval.

RESEARCH DESIGN AND METHODS—We performed association studies in 12 independent cohorts comprising >45,000 individuals representing several ancestral groups (whites from Northern and Southern Europe, whites from the U.S., African Americans from the U.S., Hispanics of Caribbean origin, and Chinese, Malays, and Asian Indians from Singapore). We conducted genetic fine-mapping across the ~417-kb region of linkage disequilibrium spanning GCKR and 16 other genes on chromosome 2p23 by imputing untyped HapMap single nucleotide polymorphisms (SNPs) and genotyping 104 SNPs across the associated genomic interval.

RESULTS—We provide comprehensive evidence that GCKR rs780094 is associated with opposite effects on fasting plasma triglyceride (Pmeta = 3 × 10^-56) and glucose (Pmeta = 1 × 10^-15) concentrations. In addition, we confirmed recent reports that the same SNP is associated with C-reactive protein (CRP) level (P = 5 × 10^-10). Both fine-mapping approaches revealed a common missense GCKR variant (rs1260326, Pro446Leu, 34% frequency, r^2 = 0.93 with rs780094) as the strongest association signal in the region.

CONCLUSIONS—These findings point to a molecular mechanism in humans by which higher triglycerides and CRP can be coupled with lower plasma glucose concentrations and position GCKR in central pathways regulating both hepatic triglyceride and glucose metabolism. Diabetes 57:3112–3121, 2008

Recently, in the genome-wide association Diabetes Genetics Initiative (DGI) Study for 19 traits, including plasma lipids, we provided evidence that the glucokinase (GCK) regulatory protein gene (GCKR) region was a novel quantitative trait locus associated with plasma triglyceride concentration (1). Of all single nucleotide polymorphisms (SNPs) tested, an intronic SNP at GCKR (rs780094) explained the greatest proportion of interindividual variability in plasma triglycerides (1).

GCKR regulates GCK, which functions as a glucose sensor responsible for glucose phosphorylation in the first step of glycolysis. The discoveries that inactivating muta-
tions in GCK cause maturity onset diabetes of the young type 2 (2) and activating GCK mutations lead to permanent hyperinsulinemic hypoglycemia (3) emphasize that GCK plays a major role in glucose metabolism. GCKR-deficient mice have reduced GCK expression but maintain nearly normal GCK activity and show impaired glucose clearance (4). Furthermore, adenoaviral-mediated overexpression of GCKR in mouse liver increased GCK activity and lowered fasting blood glucose (5) and overexpression of GCK in liver led to lowered blood glucose and increased triglyceride concentrations (6,7). Thus, experimental evidence suggests that perturbation of the GCKR pathway has opposing effects of triglyceride and glucose metabolism.

In our original report, SNP rs780094 in GCKR was associated with fasting triglyceride levels in two independent samples, each of Northern European ancestry (P = 3.7 × 10⁻⁸ and 8.7 × 10⁻⁸, respectively) (1). After initial identification and replication of a chromosomal region associated with a trait, key next steps include extension of the association finding to related phenotypes, validation of the association finding in different ethnicities, and fine-mapping to identify the putative causal variant. Recently, our initial finding was replicated in a Danish study in which a strong association was found between the rs780094 T allele and elevated fasting triglyceride levels but also lower insulin levels, better insulin sensitivity, and a moderately decreased risk of type 2 diabetes (8). In addition, recent genome-wide association studies identified an association between the same GCKR intronic SNP and C-reactive protein (CRP) levels (9,10).

Hereby, we sought to examine the effect of SNP rs780094 on triglycerides and related metabolic traits, including fasting glucose concentrations, in 12 samples representing a range of ancestral groups and including a large prospective study with a mean follow-up time of 23 years. In addition, we performed fine-mapping in one of these samples to identify the strongest association signal in the region.

**RESEARCH DESIGN AND METHODS**

The genetic association studies were performed in 12 study samples as shown in Table 1. For all studies, informed consent was obtained from the study subjects, and the study protocols were approved by local ethics committees. All study cohorts genotyped for GCKR as part of this experiment are included in this report.

The DGI Study material consisting of 2,931 individuals from Finland and Sweden (1,464 patients with type 2 diabetes and 1,467 nondiabetic control subjects) was ascertained as previously described (1,11,12). DGI samples were genotyped on the Affymetrix 500K chip (1) and were used in the present study for the in silico fine-mapping. In addition, the DGI samples were used for the genotype fine-mapping of the GCKR locus and for the analyses of apolipoprotein B (apoB) and free fatty acids (FFAs) according to rs780094.

The Malmo Diet and Cancer Study–Cardiovascular Cohort (MDC-CC) represents 6,103 people that were randomly selected to participate in a study for the in silico fine-mapping. In addition, the DGI samples were used for the identification of the strongest association signal in the region.

In the Malmo Preventive Project (MPP), 33,346 citizens from Malmo in southern Sweden participated in a health screening during 1974–1992 (16). Of individuals participating in the initial screening, 4,931 had died, and 551 were lost from follow-up. Of the eligible individuals, 25,000 were invited to a rescreening visit during 2002–2006, which included a physical examination and assessment of blood samples for measurement of lipids. Of the invited subjects, 17,284 people participated in the rescreening, and of these, 17,037 individuals had DNA, 16,908 individuals had fasting triglyceride data, and 16,976 individuals had fasting glucose data available for the study. The mean follow-up time was 23.4 ± 4.6 years. Of the 17,037 individuals, 5,946 were also participants of the MDC-CC, bringing the number of unique individuals in the MPP (i.e., not included in MDC-CC) to 11,091. Because of the overlap between MDC-CC and MPP materials, all of the phenotype association analyses were performed in the unique cohort of 11,091 participants, whereas the prospective analyses were performed in the whole MPP cohort with DNA and available phenotypes.

Individuals on lipid-lowering medication (n = 3,510) were excluded from the triglyceride analyses, and individuals with diabetes either at baseline or follow-up (n = 3,029) were excluded from the analyses of fasting blood glucose. For the analysis of progression to type 2 diabetes in MPP, 16,061 incident cases and 102,073 type 2 diabetes, were included. Diagnosis of diabetes was confirmed from patient records or based on a fasting plasma glucose concentration >7.0 mmol/L.

During the follow-up visit, a subgroup from the MPP study participated in more extensive metabolic studies, including a euglycemic-hyperinsulinemic clamp combined with indirect calorimetry to measure hepatic glucose output. All individuals underwent a physical examination, and a subgroup of 199 men (Malmo men, MPP-MM) with impaired glucose tolerance (IGT) underwent more extensive metabolic studies, including a hyperinsulinemic-euglycemic clamp combined with infusion of [3-H³]glucose to measure hepatic glucose output (17). After the follow-up time, 66 of the 199 men with IGT at baseline had normal glucose tolerance, 52 had impaired fasting glucose and/or IGT, and 81 had type 2 diabetes. Type 2 diabetic patients were treated either with diet alone (42%) or with oral hypoglycemic agents, which were withheld the day before the clamp.

The Nordic Dietlinazem (NORDIL) Study is an intervention study in 10,881 patients with hypertension from Sweden and Norway (diastolic blood pressure ¼100 mmHg at least twice, mean age 60 ± 7 years) of whom 5,152 Swedish patients provided DNA and were included in the present study (18). The study participants had been randomized to antihypertensive treatment with either the calcium-antagonist diltiazem or ß-blocker/thiazide diuretic to compare efficacy of the two drug therapies to prevent cardiovascular end points (18). Lipoprotein and lipid measurements at the baseline examination (before the initiation of antihypertensive therapy) were studied in this report.

The Diabetes Registry sample consists of 2,777 type 2 diabetic patients from the Skania Diabetes 2000 Registry, a local diabetes registry in southern Sweden (19). The majority of registered patients came from the city of Malmo in southern Sweden. Type 2 diabetes was diagnosed according to 1997 World Health Organization (WHO) criteria (20).

The Botnia-Prevalence, Prediction, and Prevention of Diabetes (Botnia-PPP) Study is a population-based study from the Botnia region of western Finland. The current study was initiated in 2004 in a population comprising ~135,000 individuals. Using a population registry, a random sample of subjects aged 18–75 years was selected: In age-groups 18–29 and 60–74 years, 1 of 10 individuals was randomly selected; and in age-group 30–59 years, 1 of 15 individuals was randomly selected. Altogether, 6,075 individuals were invited to participate in the study, and 3,821 took part. Of these individuals, 3,405 have data and DNA available for the current study.

The FINRISK97 study is a population-based cross-sectional survey which consists of 8,191 people aged 25–74 years from five geographical areas of Finland (the Helsinki, southwestern Finland, North Karelia, Oulu, and Kuopio regions) (21). The study followed the protocol of the WHO Multinational Monitoring of Trends and Determinants in Cardiovascular Disease (MONICA) Project, an international study initiated under the auspices of the WHO.

The Valencia study population comprised 1,608 individuals randomly selected from the Valencia region on the eastern Mediterranean coast of Spain and examined between 1999 and 2003 (22). This sample comprised randomly
selected workers, using a continuously updated computerized population register, and subjects randomly selected from the general population.

The Dallas Heart Study is a multiethnic, probability-based sample of Dallas County, weighted such that 50% of the study population was black (23). Ethnicity was self-reported and consisted of non-Hispanic blacks, non-Hispanic whites, and Hispanics. The study population included 3,469 individuals from one of these three ethnic groups with fasting venous blood samples.

The Genetics of Lipid-Lowering Drugs and Diet Network (GOLDN) Study sample consists of 1,062 individuals and is part of the Program for Genetic Interactions Network (24). The majority of participants in the GOLDN Study were re-recruited from three-generational pedigrees from two National Heart, Lung, and Blood Institute (NHLBI) Family Heart Study (FHS) field centers (Minneapolis, MN, and Salt Lake City, UT) (24). The NHLBI FHS is a multicenter, population-based study of genetic and environmental determinants of cardiovascular disease (CVD) and associated risk factors. Nearly all subjects were of European ancestry.

The longitudinal Boston Puerto Rican Health Study includes 837 Puerto Rican (Hispanics of Caribbean origin) men and women aged 45–75 years in the greater Boston area (25,26). As one of eight nationally funded National Institutes of Health (NIH) Centers on Population Health and Health Disparities, the study is investigating health disparities in the Puerto Rican population. Participants were recruited from the Boston area through door-to-door enumeration, following a sampling scheme based on identification of the 2000 U.S. Census blocks containing Hispanics, and in partnership with community organizations.

The Singaporean National Health Survey 98 (Singaporean NHS-98) study was an initiative to determine the risk factors for the major noncommunicable diseases in Singapore (27,28). A total of 3,973 subjects who participated in the Singaporean NHS-98 and had the data needed for the current study were included in this study. In brief, 11,200 individuals from addresses representing the house-type (a proxy for socioeconomic status) distribution of the entire Singapore housing population were selected from the National Database on Dwellings. From these individuals, a random sample was selected by disproportionate stratification and systematic sampling. The Malays and Indians were oversampled to ensure that prevalence estimates for these minority groups were reliable.

### Lipid and glucose phenotypes.

Plasma total cholesterol, HDL cholesterol, and triglycerides in each study were measured using standard enzymatic methods from fasting blood samples, with the exception of FINRISK7, in which the lipid measurements were performed from blood samples collected in a “semifasting” state; i.e., the participants were instructed to fast for 4 h and to avoid fatty meals earlier during the day. ApoB in the DGI was measured using a radioimmunoassay (Orion Diagnostica, Espoo, Finland), and FFA was measured using an enzymatic colorimetric ACS-ACOD-MEHA method (Wako Chemicals, Neuss, Germany).

Fasting blood or plasma glucose was measured by glucose oxidase methods as previously described (1,16–18, 22–27), and blood glucose was converted to plasma glucose using a correction factor of 1.13. Fasting serum insulin was measured using radioimmunoassay in DGI, MDC-CC, MPP-MM, GOLDN, and Singaporean NHS-98 samples (1,11–13,21,24,25). Homeostasis model assessment (HOMA) insulin resistance index was calculated using the following formula: fasting plasma glucose × fasting insulin/22.5. The insulinogenic index was calculated as ([insulin 30 – fasting insulin]/glucose 30 – fasting glucose).

### Genotyping.

Genotyping was performed either by matrix-assisted laser desorption ionization-time of flight mass spectrometry on the Sequenom MassARRAY platform (San Diego, CA) or by allelic discrimination method on the ABI 7000 instrument (Applied Biosystems, Foster City, CA). The studied SNPs were in Hardy-Weinberg equilibrium in all studied populations (P > 0.01) except rs780094 in the Singapore Chinese population (P = 0.000063). Because of deviation from Hardy-Weinberg equilibrium, a random sample of 12% of the Chinese samples was reanalyzed in a separate assay, and the genotyping error rate was 0.3%.

### Genotype fine-mapping.

**GCKR** rs780094 lies in a large region of linkage disequilibrium on chromosome 2. We defined the associated interval to be ~417 kb based on linkage disequilibrium between the index SNP (rs780094) and SNPs upstream and downstream of the index SNP. rs1049817 was furthest upstream with an r2 >0.25 with the index SNP, and rs10823194 was the furthest downstream with an r2 ≥0.25 with the index SNP. The interval between rs1049817 and rs10823194 spans 416,543 bases (National Center for Biotechnology Information human genome sequence Build 35). In this interval, there are 17 annotated genes: MPV17, GTF3C2, EIF2B4, SNX17, ZNF513, PPML1G, NRBP1, KRTCAP3, IFT172, FNDC9, GCKR, C2ORF16, ZNF512, CCD121X1, XAB1, SUPTL7, and SLC4A1AP. To fine-map the association signal across this interval, we selected 120 SNPs for genotyping based on the following criteria: 1) tag SNPs (n = 33) that captured all common alleles (minor allele frequency >0.05) across the ~417 kb interval at an r2 >0.8; 2) all coding SNPs (n = 83) present in HapMap CEU for the given reference sample; and 3) a set of SNPs predicted to be microRNA binding sites (n = 4). Of these 120 SNPs, 104 were successfully designed for genotyping assays and were genotyped in the DGI sample.

### In silico fine-mapping.

We also conducted fine-mapping using a second approach: imputation of untyped SNPs. We imputed untyped SNPs across the region using a recently developed Markov Chain Haplotyping algorithm (MACH 1.0) (29). This method predicts genotypes for untyped SNPs in a given study using two inputs: genotypes at typed SNPs in the study sample and the entire set of genotypes in HapMap (www.hapmap.org) for a given reference sample. Here, the inputs were the following: genotypes from the Affymetrix 500K array in the DGI study and ~2.2 million SNPs in the HapMap CEU samples (a reference sample of European ancestry).

### Expression studies.

To evaluate whether the transcript level of **GCKR** or **GCK** varied by genotype, we studied 101 human liver samples with both expression studies. To evaluate whether the transcript level of **GCKR** or **GCK** varied by genotype, we studied 101 human liver samples with both expression studies.
laser desorption ionization-time of flight mass spectroscopy using a Sequenom platform. RNA was extracted using the Qiagen RNeasy mini kit on pulverized tissue according to the manufacturer’s protocols (Qiagen). RNA was reverse transcribed to cDNA using the Superscript III first-strand synthesis from tissue according to the manufacturer’s protocols (Qiagen). RNA was reverse transcribed to cDNA using the Superscript III first-strand synthesis from tissue according to the manufacturer’s protocols (Qiagen). RNA was reverse transcribed to cDNA using the Superscript III first-strand synthesis from tissue according to the manufacturer’s protocols (Qiagen).

### Statistical analyses

SNP-phenotype association analyses were performed by multivariate linear regression using an additive genetic model. Because of deviation from normal distribution, fasting triglyceride, blood glucose, and CRP concentrations, and carotid IMT measurements were log transformed before analyses. For each participant, residual values were created for log-triglycerides, HDL cholesterol, and LDL cholesterol after adjustment for age, sex, and diabetes status. Residuals were created for BMI after adjustment for age and sex. For analyses of lipid traits, individuals on lipid-lowering medication were excluded (except in NORDIL and Diabetes Registry, where this information was not available). Log-fasting glucose, HOMA insulin resistance index, and insulinogenic index were studied only in nondiabetic individuals, and residuals were created after adjustment for age and sex. To limit the undue influence of outliers in the regression analysis, for each trait we excluded the bottom and top 0.5% of the trait-level distribution in each study sample. We tested the null hypothesis that the trait residuals do not differ by the analyzed genotypes.

To summarize the statistical evidence across the multiple cohorts, we conducted a fixed-effects variance-weighted meta-analysis. We computed a weighted average of the β-coefficient estimates and SEs (from the linear regression models) using the inverse of the variance in each cohort as weights. Heterogeneity between the studies was tested using the Cochran test. One-way ANOVA was used to compare expression levels of GCK and GCKR in the human liver samples according to genotype. Survival analyses were performed using Cox regression analysis with either age and sex or age, sex, LDL cholesterol, HDL cholesterol, triglycerides, BMI, systolic blood pressure (sBP), diastolic blood pressure (dBP), smoking, family history of myocardial infarction, lipid-lowering medication, antihypertensive medication, and log-CRP as independent predictors.

### RESULTS

#### Association with plasma triglyceride concentrations.

The initial association of rs780094 with triglyceride concentrations was studied in each of 12 study samples, representing a range of ancestral groups (Table 2). In all but one sample, the T allele was associated with higher triglycerides (Table 2; \( P \) for association ranging from 0.29 to 6 \( \times 10^{-10} \)); for example, in the MDC-CC, each copy of the T allele was associated with \( \sim 5.5 \) mg/dL higher triglycerides. Across the studies, SNP rs780094 explained between 0.1 and 1.2% of triglyceride variance (after accounting for age, sex, and diabetes status).

Combining the data from the studied 46,549 individuals provided robust evidence for association between the minor T allele at rs780094 and higher triglyceride levels (meta-analysis \( P = 3 \times 10^{-56} \), Table 2). The minor T-allele frequency varied from 16.1% in U.S. blacks to 47.6% in the Spanish from Valencia, and the effect size per T allele ranged from 0.6 mg/dL in Dallas Heart Study Hispanics to \( \sim 6.2 \) mg/dL in the MPP (Table 2). Mean effect size was \(-4.2 \) mg/dL, and we did not observe significant heterogeneity between the studies (\( P = 0.15 \)). We found that the T allele was associated with higher triglycerides in population-based samples and in cohorts of patients with diabetes and hypertension. Furthermore, the T allele was associated with higher triglycerides regardless of the mean triglyceride level, we had a power of 0.88 to detect an association for a SNP genotype that explained 0.2% of the variance in carotid IMT.

The study sample included a modest number of CVD events (321 incident case subjects), we assessed power to detect an association between SNP rs1260326 has an allele frequency of 34%). For the association of Pro446Leu with carotid IMT (\( n = 4,850 \) individuals analyzed), at 5% significance level, we had a power of 0.88 to detect an association for a SNP genotype that explained 0.2% of the variance in carotid IMT.

### TABLE 1

Continued

<table>
<thead>
<tr>
<th>Spain</th>
<th>Dallas Heart Study</th>
<th>U.S.</th>
<th>Singapore</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valencia</td>
<td>Blacks</td>
<td>Hispanics</td>
<td>Whites</td>
</tr>
<tr>
<td>1,608</td>
<td>1,825</td>
<td>601</td>
<td>1,043</td>
</tr>
<tr>
<td>760/848</td>
<td>770/1,055</td>
<td>251/350</td>
<td>500/543</td>
</tr>
<tr>
<td>42 ± 14</td>
<td>45 ± 10</td>
<td>40 ± 9</td>
<td>45 ± 10</td>
</tr>
<tr>
<td>26 ± 5</td>
<td>32 ± 8</td>
<td>31 ± 7</td>
<td>29 ± 7</td>
</tr>
<tr>
<td>112 ± 66</td>
<td>107 ± 95</td>
<td>151 ± 130</td>
<td>139 ± 107</td>
</tr>
<tr>
<td>51 ± 11</td>
<td>52 ± 15</td>
<td>46 ± 11</td>
<td>48 ± 15</td>
</tr>
<tr>
<td>131 ± 36</td>
<td>105 ± 37</td>
<td>107 ± 33</td>
<td>108 ± 34</td>
</tr>
<tr>
<td>96 ± 24</td>
<td>105 ± 49</td>
<td>106 ± 43</td>
<td>98 ± 34</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>3.9</td>
<td>14.2</td>
<td>12.0</td>
<td>6.6</td>
</tr>
</tbody>
</table>

Data are means ± SD (continuous measures). To convert the values to millimoles per liter, multiply triglycerides by 0.01129, HDL and LDL by 0.02586, and fasting blood glucose (fB-glucose) by 0.0556. BPRHS, Boston Puerto Rican Health Study.
ide level in the sample (e.g., the mean triglyceride concentration is considerably lower in a population-based sample, such as MDC-CC [at 122 mg/dl] compared with the Diabetes Registry cohort [at 231 mg/dl] comprising entirely individuals with type 2 diabetes).

**Association with plasma triglyceride–related metabolic traits.** We next explored the relationship between SNP rs780094 and related metabolic traits, including plasma LDL cholesterol, HDL cholesterol, apolipoprotein concentrations, FFA concentrations, and BMI. SNP rs780094 was not associated with LDL cholesterol or HDL cholesterol in any of the samples (Supplementary Table 1, which is available in an online appendix at http://dx.doi.org/10.1001/db06-0516). As expected, given the correlation between triglyceride and apoB concentration (r = 0.42 [P < 0.0001] and 0.57 [P < 0.0001] in DGI and FINRISK97, respectively), rs780094 was associated with apoB concentration in a meta-analysis of the DGI and FINRISK97 cohorts (P = 7.5 × 10^{-3}), with the T-allele carriers having the highest apoB concentration. Fasting FFA and triglyceride concentrations are weakly correlated (r = 0.25 and 0.20 in DGI and MPP-MM, respectively), and we did not observe any association between rs780094 and fasting FFA (P = 0.39 and 0.90 in DGI and MPP-MM cohorts, respectively) or suppression of FFA levels at 2 h in an OGTT (P = 0.70 and 0.98 in DGI and MPP-MM cohorts, respectively).

**Association with metabolic traits.** In the meta-analysis of the DGI and FINRISK97, rs780094 was nominally associated with BMI in the Singaporean NHS-98 study (P = 0.04) but not in any of the other samples (Supplementary Table 2).

Because two genome-wide association studies recently reported association between CRP levels and rs780094 and rs1260326 (9,10), we tested for association between GCKR rs1260326 and rs780094 and CRP in MDC-CC. Both SNPs were strongly associated with CRP levels with the T-allele carriers having significantly higher levels (CC 2.5 ± 4.7, CT 2.6 ± 4.2, and TT 2.9 ± 4.3 mg/l, P = 4.5 × 10^{-5} in linear regression analysis of rs1260326 adjusted for age and sex).

**Association with insulin sensitivity.** We studied the association of SNP rs780094 with fasting glucose concentrations in 33,995 nondiabetic individuals and HOMA estimates of insulin resistance in 11,084 nondiabetic individuals. Despite the fact that the T allele was consistently associated with higher triglycerides, T-allele carriers had significantly lower fasting plasma glucose levels in six of the studied populations, and a similar trend was observed in the other populations (Table 3, Pmeta = 1 × 10^{-15}; for example, in the MDC-CC, each copy of the T allele was associated with −0.5 mg/dl lower fasting blood glucose. In addition, T-allele carriers were more insulin sensitive as estimated by the HOMA insulin resistance index (Pmeta = 5.0 × 10^{-8}). In DGI and Botnia PPP cohorts, we could calculate the insulinogenic index during an OGTT in 999 and 3,184 nondiabetic individuals, respectively. Insulin secretion capacity calculated as the insulinogenic index did not differ significantly between the different GCKR genotype carriers (P = 0.72 and 0.27, respectively).

Encouraged by these results for intermediate traits, we also tested association of rs780094 with type 2 diabetes in our Nordic cohorts with similar minor allele frequency of the SNP (DGI, MDC-CC, NORDIL, Diabetes Registry, and FINRISK97). Of 24,034 individuals, 5,578 had type 2 diabetes.
<table>
<thead>
<tr>
<th>Country/Study</th>
<th>CC (n = 60)</th>
<th>CT (n = 60)</th>
<th>TT (n = 60)</th>
<th>HOMA (insulin resistance)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finland &amp; Sweden DGI 101</td>
<td>31 (609)</td>
<td>99</td>
<td>31 (170)</td>
<td>0.25</td>
</tr>
<tr>
<td>Sweden MDC-CC 99</td>
<td>9 (2,043)</td>
<td>99</td>
<td>17 (610)</td>
<td>0.02</td>
</tr>
<tr>
<td>Sweden MPP 87</td>
<td>9 (3,893)</td>
<td>87</td>
<td>8 (1,364)</td>
<td>0.00048</td>
</tr>
<tr>
<td>Sweden NORDIL 102</td>
<td>16 (2,024)</td>
<td>102</td>
<td>16 (538)</td>
<td>0.004</td>
</tr>
<tr>
<td>Finland Botnia-PPP 83</td>
<td>10 (1,383)</td>
<td>82</td>
<td>12 (444)</td>
<td>0.00022</td>
</tr>
<tr>
<td>Spain Valencia 94</td>
<td>17 (433)</td>
<td>92</td>
<td>19 (359)</td>
<td>0.01</td>
</tr>
<tr>
<td>U.S. Dallas Heart Study blacks</td>
<td>13 (1,102)</td>
<td>92</td>
<td>10 (42)</td>
<td>0.23</td>
</tr>
<tr>
<td>U.S. Dallas Heart Study Hispanics</td>
<td>11 (222)</td>
<td>94</td>
<td>11 (59)</td>
<td>0.22</td>
</tr>
<tr>
<td>U.S. Dallas Heart Study whites</td>
<td>12 (347)</td>
<td>91</td>
<td>13 (148)</td>
<td>0.32</td>
</tr>
<tr>
<td>U.S. GOLDN 99</td>
<td>13 (353)</td>
<td>100</td>
<td>9 (141)</td>
<td>0.47</td>
</tr>
<tr>
<td>U.S. BPRHS 104</td>
<td>31 (262)</td>
<td>103</td>
<td>18 (218)</td>
<td>0.16</td>
</tr>
<tr>
<td>Singapore Singapore NHS-98 Chinese</td>
<td>19 (823)</td>
<td>99</td>
<td>19 (604)</td>
<td>0.001</td>
</tr>
<tr>
<td>Singapore Singapore NHS-98 Malays</td>
<td>25 (256)</td>
<td>105</td>
<td>14 (128)</td>
<td>0.40</td>
</tr>
<tr>
<td>Singapore Singapore NHS-98 Asian Indians</td>
<td>33 (298)</td>
<td>102</td>
<td>19 (27)</td>
<td>0.22</td>
</tr>
<tr>
<td>Total: 33,995</td>
<td>14,048</td>
<td>15,258</td>
<td>4,689</td>
<td></td>
</tr>
</tbody>
</table>

Data are means ± SD (n) or, for continuous raw measures of glucose, converted to plasma glucose using a correction factor of 1.13. Association analyses were conducted with an outcome variable of residual log-glucose concentration after adjustment for age and sex. For ease of interpretation, glucose concentrations are presented in the table. HOMA insulin resistance index was calculated using the following formula: fasting plasma glucose/millimoles per liter / fasting insulin/22.5. Two-sided P values are shown for all cohorts. To convert the plasma glucose values to millimoles per liter, multiply by 0.055.
gest signal for association with triglycerides (GCKR studied the association of these SNPs with triglyceride levels. We genotyped 104 SNPs across the imputed SNPs (or so-called in silico fine-mapping) (Fig. 1).

GCKR Fine-mapping of the GCKR locus. To define the strongest signal for the association on chromosome 2p23 for triglycerides, a region spanning ~417 kb and containing 17 annotated genes was fine-mapped by two different approaches, imputation of untyped SNPs (29) (or so-called in silico fine-mapping) (A) and genotyping tagging SNPs across the region (B). Both in silico and genotype fine-mapping methods indicated the Pro446Leu as the variant with the strongest association with triglyceride levels. The genotype consensus rate between the imputed genotypes and genotyped genotypes was ~95.7%.

We next examined whether the rs780094 genotype was associated with hepatic glucose output in 125 men who had undergone a hyperinsulinemic-euglycemic clamp with infusion of [3-H]glucose and assessment of basal and clamp clamp hepatic glucose production. GCKR rs780094 T-allele carriers did not have a significantly lower basal rate of hepatic glucose production (CC 2.26 ± 0.27, CT 2.18 ± 0.20, and TT 2.16 ± 0.09 mg · kg⁻¹ · min⁻¹, P = 0.10), but their hepatic glucose output during the hyperinsulinemic state was slightly lower compared with that of C-allele carriers (CC 0.26 ± 0.59, CT 0.20 ± 0.40, and TT 0.09 ± 0.22 mg · kg⁻¹ · min⁻¹, P = 0.01).

Fine-mapping of the GCKR locus. We fine-mapped the GCKR region with two different approaches, by genotyping and coding SNPs across the region and by imputing untyped SNPs (or so-called in silico fine-mapping) (Fig. 1). We genotyped 104 SNPs across the ~417-kb region and studied the association of these SNPs with triglyceride concentrations (supplementary Table 3). A common GCKR coding SNP rs1260326 (Pro446Leu) gave the strongest signal for association with triglycerides (P = 9.4 × 10⁻¹⁰).

Fine-mapping by imputation also revealed that GCKR coding SNP rs1260326 (Pro446Leu) gave the strongest signal for triglyceride concentrations (P = 1.5 × 10⁻¹⁰) in the associated interval on chromosome 2p23. In HapMap CEU, GCKR coding SNP rs1260326 shows strong linkage disequilibrium to the intronic SNP rs780094 (r² = 0.93). We performed regression analysis, including both rs1260326 and rs780094 as predictors of triglyceride levels in MDC-CC, but because of the strong correlation between the SNPs, none of the two were significant in this analysis (P = 0.18 and 0.80 for rs1260326 and rs780094, respectively).

Figure 1 summarizes the results of both fine-mapping approaches. Both the genotyping and the in silico fine-mapping methods indicated the Pro446Leu as the variant with strongest association with triglyceride levels. To evaluate the fidelity of the MACH imputation algorithm (24), we compared the genotypes generated by Sequenom genotyping for 57 SNPs with that predicted by imputation. The genotype consensus rate was 95.7%.

Longitudinal changes in fasting triglyceride and glucose stratified by GCKR Pro446Leu genotype. In the MPP cohort, the Pro446Leu was strongly associated with higher triglycerides and lower fasting blood glucose both at baseline (P = 6 × 10⁻⁴⁰ and 0.0005, respectively) and after the mean follow-up period of 23.4 years (P = 3 × 10⁻¹⁰ and 0.004, respectively) (Fig. 2). In addition, the triglyceride levels of the Leu446 carriers increased more over time compared with those of homozygous Pro446 carriers (P = 8 × 10⁻¹⁰), whereas change in fasting glucose over time did not differ by genotype status (Fig. 2).

In the MPP study, among 17,037 individuals free of type 2 diabetes at baseline, 2,063 (12.1%) individuals developed type 2 diabetes during the follow-up period. Carriage of the Leu allele trended to protect from development of type 2 diabetes (OR 0.96 [95% CI 0.91–1.02], P = 0.27).

Association of GCKR variation with CVD and carotid IMT. In MDC-CC, 321 individuals experienced the first CVD end point during the mean follow-up time of 10.5 ± 1.8 years. Neither rs780094 nor rs1260326 predicted CVD (P = 0.85 and 0.45, respectively). The results were similar when age, sex, LDL cholesterol, HDL cholesterol, triglycerides, BMI, sBP, dBP, smoking, family history of myocardial infarction, lipid-lowering medication, antihypertensive medication, and CRP were included as covariates. GCKR variants were also not associated with carotid IMT in MDC-CC. No association was detected between GCKR variants and common carotid artery IMT (P = 0.94 and 0.63 for rs780094 and rs1260326, respectively).

Hepatic expression of GCK and GCKR according to GCKR genotypes. We next examined whether rs780094 or Pro446Leu was associated with transcript levels of GCKR and/or GCK in human liver. In a modest number of liver samples (n = 60), neither rs780094 nor Pro446Leu genotype was associated with transcript levels of GCK or GCKR (supplementary Table 4).
DISCUSSION

In line with the opposite effects of GCKR-pathway manipulation on glucose and triglyceride concentrations in rodent models and a recent association study in Danes (8), our study provides compelling evidence that common DNA sequence variants in GCKR are associated with opposite effects on fasting triglyceride and glucose concentrations in humans and modest protection from type 2 diabetes. Both imputation and genotype fine-mapping of the GCKR locus yielded a nonsynonymous coding SNP (Pro446Leu) as the strongest association signal, suggesting the hypothesis that this nonsynonymous coding SNP is the causal variant for the observed associations. We also provide evidence that the Leu446 allele carriers increase their triglyceride levels more over time compared with noncarriers. In addition, our data suggest that, at least within regions of high linkage disequilibrium, genotypes predicted by imputation are highly accurate and may provide a good starting point for genotype fine-mapping.

The exact mechanism for the effect of the GCKR variant on blood glucose, triglycerides, and CRP remains to be defined. A potential explanation is the opposite and overriding effects of increased glucose utilization and glycolytic flux on liver glucose and lipid metabolism. With increased glucose utilization and glycolytic flux, PEPCK and glucose-6-phosphatase are downregulated, whereas GCK, phosphofructokinase, and fatty acid synthase are upregulated. These changes increase glycogen synthesis and malonyl CoA concentration and direct fatty-acyl-CoA into de novo lipogenesis and VLDL triglyceride production (33). However, the consequence of in vivo glucose metabolism is enhanced suppression of hepatic glucose output (33). Our observation that the GCKR variant T allele associates with higher triglycerides, lower fasting glucose, and lower hepatic glucose output during a euglycemic-hyperinsulinemic clamp agrees with this hypothesis. Finally, the T-allele carriers had insulin secretion capacity similar to that of noncarriers. Thus, the association is similar to that of GCK-30G/A, which affects the glucose levels needed to induce insulin secretion.

Our human studies propose the hypothesis that GCKR Pro446Leu may mimic the consequences of GCK overexpression in rodent models with upregulation of glucose utilization and VLDL-triglyceride synthesis and downregulation of gluconeogenesis. The Pro446Leu variant has been introduced into rat cDNA but was not found to affect the functional properties of the rat protein when prepared by overexpression in Escherichia coli (34). Unfortunately, the human protein could not be produced using that expression system (32). Although human and rat GCKR share 88% identity, the human protein is importantly different from rat GCKR: human GCKR is a more potent inhibitor of GCK than rat GCKR in the absence of fructose-6-phosphate, and human GCKR has higher affinity for fructose-6-phosphate (35). Thus, the potential impact of the amino acid difference on overall structure and function of human GCKR remains to be defined.

There are conflicting data on the association between circulating triglyceride concentrations and risk of CVD (36,37). Our data combined with data for other common variants suggest a potential explanation for the varying risk associated with high triglycerides. DNA sequence variants in some genes (e.g., APOB) have been associated with both increased triglycerides and markers of increased atherosclerosis risk, such as elevated LDL cholesterol (38). Similarly, common genetic variations in both lipoprotein lipase (LPL) and apoA5 (APOA5) genes (rs328 and rs3133506, respectively) are associated with both increased triglycerides and markers of increased atherosclerosis risk, such as decreased HDL cholesterol (39,40). However, at GCKR Pro446Leu, the variant allele is associated with higher triglycerides and higher CRP levels but also with a favorable metabolic marker, namely decreased glucose. Our finding of no association between the GCKR variant and CVD events or carotid IMT is thus not surprising but instead proposes that the risk of CVD associated with
REFERENCES


