Journal of Human Genetics (2009) 54, 162–168

www.nature.com/jhg

ORIGINAL ARTICLE

Testing the association of novel meta-analysis-derived diabetes risk genes with type II diabetes and related metabolic traits in Asian Indian Sikhs

Dharambir K Sanghera[1], Latonya Been[1], Lyda Ortega[1], Gurpreet S Wander[2], Narinder K Mehra[3], Christopher E Aston[1,4], John J Mulvihill[1] and Sarju Ralhan[2]

A recent meta-analysis on three genome-wide association (GWA) scans identified six loci (*NOTCH2*, *THADA*, *ADAMTS9*, *JAZF1*, *CDC123/CAMKID* and *TSPAN8/LGRS*) highly associated with type II diabetes (T2D) in Caucasians. This investigation seeks to confirm this association with diabetes and related metabolic traits in Khatri Sikh diabetics of North India. We genotyped highly significant variants from each locus in a case–control cohort consisting of 680 T2D cases and 637 normoglycemic (NG) controls. Only *CDC123/CAMKID* (rs12779790) replicated earlier evidence of association with T2D under a dominant model (odds ratio (OR): 1.27; 95% confidence interval (CI): 1.02–1.57; *P*=0.031) during initial testing. However, we could not confirm this association using multiple testing corrections. In a multiple linear-regression analysis, the same variant in the *CDC123/CAMKID* revealed a marked decrease in fasting insulin levels among 'G' (risk) allele carriers independently in NG controls (*P*=0.030) and in T2D cases (*P*=0.009), as well as in the combined sample (*P*=0.003) after adjusting for covariates. Evidence of impaired β-cell function was also observed among 'G' (risk) allele carriers in T2D cases (*P*=0.008) and in a combined cohort (*P*=0.026). Our data could not confirm the role of the remaining variants with risk either for T2D or quantitative phenotypes measuring insulin secretion or insulin resistance. These findings suggest that *CDC123/CAMKID* could be a major risk factor for the development of T2D in Sikhs by affecting β-cell function. To our knowledge, this is the first study reporting the role of recently emerging loci in this high-risk population from the South Asian subcontinent.
Journal of Human Genetics (2009) **54,** 162–168; doi:10.1038/jhg.2009.7; published online 27 February 2009

Keywords: GWAS loci; polymorphisms; type II diabetes; fasting insulin; β-cell dysfunction

INTRODUCTION

Recent genome-wide association (GWA) studies and their meta-analyses of type II diabetes (T2D) in populations of European origin have lead to the identification of multiple new loci with no prior implications in T2D pathogenesis.[1–6] The association of some of these new loci detected during the first round of GWA studies was replicated in other populations of non-European origin. First detected by linkage scan,[7] the *TCF7L2* gene was identified as the strongest candidate gene to be associated with T2D through GWA studies (odds ratio (OR): ~1.35) and has been replicated in subsequent studies in many different ethnicities.[8–16] The association of other newly emerged GWAS loci, such as *IGF2BP2*, *CDKAL1*, *CDKN2A/B*, *HHEX*, *SLC30A8* and *KCNJ11*, was replicated in Japanese;[17,18] *IGF2BP2*, *SLC30A8*, *HHEX*, *CDKAL1*, *CDKN2A/B* and *FTO* in Asians from Hong Kong and Korea;[19] *IGF2BP2*, *CDKAL1*, *CDKN2A/B*, *HHEX* and *SLC30A8* in Han Chinese;[20] and *IGF2BP2*, *PPARG2* and *FTO* in Indian Sikhs.[21]

More recently, a meta-analysis of three large GWA studies by the Diabetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium identified six additional T2D loci (*NOTCH2*, *THADA*, *ADAMTS9*, *JAZF1*, *CDC123/CAMKID* and *TSPAN8/LGRS*) to be strongly associated with increased susceptibility to T2D with modest ORs (1.15−1.3).[6] Three of these six loci (*JAZF1*, *CDC123/CAMKID* and *TSPAN8/LGRS*) revealed significant association with surrogate measures of insulin release and impaired β-cell function in an independent study of Caucasians from Denmark.[22] Another Caucasian study from Southern Germany could not replicate the association of any of these six loci with T2D in their large population-based cohort, except that a weak association was detected in *ADAMTS9* with insulin sensitivity and insulin secretion.[23]

Given the considerable differences in allele frequencies among different ethnic groups,[3,24–26] and variations in diabetic phenotype across ethnicities,[27] an evaluation of these novel loci in T2D in other

[1]Department of Pediatrics, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA; [2]Section of Cardiology, Hero Dayanand Medical College and Heart Institute, Ludhiana, Punjab, India; [3]Department of Transplant Immunology and Immunogenetics, All India Institute of Medical Sciences, New Delhi, India and [4]Section of Biostatistics and Informatics, General Clinical Research Center, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA
Correspondence: Dr DK Sanghera, Department of Pediatrics, Section of Genetics University of Oklahoma Health Sciences Center, 975 NE 10th, BRC-254A, Oklahoma City, OK 73104, USA.
E-mail: dharambir-sanghera@ouhsc.edu
Received 13 October 2008; revised 27 December 2008; accepted 4 January 2009; published online 27 February 2009

ethnic populations is important. The aim of this replication study is to determine whether the variations in these six new loci contribute to the risk of T2D in a high-risk community of Asian Indian Sikhs.

MATERIALS AND METHODS

Human subjects

The study subjects are part of the ongoing Sikh Diabetes Study (SDS).[28] This study focuses on an endogamous community of Khatri Sikhs living in the Northern states of India, including Punjab, Haryana, Himachal Pradesh, Delhi and Jammu and Kashmir. The DNA and serum samples of 680 T2D case (375 men and 305 women) and 637 normoglycemic (NG) control (320 men and 317 women) subjects were used in this investigation. Of these 680 cases, 357 subjects were from family material (one index case from each family) and the remaining 323 individuals were unrelated T2D cases from the same Sikh community and geographic location. The cases were $\geqslant$25 years with mean age at the time of recruitment (mean ± s.d.) of 55.4 ± 11.0 years. The diagnosis of T2D was confirmed by scrutinizing medical records for symptoms, use of medication and measuring fasting glucose levels following the guidelines of the American Diabetes Association.[29,30] T2D was defined as a medical record indicating either (1) a fasting plasma glucose level $\geqslant$126 mg per 100 ml or $\geqslant$7.0 mmol l^{-1} after a minimum 12-h fast, or (2) a 2-h post glucose level (2-h oral glucose tolerance test (OGTT)) $\geqslant$200 mg per 100 ml or $\geqslant$11.1 mmol l^{-1} on more than one occasion with symptoms of diabetes. Impaired fasting glucose was defined as fasting plasma glucose $>$100 mg per 100 ml (5.6 mmol l^{-1}) but $<$126 mg per 100 ml (7.0 mmol l^{-1}). Impaired glucose tolerance was defined as a 2-h OGTT $\geqslant$140 mg per 100 ml (7.8 mmol l^{-1}) but $<$200 mg per 100 ml (11.1 mmol l^{-1}). In the absence of medical record information, we confirmed self-reported T2D cases by performing a 2-h OGTT following the criteria of the World Health Organization (WHO) (75 g oral load of glucose). Subjects with impaired fasting glucose or impaired glucose tolerance were considered as 'pre-diabetics' and were not included in this study. Body mass index (BMI) was calculated as (weight (kg)/height (meter)^2). Waist and hip circumferences were taken with a metal tape at the abdomen and at the hip.

The NG participants were from the same Khatri Sikh community and were recruited from the same geographic origin as the T2D patients, as described earlier.[28] The majority of the subjects were recruited from the state of Punjab in Northern India. Individuals of South, East and Central Indian origin or those with type I diabetes or family members with type I diabetes or rare forms of T2D denoted maturity-onset diabetes of young and secondary diabetes (for example, diabetes due to hemochromatosis or pancreatitis) were excluded. Of the 637 NG controls (320 men and 317 women), 262 were non-diabetic spouses of diabetic patients, and the remaining 375 subjects were unrelated non-T2D individuals, who had no family history of T2D, but some had other chronic illnesses such as hypertension, coronary heart disease or arthritis. The selection of controls was based on a fasting glycemia $<$100 mg per 100 ml ($<$5.6 mmol l^{-1}) or a 2-h glucose $<$140 mg per 100 ml ($<$7.8 mmol l^{-1}). The average age of NG controls (mean ± s.d.) was 50.6 ± 13.5 years. In general, Sikhs do not smoke for religious and cultural reasons and about 50% of them are lifelong vegetarians. Clinical characteristics of the SDS subjects used for this investigation are summarized in Table 1. All blood samples were obtained at the baseline visit and all participants provided written informed consent for investigations. All SDS protocols and consent documents were reviewed and approved by the University of Oklahoma Institutional Review Boards as well as by the Human Subject Protection Committees at the participating hospitals and institutes in India.

Table 1 Clinical characteristics of the study population stratified by gender and disease (mean ± s.d.)

	NG controls (n=637)	*T2D cases (n=680)*	
Gender	*(320 males/ 317 females)*	*(375 males/ 305 females)*	P-*value*[a] *(two-tailed)*
Age (years)			
Males	51.1 ± 15.1	55.8 ± 11.2	$<$0.001
Females	50.2 ± 13.2	56.9 ± 10.8	$<$0.001
Age at diagnosis (years)			
Males	—	45.8 ± 10.6	—
Females	—	48.1 ± 10.5	—
Duration of diabetes (years)			
Males	—	10.3 ± 6.6	—
Females	—	9.1 ± 5.5	—
BMI			
Males	26.3 ± 4.5	26.5 ± 4.7	0.314
Females	27.0 ± 5.1	28.0 ± 5.0	0.008
Waist (cm)			
Males	93.2 ± 11.5	94.8 ± 10.5	0.058
Females	87.6 ± 10.8	92.8 ± 11.4	$<$0.0001
Hip (cm)			
Males	96.2 ± 8.8	96.0 ± 8.4	0.818
Females	97.0 ± 10.2	99.5 ± 10.3	0.002
Fasting glucose (mmol l^{-1})			
Males	5.4 ± 0.7	10.3 ± 3.7	$<$0.0001
Females	5.4 ± 0.7	10.6 ± 3.8	$<$0.0001
2-h Glucose[b] *(mmol l^{-1})*			
Males	5.8 ± 1.1	13.0 ± 4.4	$<$0.0001
Females	6.0 ± 0.9	13.1 ± 4.6	$<$0.0001
Fasting insulin (pmol l^{-1})			
Males	58.0 ± 51.5	45.2 ± 41.0	0.002
Females	61.5 ± 56.3	44.7 ± 36.5	$<$0.001

Abbreviations: BMI, body mass index; NG, normoglycemic; T2D, type II diabetes.
[a]Difference between T2D cases and NG controls.
[b]Quantitative data on 2-h glucose was available in 671 subjects (441 NG controls and 230 T2D cases).

Metabolic assays

Quantitation of lipids (total cholesterol, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, very low-density lipoprotein cholesterol and triglycerides) and insulin levels was measured in serum. Lipids were quantified by using standard enzymatic methods (Roche, Basel, Switzerland). Insulin was measured by a radioimmunoassay (Diagnostic Products, Los Angeles, CA, USA). All quantitative parameters were determined by following the manufacturer's instructions using a Hitachi 902 auto-analyzer (Roche, Basel, Switzerland). Homeostasis model assessment (HOMA) for insulin resistance (HOMA-IR) was calculated as (fasting glucose$\times$fasting insulin/22.5). Pancreatic β-cell function (HOMA-B) was calculated as (fasting insulin$\times$20)/(fasting glucose$-$3.5), as described by Matthews *et al.*[31]

SNP genotyping

DNA was extracted from buffy coats using QiaAmp blood kits (Qiagen, Chatsworth, CA, USA) or by the salting out procedure.[32] Genotyping for all investigated single nucleotide polymorphisms (SNPs) was performed using TaqMan Pre-Designed or TaqMan Made-to-Order SNP genotyping assays from Applied Biosystems (ABI, Foster City, CA, USA). TaqMan genotyping reactions were performed on an ABI 7900HT genetic analyzer using 2 µl of (10 ng μl^{-1}) genomic DNA following the manufacturer's instructions. Fluorescence was detected on an ABI prism 7900HT sequence detection system (ABI). Genotypes were scored by analyzing data on both real-time and allele discrimination assay platforms using SDS software provided by the ABI. Data quality for SNP

genotyping was checked by establishing reproducibility of control DNA samples. For quality control, 30 replicate positive controls and 8 negative controls were included in each run to match the concordance, and the discrepancy in the concordance was $<0.2\%$. Genotyping success rate ranged between 95 and 99.3% for all the investigated SNPs.

Statistical analysis

The Hardy–Weinberg equilibrium was evaluated using Pearson's χ^2 1 d.f. goodness-of-fit test separately for cases and controls. The allele frequencies in T2D cases were compared with those in controls using a χ^2 test or Fisher's exact probability test, where appropriate. Statistical evaluations for testing genetic effects of association between the case–control status and each individual SNP, measured by the ORs and its corresponding 95% confidence limits, were estimated using unconditional logistic regression before and after adjusting for age, gender and other covariates. Association analyses were performed assuming codominant, dominant and recessive models using SNPassoc[33] and SPSS (v. 15.0) for Windows. In all analyses, the common homozygote genotype in the control population was defined as the reference category. Statistical power was assessed using the Genetic Power Calculator[34]. We also applied Bonferroni's correction for multiple testing.

A multiple linear regression analysis was used to examine the impact of these variants on quantitative risk variables for T2D, including fasting insulin, glucose and lipid levels. Skewed variables were detected by Shapiro-Wilk's test for the continuous traits. Subsequently, triglycerides, total cholesterol, low-density lipoprotein cholesterol, fasting glucose, fasting insulin, HOMA-IR and HOMA-B were normalized by log-transformation before statistical comparisons, and all *P*-values were derived from analyses of transformed data. Significant covariates for each dependent trait were identified by using Spearman's correlation and step-wise multiple linear regression with an overall 5% level of significance. These analyses were performed using dominant, codominant and recessive genetic models and, the best fitting model was chosen for each SNP. Analyses were adjusted for the confounding effects of age, BMI, sex, medication and also for disease status in the combined analysis, where appropriate. For lipid traits, the individuals on lipid-lowering medications were excluded.

Table 2 Association of candidate SNPs with T2D

					OR (95% CI)[d]		
Gene/SNP	*Genomic position*[a]	*Genotype*[b]	*NG*[c] *(%)*	*T2D (%)*	*Dominant*	*Codominant*	*Recessive*
NOTCH2	Chromosome 1						
rs10923931	120319482	GG	393 (63.3)	403 (60.6)	1.11 (0.92–1.34)	1.21(1.00–1.47)	1.46 (0.94–2.28)
		GT	203 (32.7)	220 (33.1)			
		TT	24 (3.9)	42 (6.3)	*P*=0.268	*P*=0.053	*P*=0.095
		G/**T**	0.80/0.20	0.77/0.23			
THADA	Chromosome 2						
rs7578597	43586327	TT	444 (72.5)	514 (76.1)	0.88 (0.71–1.08)	0.86 (0.68–1.08)	1.21 (0.58–2.51)
		CT	158 (25.8)	148 (21.9)			
		CC	10 (1.6)	13 (1.9)	*P*=0.212	*P*=0.190	*P*=0.605
		T/C	0.85/0.15	0.86/0.14			
ADAMTS9	Chromosome 3						
rs4607103	64686944	CC	161 (26.3)	194 (29.4)	0.98 (0.80–1.20)	0.93 (0.80–1.09)	1.02 (0.82–1.25)
		CT	287 (47.0)	284 (43.0)			
		TT	162 (26.6)	182 (27.6)	*P*=0.859	*P*=0.370	*P*=0.887
		C/T	0.50/0.50	0.51/0.49			
JAZF1	Chromosome 7						
rs864745	28147081	TT	284 (46.3)	286 (43.3)	1.01 (0.84–1.22)	1.07 (0.90–1.26)	1.01 (0.76–1.34)
		CT	260 (42.4)	294 (44.5)			
		CC	69 (11.2)	80 (12.1)	*P*=0.898	*P*=0.470	*P*=0.964
		T/C	0.68/0.32	0.66/0.34			
CDC123/CAMKID	Chromosome 10						
rs12779790	12368016	AA	474 (76.6)	457 (71.4)	1.27 (1.02–1.57)	1.27 (1.00–1.67)	0.94 (0.46–1.90)
		AG	134 (21.6)	170 (26.6)			
		GG	11 (1.7)	13 (2.0)	***P*=0.031**	*P*=0.052	*P*=0.853
		A/**G**	0.87/0.13	0.85/0.15			
TSPAN8/LGR5	Chromosome 12						
rs7961581	69949369	TT	258 (42.4)	306 (46.3)	0.99 (0.83–1.19)	0.96 (0.81–1.13)	1.29 (0.98–1.70)
		CT	272 (44.7)	266 (40.2)			
		CC	78 (12.8)	89 (13.5)	*P*=0.935	*P*=0.610	*P*=0.067
		T/**C**	0.65/0.35	0.66/0.34			

Abbreviations: CI, confidence interval; BMI, body mass index; NG, normoglycemic; OR, odds ratio; SNP, single nucleotide polymorphism; T2D, type II diabetes
[a]Position of SNPs on chromosome has been taken from NCBI (Build 36.3).
[b]Risk allele (based on an earlier meta-analysis scan[6]) is denoted in bold.
[c]Normoglycemic controls.
[d]ORs were adjusted for age, sex, and BMI.
Bold fonts indicate significant *P*-values.

RESULTS

We genotyped 680 T2D patients and 637 NG controls. These subjects were genotyped for *NOTCH2* (rs10923931), *THADA* (rs7578597), *ADAMTS9* (rs4607103), *JAZF1* (rs864745), *CDC123/CAMKID* (rs12779790) and *TSPAN/LGR5* (rs7961581) SNPs and analyzed for association with T2D and quantitative sub-phenotypes of T2D. The case–control cohort used in this investigation was matched for ethnicity, culture and geographical locations. Table 1 shows the clinical and physical characteristics of study subjects stratified by gender and disease. The information on the genomic location of the investigated SNPs and minor allele frequencies is summarized in Table 2. The genotype distribution of all investigated SNPs was in the Hardy–Weinberg equilibrium in both cases and controls.

Of these six SNPs, only *CDC123/CAMKID* revealed a significant evidence of association with T2D in Sikhs after appropriately adjusting for the confounding variables of age, sex and BMI. The adjusted OR for *CDC123/CAMKID* under dominant genetic model was 1.27 (95% CI: 1.02–1.57; P=0.031). A marginally significant trend was also observed under the codominant model (OR: 1.27; 95% CI: 1.00–1.67; P=0.052). The variant in *NOTCH2* also revealed a weak association with T2D under the codominant model (OR: 1.21; 95% CI: 1.00–1.47; P=0.053) (Table 2). However, after applying Bonferroni's correction testing for six SNPs for multiple comparisons, none of the investigated SNPs remained statistically significant. No other variant from the remaining loci was associated with T2D in our study. The general estimates of power in this sample to detect the effects of variants with ORs between 1.2 and 1.5 at α=0.05 were 65, 70 and >82% when the frequency of the risk allele was 0.15, 0.20 and >0.30, respectively. Decreasing the OR to 1.1 reduced these power estimates substantially to 30, 33 and >40%.

We next explored the association of these investigated SNPs with T2D-related quantitative traits controlling insulin action (HOMA-IR) and insulin secretion (HOMA-B), adiposity and lipid-related traits. The multiple linear regression analysis after adjusting confounding variable of age, BMI and sex revealed a significant association of the *CDC123/CAMKID* (rs12779790) variant with fasting insulin levels both in T2D cases (P=0.009) and NG controls (P=0.030), and also in a combined sample of T2D cases+NG controls after adjusting for the effects of disease status along with other significant covariates (P=0.003). As shown in Table 3, the mean values of fasting insulin levels (95% CI) (pmoll^{-1}) markedly decreased among GG homozygous individuals with G (risk) allele in both NG controls (AA: 62.0 (56.3–67.4); AG: 61.0 (52.8–75.7); and GG: 26.4 (18.1–37.5); P=0.030) and T2D cases (AA: 49.3 (45.1–53.5); AG: 40.3 (34.7–46.5); and GG: 34.0 (19.9–51.6); P=0. 009). Our results also revealed significant evidence of impaired β-cell function among G-allele carriers in both T2D cases (P=0.008) and a combined sample (P=0.026) but not in NG controls (P=0.572) (Table 3). There was a gradual decline in pancreatic β-cell function among G (risk) allele carriers, as suggested by mean (95% CI) HOMA-B levels (AA: 20.5 (18.6–22.7); AG: 16.0 (13.3–19.1); and GG: 13.5 (5.2–34.9); P=0.008) among T2D cases, and (AA: 38.9 (35.6–42.7); AG: 32.5 (27.1–38.8); and GG: 23.8 (13.2–40.9); P=0.026) in the combined cohort after controlling for the effects of disease status and medication. We also observed a weak association of *NOTCH2* variants with fasting insulin levels among T2D patients (P=0.033) but not in NG controls (P=0.720) (Supplementary Table 1s). None of the remaining SNPs revealed association with any metabolic trait related to insulin resistance, insulin secretion, T2D or obesity (Supplementary Tables 1s–5s). Conservative estimates of the statistical power for quantitative analysis were calculated assuming an additive genetic model and a 0.005 level of significance after Bonferroni's adjustment for considering 10 independent variables. The power to detect inter-genotypic differences in insulin levels as small as 10% was well in excess of 85% at allele frequency ranging from 0.13 to 0.40. HOMA-B levels with differences of 20% were detectable with 57% power for allele frequencies of 0.13 and they increased to 80% for allele frequencies of about 0.20.

DISCUSSION

In this investigation, we have tested the association of six latest candidate SNPs identified by the meta-analysis of GWA data[6] in our T2D case–control cohort of Khatri Sikhs of Asian Indian origin. Of the six SNPs, only *CDC123/CAMKID*; (rs12779790) replicated earlier evidence of association with T2D and other metabolic sub-phenotypes controlling insulin secretion in our sample. By applying Bonferroni's correction, the observed association with T2D phenotype would no longer be statistically significant for *CDC123/CAMKID*. However, the observed association of this variant with quantitative measures of insulin secretion in our sample suggests that the association is true. Our results in the multiple linear regression analysis revealed a significant association of this SNP with fasting insulin in NG controls (P=0.030; R^2=0.017), T2D cases (P=0.009; R^2=0.020) as well as in a combined sample (P=0.003; R^2=0.016) under an additive genetic model. Moreover, the P-value associated with fasting insulin levels remained statistically significant after applying Bonferroni's correction in the combined cohort (P=0.03). It is noted that the impairment in the β-cell function was more pronounced among T2D patients (P=0.008; R^2=0.017) but not in NG controls (P=0.572; R^2=0.001), perhaps because of the increased demand for insulin during acute conditions of hyperglycemia.

From these results, it appears that the genetic variation in *CDC123/CAMKID* affects glucose-stimulated insulin release and may be responsible in impaired β-cell function that may subsequently contribute to putative T2D risk in Sikhs. However, this variant explained only 2% of the overall contribution of this gene in controlling fasting insulin levels in this sample. These findings suggest that perhaps the causal variant in this gene still remains to be identified. Interestingly, the same SNP has been shown to affect pancreatic β-cell function in a large population-based cohort from Denmark.[22] The authors showed association of G (risk) allele with insulin secretion under additive (P=0.02) and recessive (P=8×10^{-5}) genetic models. However, at the same time, the association of this variant could not be reproduced with T2D or impaired β-cell function in another Caucasian study from Southern Germany.[23] The possible reason for this inconsistency in confirming the role of these loci with diabetic and pre-diabetic phenotypes within Euro-Caucasians could be due to the genuine genetic heterogeneity existing among 'Caucasians' because of wide geographic locations and genetic histories. In the current sample of Asian Indian Sikhs, false-positive or false-negative associations arising from genetic heterogeneity owing to population admixture are less likely to be the reason for non-replication of association with these loci in our sample. Our carefully ascertained, relatively homogeneous case–control sample of Asian Indian Sikhs belongs to a single geographic location and is collected from an endogamous community of urban Sikhs from the Khatri community.

Another possible reason for inconsistency among the Euro-Caucasian-based studies is a lack of sufficient statistical power to detect the small contribution of these loci (ORs: ~1.15–1.3)[6] to the complex phenotype of T2D.[35] Unfortunately, this case–control sample of Sikhs is similarly hampered by low statistical power at the lower end of this range of ORs. The minimum ORs detectable with 80% power for risk

Table 3 Association of *CDC123/CAMKID* (rs12779790) with anthropometric and metabolic traits controlling glucose homeostasis and lipid metabolism

	Controls (n=*619*)				*Cases* (n=*640*)			
	AA	*AG*	*GG*	P *corrected*[a]	*AA*	*AG*	*GG*	P *corrected*[a]
N (males/females)	474 (250/224)	134 (54/80)	11 (5/6)		457 (253/204)	170 (88/82)	13 (7/6)	
Age (years)	51.4 (50.2–52.8)	47.8 (45.4–50.2)	56.7 (46.2–67.3)	0.040	55.8 (54.8–56.8)	57.2 (55.5–58.9)	59.41 (52.4–66.4)	0.109
BMI (kg m^{-2})	26.5 (25.9–26.8)	26.8 (26.0–27.6)	25.9 (22.5–29.3)	0.603	27.2 (26.7–27.6)	27.3 (26.5–28.0)	25.99 (23.4–28.6)	0.942
Waist (cm)	90.7 (89.6–91.8)	89.8 (87.9–91.6)	94.2 (85.9–102.4)	0.557	93.9 (92.9–94.9)	93.3 (91.6–94.9)	90.6 (83.9–97.3)	0.363
Hip (cm)	96.4 (95.5–97.3)	96.7 (95.2–98.1)	99.8 (91.3–108.3)	0.538	97.2 (96.4–98.1)	98.0 (96.6–99.4)	97.9 (91.1–104.8)	0.632
Fasting glucose (mmol l^{-1})[a]	5.4 (5.3–5.4)	5.3 (5.1–5.4)	5.4 (5.2–5.6)	0.170	9.7(9.4 –10.1)	10.4 (9.8–11.0)	10.6 (6.8–16.6)	0.060
Fasting insulin (pmol l^{-1})[a]	62.0 (56.3–67.4)	61.0 (52.8–75.7)	26.4 (18.1–37.5)	**0.030**[c]	49.3 (45.1–53.5)	40.3 (34.7–46.5)	34.0 (19.9–51.6)	**0.009**[c]
HOMA-IR[a]	1.7 (1.5–1.9)	1.8 (1.5–2.1)	1.0 (0.7–1.4)	0.877	2.7 (2.4–3.0)	2.4 (2.0–2.8)	2.1 (1.0–4.2)	0.162
HOMA-B[a]	77.5 (69.7–85.8)	86.5 (70.0–106.7)	45.6 (29.1–70.8)	0.572	20.5 (18.6–22.7)	16.0 (13.3–19.1)	13.5 (5.2–34.9)	**0.008**[c]
Total cholesterol (mg per 100 ml)[a]	172.4 (169.2–180.1)	175.9 (169.2–185.2)	186.8 (151.8–232.0)	0.541	181.3 (175.8–185.8)	178.7 (169.8–184.7)	186.8 (166.9–207.5)	0.522
Triglycerides (mg per 100 ml)[a]	147.0 (140.3–153.9)	153.1 (139.7–167.9)	142.4 (97.9–207.3)	0.574	167.9 (159.4–177.0)	147.8 (137.0–159.4)	143.2 (121.0–169.4)	0.026
LDL-cholesterol (mg per 100 ml)[a]	99.7 (96.3–403.0)	101.2 (94.9–229.0)	110.5 (87.3–190.0)	0.429	100.6 (96.9–104.5)	104.1 (98.2–110.5)	113.0 (96.0–133.1)	0.406

	Combined (n=*1259*)			
N (males/females)	931 (503/428)	304 (142/162)	24 (12/12)	
Age (years)	53.6 (52.8–54.4)	53.1 (51.6–54.6)	58.2 (52.5–63.9)	0.081
BMI (kg m^{-2})	26.8 (26.4–27.1)	27.1 (26.5–27.6)	26.0 (24.0–27.9)	0.361
Waist (cm)	92.3 (91.6–93.0)	91.7 (90.5–92.9)	92.2 (87.3–97.0)	0.449
Hip (cm)	96.8 (96.2–97.4)	97.4 (96.4–98.4)	98.7 (93.8–103.6)	0.426
Fasting glucose (mmol l^{-1})[b]	7.4 (7.2–7.6)	7.8 (7.4–8.3)	7.3 (5.7–9.3)	0.063
Fasting insulin (pmol l^{-1})[b]	55.3 (52.1–58.8)	49.5 (44.1–55.6)	30.5 (22.1–42.2)	**0.003**[c]
HOMA-IR[b]	2.2 (2.0–2.3)	2.1 (1.8–2.4)	1.4 (0.9–2.2)	0.092
HOMA-B[b]	38.9 (35.6–42.7)	32.5 (27.1–38.8)	23.8 (13.2–40.9)	**0.026**[c]
Total cholesterol (mg per 100 ml)[b]	177.7 (173.9–181.3)	177.1 (171.7–182.6)	186.8 (169.3–206.1)	0.389
Triglycerides (mg per 100 ml)[b]	157.0 (151.6–162.6)	150.1 (141.7–159.1)	142.8 (120.9–168.7)	0.167
LDL-cholesterol (mg per 100 ml)[b]	100.2 (97.6–102.8)	102.9 (98.5–107.4)	111.9 (98.9–126.7)	0.186

Abbreviations: ANOVA, analysis of variance; BMI, body mass index; HOMA-IR, homeostasis model assessment-insulin resistance; HOMA-B, homeostasis model assessment-β-cell function; LDL, low-density lipoprotein.
Data represent mean and its 95% confidence interval from ANOVA.
[a]*P*-values were derived from multiple linear regression analysis using the best genetic model after adjusting for age, sex, BMI, medication and disease status, where appropriate.
[b]Log-transformed traits; mean values of log-transformed variables are presented as geometric mean and other un-transformed variables are presented as arithmetic mean.
[c]Bold fonts indicate statistical significance.

allele frequencies of 0.15, 0.20 and 0.30 were 1.25, 1.22 and 1.19, respectively. Consequently, the lack of association of some previously reported loci with T2D in this Sikh population may not be conclusive that these genetic variants have no effect on risk for T2D.

A further reason for the lack of association of these SNPs with risk for T2D could be that in the Euro-Caucasian studies the reported SNPs were not functional but rather were in linkage disequilibrium (LD) with nearby SNPs, and perhaps because of ethnic differences, this LD was not present in our sample. Clearly, there was a significant difference ($P \leqslant 0.001$) in the distribution of risk allele frequencies in our Sikh sample compared with Caucasians reported in a meta-analysis study[6] (*NOTCH2*: 0.20 vs 0.10; *THADA*: 0.85 vs 0.90; *ADAMTS9*: 0.50 vs 0.76; *JAZF1*: 0.68 vs 0.50; *CDC123/CAMKID*: 0.13 vs 0.18; *TSPAN8/LGR5*: 0.35 vs 0.27, respectively, for Sikhs and Caucasians). Therefore, it is likely that the causal variant in these loci could be ethnic-specific and could be present elsewhere in the same or nearby gene. Furthermore, genetic differences in phenotype heterogeneity and characterization also cannot be ruled out.

GWA studies allow the detection of novel genes with previously unknown roles in the biology of a disease. The detection of new loci may help unravel the disease pathophysiology in greater detail.[35] However, it is still challenging to detect the true candidates using GWA studies in complex diseases, especially for the loci that are common because of ethnic variations.[36] *CDC123/CAMKID* is an evolutionary conserved protein involved in controlling cell cycle and its levels of expression are regulated by nutrients in *Saccharomyces cerevisiae*.[20] However, its role in T2D pathophysiology is unclear. This variant (rs12779790) lies in the inter-genic region (about 90 kb) from *CDC123/CAMKID* and (~63.5 kb) from *CAMKID*. It is also possible that the causative variant is located in *CAMKID* or other nearby gene regions that may be affecting insulin secretion or insulin sensitivity. Given the sparse knowledge available on the biological function of these genes, it is unclear how these genes are involved in the T2D pathogenesis. Unlike *CDC123/CAMKID*, no other variant revealed significant association with quantitative variables involved in glucose homeostasis (fasting glucose, fasting insulin, HOMA-IR or HOMA-B), adiposity or lipid levels, except for the weak association of *NOTCH2* variants affecting fasting insulin levels (P=0.033) in T2D cases (Supplementary Table 1s). It is noted that *NOTCH2* is expressed in ductal cells of branching pancreatic buds and plays a key role in fetal pancreatic development. In addition, modulation of NOTCH signaling can affect pancreatic β-cell formation. Although our data could not replicate the positive association of *NOTCH2* with T2D, marginally significant association of T (risk) allele with reduced insulin secretion suggests the possible role of *NOTCH2* variation in affecting β-cell function in Sikhs. Perhaps a causal variant in this gene with a greater risk is yet to be identified. Our findings also confirmed the negative association revealed in *JAZF1*, *THADA* and *ADAMTS9* with pre-diabetic phenotypes in another Caucasian sample from Southern Germany.[23] Our results also agree with the negative findings reported in *THADA* and *ADAMTS9* with surrogate measures of insulin release and insulin action in Caucasian study from Denmark.[22] The inability of this study as well as of the earlier studies to clearly replicate meta-analysis findings may suggest a weak association of these loci to T2D.

Nevertheless, false-negative results in our data set also cannot be ruled out because of limited statistical power to detect the role of some of these SNPs in T2D or in β-cell function (insulin secretion) or insulin sensitivity. Precise phenotype data, including the measurements of insulin secretion (such as intravenous glucose tolerance test) or hyperglycemic clamps and C-peptide for more efficient estimation of insulin release, may better define the role of these variants in β-cell function and insulin sensitivity. Positive association detected in the *CDC123/CAMKID* variant with fasting insulin levels and impaired β-cell function emphasizes the importance of abnormal β-cell function in the pathogenesis of T2D. These results await further confirmation by an ethnicity-matched larger study. To our knowledge, ours is the first replication study reporting the role of these six new loci with T2D in a population sample from the Indian subcontinent.

ACKNOWLEDGEMENTS

This work was supported by NIH Grant no. KO1 TW006087 from the Fogarty International Center (National Institute of Health, USA) and a seed grant from the University of Oklahoma Health Sciences Center (Oklahoma City, OK, USA). The authors thank all the participants of SDS and are grateful for their contribution in this study.

1 Sladek, R., Rocheleau, G., Rung, J., Dina, C., Shen, L., Serre, D. *et al*. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* **445,** 881–885 (2007).
2 Scott, L. J., Mohlke, K. L., Bonnycastle, L. L., Willer, C. J., Li, Y., Duren, W. L. *et al*. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* **316,** 1341–1345 (2007).
3 Steinthorsdottir, V., Thorleifsson, G., Reynisdottir, I., Benediktsson, R., Jonsdottir, T., Walters, G. B. *et al*. A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat. Genet.* **39,** 770–775 (2007).
4 Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, and Novartis Institutes of BioMedical ResearchSaxena, R., Voight, B. F., Lyssenko, V., Burtt, N. P., de Bakker, P. I. & Chen, H. *et al*. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* **316,** 1331–1336 (2007).
5 Wellcome Trust Case Control Consortium. Genome-wide association study of 14 000 cases of seven common diseases and 3000 shared controls. *Nature* **447,** 661–678 (2007).
6 Zeggini, E., Scott, L. J., Saxena, R., Voight, B. F., Marchini, J. L., Hu, T *et al*. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat. Genet.* **40,** 638–645 (2008).
7 Grant, S. F., Thorleifsson, G., Reynisdottir, I., Benediktsson, R., Manolescu, A., Sainz, J. *et al*. Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat. Genet.* **38,** 320–323 (2006).
8 Cauchi, S., Meyre, D., Dina, C., Choquet, H., Samson, C., Gallina, S. *et al*. Transcription factor TCF7L2 genetic study in the French population: expression in human beta-cells and adipose tissue and strong association with type 2 diabetes. *Diabetes* **55,** 2903–2908 (2006).
9 Damcott, C. M., Pollin, T. I., Reinhart, L. J., Ott, S. H., Shen, H., Silver, K. D. *et al*. Polymorphisms in the transcription factor 7-like 2 (TCF7L2) gene are associated with type 2 diabetes in the Amish: replication and evidence for a role in both insulin secretion and insulin resistance. *Diabetes* **55,** 2654–2659 (2006).
10 Florez, J. C., Jablonski, K. A., Bayley, N., Pollin, T. I., de Bakker, P. I., Shuldiner, A. R. *et al*. TCF7L2 polymorphisms and progression to diabetes in the Diabetes Prevention Program. *N. Engl. J. Med.* **355,** 241–250 (2006).
11 Helgason, A., Palsson, S., Thorleifsson, G., Grant, S. F., Emilsson, V., Gunnarsdottir, S. *et al*. Refining the impact of TCF7L2 gene variants on type 2 diabetes and adaptive evolution. *Nat. Genet.* **39,** 218–225 (2007).
12 Zhang, C., Qi, L., Hunter, D. J., Meigs, J. B., Manson, J. E., van Dam, R. M. *et al*. Variant of transcription factor 7-like 2 (TCF7L2) gene and the risk of type 2 diabetes in large cohorts of U.S. women and men. *Diabetes* **55,** 2645–2648 (2006).
13 Lehman, D. M., Hunt, K. J., Leach, R. J., Hamlington, J., Arya, R., Abboud, HE. *et al*. Haplotypes of transcription factor 7-like 2 (TCF7L2) gene and its upstream region are associated with type 2 diabetes and age of onset in Mexican Americans. *Diabetes* **56,** 389–393 (2007).
14 Chandak, G. R., Janipalli, C. S., Bhaskar, S., Kulkarni, S. R., Mohankrishna, P., Hattersley, A. T. *et al*. Common variants in the TCF7L2 gene are strongly associated with type 2 diabetes mellitus in the Indian population. *Diabetologia* **50,** 63–67 (2007).
15 Sanghera, D. K., Nath, S. K., Ortega, L., Gambarelli, M., Kim-Howard, X., Singh, J. R. *et al*. TCF7L2 polymorphisms are associated with type 2 diabetes in Khatri Sikhs from North India: genetic variation affects lipid levels. *Ann. Hum. Genet.* **72,** 499–509 (2008).
16 Lewis, J. P., Palmer, N. D., Hicks, P. J., Sale, M. M., Langefeld, C. D., Freedman, B. I. *et al*. Association analysis in African Americans of European-derived type 2 diabetes single nucleotide polymorphisms from whole-genome association studies. *Diabetes* **57,** 2220–2225 (2008).
17 Omori, S., Tanaka, Y., Takahashi, A., Hirose, H., Kashiwagi, A., Kaku, K. *et al*. Association of CDKAL1, IGF2BP2, CDKN2A/B, HHEX, SLC30A8, and KCNJ11 with

susceptibility to type 2 diabetes in a Japanese population. *Diabetes* **57,** 791–795 (2008).
18 Horikawa, Y., Miyake, K., Yasuda, K., Enya, M., Hirota, Y., Yamagata, K. *et al.* Replication of genome-wide association studies of type 2 diabetes susceptibility in Japan. *J. Clin. Endocrinol. Metab.* **93,** 3136–3141 (2008).
19 Ng, M. C., Park, K. S., Oh, B., Tam, C. H., Cho, Y. M., Shin, H. D. *et al.* Implication of genetic variants near TCF7L2, SLC30A8, HHEX, CDKAL1, CDKN2A/B, IGF2BP2, and FTO in type 2 diabetes and obesity in 6,719 Asians. *Diabetes* **57,** 2226–2233 (2008).
20 Wu, Y., Li, H., Loos, R. J., Yu, Z., Ye, X., Chen, L. *et al.* Common variants in CDKAL1, CDKN2A/B, IGF2BP2, SLC30A8, and HHEX/IDE genes are associated with type 2 diabetes and impaired fasting glucose in a Chinese Han population. *Diabetes* **57,** 2834–2842 (2008).
21 Sanghera, D. K., Ortega, L., Han, S., Singh, J., Ralhan, S. K., Wander, G. S. *et al.* Impact of nine common type 2 diabetes risk polymorphisms in Asian Indian Sikhs: PPARG2 (Pro12Ala), IGF2BP2, TCF7L2 and FTO variants confer a significant risk. *BMC Med. Genet.* **9,** 59 (2008).
22 Grarup, N., Andersen, G., Krarup, N. T., Albrechtsen, A., Schmitz, O., Jørgensen, T. *et al.* Association testing of novel type 2 diabetes risk alleles in the JAZF1, CDC123/CAMK1D, TSPAN8, THADA, ADAMTS9, and NOTCH2 loci with insulin release, insulin sensitivity, and obesity in a population-based sample of 4,516 glucose-tolerant middle-aged Danes. *Diabetes* **57,** 2534–2540 (2008).
23 Staiger, H., Machicao, F., Kantartzis, K., Schafer, S. A., Kirchhoff, K., Guthoff, M. *et al.* Novel meta-analysis-derived type 2 diabetes risk loci do not determine prediabetic phenotypes. *PLoS ONE* **3,** e3019 (2008).
24 Horikoshi, M., Hara, K., Ito, C., Nagai, R., Froguel, P., Kadowaki, T. *et al.* A genetic variation of the transcription factor 7-like 2 gene is associated with risk of type 2 diabetes in the Japanese population. *Diabetologia* **50,** 747–751 (2007).
25 Jones, D. Scanning for success. *Nat. Rev. Drug. Discov.* **6,** 590–591 (2007).
26 Mathew, C. G. New links to the pathogenesis of Crohn disease provided by genome-wide association scans. *Nat. Rev. Genet.* **9,** 9–14 (2008).
27 Nakagami, T., Qiao, Q., Carstensen, B., Nhr-Hansen, C., Hu, G., Tuomilehto, J. *et al.* Age, body mass index and type 2 diabetes-associations modified by ethnicity. *Diabetologia* **46,** 1063–1070 (2003).
28 Sanghera, D. K., Bhatti, J. S., Bhatti, G. K., Ralhan, S. K., Wander, G. S., Singh, JR. *et al.* The Khatri Sikh Diabetes Study (SDS): study design, methodology, sample collection, and initial results. *Hum. Biol.* **78,** 43–63 (2006).
29 Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care* **26** (Suppl 1), S5–S20 (2003).
30 American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care* **27** (Suppl 1), S5–S10 (2004).
31 Matthews, D. R., Hosker, J. P., Rudenski, A. S., Naylor, B. A., Treacher, D. F., Turner, R. C. *et al.* Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* **28,** 412–419 (1985).
32 Miller, S. A., Dykes, D. D., Polesky, H. F. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* **16,** 1215 (1988).
33 Gonzalez, J. R., Armengol, L., Sole, X., Guino, E., Mercader, J. M., Estivill, X. *et al.* SNPassoc: an R package to perform whole genome association studies. *Bioinformatics* **23,** 644–645 (2007).
34 Purcell, S., Cherny, S. S., Sham, P. C. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* **19** (1), 149–150 (2003).
35 Zeggini, E. A new era for type 2 diabetes genetics. *Diabet. Med.* **24,** 1181–1186 (2007).
36 Frayling, T. M. Genome-wide association studies provide new insights into type 2 diabetes aetiology. *Nat. Rev. Genet.* **8,** 657–662 (2007).

Supplementary Information accompanies the paper on Journal of Human Genetics website (http://www.nature.com/jhg)