ORIGINAL ARTICLE

Genome-Wide Association Study for Type 2 Diabetes in Indians Identifies a New Susceptibility Locus at 2q21

Rubina Tabassum,[1] Ganesh Chauhan,[1] Om Prakash Dwivedi,[1] Anubha Mahajan,[1] Alok Jaiswal,[1] Ismeet Kaur,[1] Khushdeep Bandesh,[1] Tejbir Singh,[1] Benan John Mathai,[1] Yogesh Pandey,[1] Manickam Chidambaram,[2] Amitabh Sharma,[1] Sreenivas Chavali,[1] Shantanu Sengupta,[1] Lakshmi Ramakrishnan,[3] Pradeep Venkatesh,[3] Sanjay K. Aggarwal,[3] Saurabh Ghosh,[4] Dorairaj Prabhakaran,[5] Reddy K. Srinath,[6] Madhukar Saxena,[7] Monisha Banerjee,[7] Sandeep Mathur,[8] Anil Bhansali,[9] Viral N. Shah,[9] Sri Venkata Madhu,[10] Raman K. Marwaha,[11] Analabha Basu,[12] Vinod Scaria,[13] Mark I. McCarthy,[14,15,16] DIAGRAM,* INDICO,* Radha Venkatesan,[2] Viswanathan Mohan,[2] Nikhil Tandon,[3] and Dwaipayan Bharadwaj[1]

Indians undergoing socioeconomic and lifestyle transitions will be maximally affected by epidemic of type 2 diabetes (T2D). We conducted a two-stage genome-wide association study of T2D in 12,535 Indians, a less explored but high-risk group. We identified a new type 2 diabetes–associated locus at 2q21, with the lead signal being rs6723108 (odds ratio 1.31; $P = 3.32 \times 10^{-9}$). Imputation analysis refined the signal to rs998451 (odds ratio 1.56; $P = 6.3 \times 10^{-12}$) within *TMEM163* that encodes a probable vesicular transporter in nerve terminals. *TMEM163* variants also showed association with decreased fasting plasma insulin and homeostatic model assessment of insulin resistance, indicating a plausible effect through impaired insulin secretion. The 2q21 region also harbors *RAB3GAP1* and *ACMSD*; those are involved in neurologic disorders. Forty-nine of 56 previously reported signals showed consistency in direction with similar effect sizes in Indians and previous studies, and 25 of them were also associated ($P < 0.05$). Known loci and the newly identified 2q21 locus altogether explained 7.65% variance in the risk of T2D in Indians. Our study suggests that common susceptibility variants for T2D are largely the same across populations, but also reveals a population-specific locus and provides further insights into genetic architecture and etiology of T2D. *Diabetes* **62:977–986, 2013**

From the [1]Genomics and Molecular Medicine Unit, Council for Scientific and Industrial Research-Institute of Genomics and Integrative Biology, Delhi, India; the [2]Department of Molecular Genetics, Madras Diabetes Research Foundation-Indian Council of Medical Research Advanced Centre for Genomics of Diabetes, Chennai, India; the [3]Department of Endocrinology, All India Institute of Medical Sciences, New Delhi, India; the [4]Human Genetics Unit, Indian Statistical Institute, Kolkata, India; the [5]Centre for Chronic Disease Control, New Delhi, India; the [6]Public Health Foundation of India, New Delhi, India; the [7]Department of Zoology, University of Lucknow, Lucknow, India; the [8]Department of Endocrinology, SMS Medical College and Hospital, Jaipur, India; the [9]Department of Endocrinology, Post Graduate Institute of Medical Education and Research, Sector-12, Chandigarh, India; the [10]Division of Endocrinology, University College of Medical Sciences, Delhi, India; the [11]Department of Endocrinology and Thyroid Research, Institute of Nuclear Medicine and Allied Sciences, Delhi, India; the [12]National Institute of BioMedical Genomics, Kalyani, India; the [13]GN Ramachandran Knowledge Center for Genome Informatics, Council for Scientific and Industrial Research-Institute of Genomics and Integrative Biology, Delhi, India; the [14]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; the [15]Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Churchill Hospital, Oxford, United Kingdom; and the [16]Oxford National Institute for Health Research Biomedical Research Centre, Churchill Hospital, Oxford, United Kingdom.
Corresponding authors: Dwaipayan Bharadwaj, db@igib.res.in, and Nikhil Tandon, nikhil_tandon@hotmail.com.
Received 3 April 2012 and accepted 28 August 2012.
DOI: 10.2337/db12-0406
This article contains Supplementary Data online at http://diabetes.diabetesjournals.org/lookup/suppl/doi:10.2337/db12-0406/-/DC1.
*A complete list of the members of the DIAGRAM and INDICO consortia can be found in the Supplementary Data online.
R.T., G.C., and O.P.D. contributed equally to this study.

Type 2 diabetes (T2D) has developed into a major health problem, responsible for early morbidities and mortality that affects over a billion people worldwide (1). Developing countries such as India will be maximally affected by the T2D epidemic, both in terms of morbidity/mortality and socioeconomic loss in the coming decade (1,2). India, with one-sixth of the world's population, typical risk phenotypes, and rapid socioeconomic transitions, provides an important resource for understanding the pathogenesis of T2D (1,3).

Genome-wide association studies (GWAS) and subsequent meta-analyses have identified >56 susceptibility loci for T2D that collectively explain ~10% of the disease risk (4–8). Although GWAS have greatly improved our understanding of the genetic basis of T2D, most of these studies have been performed in Europeans (9), and the studies involving South Asians are very limited (7,10). Interpopulation differences in allele frequencies and effect sizes have yielded the discovery of new loci in different populations (11). The relatively unexplored and high-risk Indian population provides an opportunity for genetic dissection of T2D and other metabolic disorders (3).

In this study, we performed a two-staged GWAS of T2D involving a total of 12,535 Indians. The study involved a genome scan of 2,465 subjects and replication of top signals in two ethnic groups of India comprising 7,221 Indo-Europeans and 2,849 Dravidian subjects. These two ethnic groups of India are genetically diverse due to different ancestral backgrounds (12); hence, data for these two ethnic groups have been provided separately with the primary focus on Indo-Europeans. Our study identified a new T2D–associated locus on 2q21, with a lead signal at rs6723108 near *TMEM163*. Imputation analysis and genotyping revealed a stronger signal on 2q21 at rs998451 within *TMEM163*. Further analysis suggested a probable mechanism of influence of the *TMEM163* variants on T2D susceptibility through insulin secretion.

RESEARCH DESIGN AND METHODS

The present GWAS was performed in a two-stage study design (Fig. 1), involving a total of 12,535 Indians comprising 6,738 T2D patients and 5,797

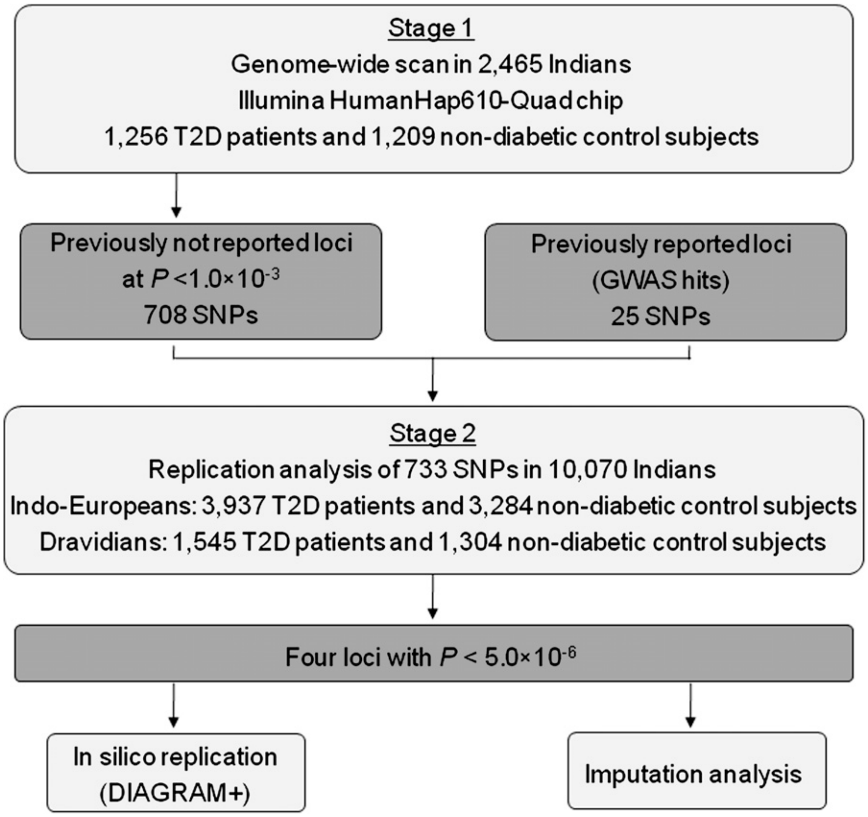

FIG. 1. Study design.

nondiabetic control subjects. We performed a genome scan for 2,465 subjects including 1,256 T2D patients and 1,209 nondiabetic control subjects of Indo-European ethnicity in stage 1. Stage 2 comprised of replication of 708 single nucleotide polymorphisms (SNPs; stage 1, $P < 1.0 \times 10^{-3}$) not reported for association with T2D by previous studies and 25 SNPs of previously reported T2D-associated loci in 7,221 Indo-Europeans (3,937 T2D patients and 3,284 control subjects) and 2,849 Dravidians (1,545 T2D patients and 1,304 control subjects).

All subjects of the current study are part of the Indian Diabetes Consortium (INDICO) (13), a sample repository of two major ethnic groups of India, namely Indo-European and Dravidian (14). Informed written consent was obtained from all of the participants of the study. The study was approved by the ethics committees of the participating institutes (13) and carried out in accordance with the principles of the Helsinki Declaration.

Study participants

Stage 1. Consecutive T2D patients attending the outpatient Department of Endocrinology clinic of All India Institute of Medical Sciences, New Delhi, India, prior to September 2008, were enrolled as case subjects. T2D was diagnosed according to the World Health Organization criteria (i.e., fasting plasma glucose level >126 mg/dL [7.0 mmol/L] or 2-h postglucose load of >200 mg/dL [11.1 mmol/L]) (15). Control group included subjects who met the following criteria: *1*) ≥40 years of age; *2*) HbA_{1c} level ≤6.0%; *3*) fasting plasma glucose level <110 mg/dL (6.1 mmol/L); *4*) no family history of diabetes in first-degree relatives; and *5*) urban dwellers. All of the control subjects were collected from Diabetes Awareness Camps organized in various localities of Delhi and nearby regions as described previously (13).

Stage 2. Stage 2 comprised subjects from two ethnic groups: Indo-European and Dravidian. The ethnicity of the subject was defined primarily on the basis of geographical location of the recruitment along with their birth place and place of origin of both the parents. Indo-European subjects were recruited from northern part of the India that included mainly Delhi and neighboring states, whereas Dravidian subjects were collected from Chennai, located in southern part of India.

Indo-European case subjects included consecutive T2D patients recruited from the outpatient departments of collaborating hospitals: All India Institute of Medical Sciences (New Delhi), Guru Teg Bahadur Hospital (Delhi), and Sawai Man Singh Hospital (Jaipur). T2D patients recruited from the All India Institute of Medical Sciences (New Delhi) for stage 2 included only those subjects who attended the clinic after September 2008. Subjects with self-reported diabetes and/or taking medication for diabetes and newly diagnosed T2D patients were also enrolled from Diabetes Awareness Camps. T2D was diagnosed according to the World Health Organization criteria. Control subjects were recruited based on the criteria as in stage 1, except that the fasting glucose level <100 mg/dL (5.6 mmol/L) was considered as the cutoff for subjects with missing information about HbA_{1c} levels. HbA_{1c} levels were not available for 1,198 subjects in stage 2 of the study.

Subjects in the Dravidian group were recruited from an ongoing epidemiology study, Chennai Urban Rural Epidemiological Study, on a representative population (>20 years of age) of Chennai (16). Subjects with self-reported diabetes taking drug treatment for diabetes or confirmed by oral glucose tolerance test to have a 2-h postload plasma glucose level >200 mg/dL (11.1 mmol/L) were classified as T2D patients. Control subjects included randomly selected subjects with a 2-h plasma glucose level <140 mg/dL (7.8 mmol/L) and fasting plasma glucose <100 mg/dL (5.6 mmol/L). All recruited subjects underwent anthropometric and clinical measurements. Height, weight, and waist and hip circumferences were measured using standard methods as described previously (13,16). Levels of glucose, HbA_{1c}, insulin, C-peptide, total cholesterol, triglycerides, HDL cholesterol, LDL cholesterol, and blood pressure were measured as described previously (13,16). Homeostatic model assessment of insulin resistance (HOMA-IR) was calculated as provided by Matthews et al. (17). Anthropometric and clinical characteristics of the subjects are provided in Supplementary Table 1.

Stage 1 genotyping and analysis. A genome scan was performed using Illumina Human610-Quad BeadChips (Illumina Inc., San Diego, CA). Genotype calls were assigned using the GenCall algorithm as implemented in Illumina GenomeStudio (version 2010.3; Illumina Inc.). Stringent Quality Control (QC) criteria for filtering SNPs and samples for analyses were applied separately in T2D case and control subjects (Supplementary Fig. 1). Samples with low call rate (<95%), sex discrepancies, and cryptic relatedness (pi_hat >0.1875) were removed. Cryptic relatedness was determined by identity-by-descent analysis using PLINK v1.07 (http://pngu.mgh.harvard.edu/~purcell/plink) (18). SNPs with minor allele frequency (MAF) <0.01, either in case or control subjects, were excluded (Supplementary Fig. 1). Further, from the SNPs with a MAF 0.01–0.05, those with a call rate <99% and Hardy-Weinberg equilibrium (HWE) of $P < 5.7 \times 10^{-4}$ were excluded, whereas from SNPs with a MAF >0.05, those with call rate <95% and HWE of $P < 1 \times 10^{-7}$ were removed.

Principal component analysis was performed in stage 1 samples (without inclusion of HapMap samples) to assess population structure and identify outliers with 116,631 independent pruned SNPs using SMARTPCA as implemented in EIGENSOFT version 3.0 (Supplementary Fig. 2). Pruning of SNPs was carried out with successfully genotyped autosomal SNPs using the –indep-pairwise option of PLINK v1.07, with a window size of 50 SNPs, step size of 5, and $r^2 < 0.2$. We also removed SNPs within regions of known high linkage disequilibrium (LD) (19). The first two principal components (PC1 and PC2) had maximum contribution with eigenvalues of 8.41 and 4.33, respectively. A total of 89 samples were identified as outliers that were removed using the outlier removal option of SMARTPCA (σ threshold of 6 with 5 iterations along first 10 principal components). Principal component analysis was also performed using our samples with reference samples from HapMap Phase III populations CEU (Utah residents with Northern and Western European ancestry from the Centre d'Etude du Polymorphisme Humain collection), GIH (Gujarati Indians in Houston, TX), YRI (Yoruban in Ibadan, Nigeria), JPT (Japanese in Tokyo, Japan), and CHB (Han Chinese in Beijing, China) using first two eigenvectors.

Association of QC-passed SNPs with T2D was tested by logistic regression analysis under a log-additive model adjusting for age, sex, and potential population stratification using PLINK v1.07. The first two principal components (PC1 and PC2) that accounted for the maximum variance were used as covariates to adjust for residual population stratification. Deviation between observed *P* values for association and the expected *P* values under the null hypothesis was investigated by constructing quantile–quantile plot. The $-\log10$ of *P* values observed for the association of SNPs under the additive model adjusted for age, sex, and first two principal components was plotted against the theoretical $-\log10$ *P* values expected under the null hypothesis using STATA version 10.1 (Stata Corporation, College Station, TX). The genomic inflation factor (λ) was calculated from the average χ^2 statistics using PLINK v1.07.

Stage 2 genotyping and analysis. Genotyping of the selected SNPs in stage 2 was performed using the Illumina Golden Gate assay following the manufacturer's instructions (http://support.illumina.com/documents/MyIllumina/2bbe5885-34be-4e2e-b655-4848b7d3e3d7/GoldenGate_Genotyping_Assay_Guide_15004065_B.pdf; Illumina Inc.). SNPs with a genotype confidence score <0.25, call frequency <0.90, GenTran score <0.60, cluster separation score <0.40, minor allele frequency <0.01, and HWE of $P < 1.0 \times 10^{-5}$ were excluded. All of the QC criteria for SNPs were applied separately in case and control subjects, and only those SNPs that qualified QC both in case and control subjects were taken forward for analysis. Samples with a <0.90 call rate were also removed. A total of 256 samples (~3%) were genotyped in duplicates, which showed >99% concordance in genotype calls. Association of SNP T2D was assessed separately in Indo-Europeans and Dravidians under the additive model using logistic regression analysis adjusting for age and sex.

Results of stages 1 and 2 were combined by inverse variance method under fixed effect model using METAL (http://www.sph.umich.edu/csg/abecasis/Metal/) (20). Association of the top four loci with quantitative clinical traits including measures of adiposity, dysglycemia, and dyslipidemia was tested using linear regression analysis under the additive model adjusted for age, sex, and BMI as appropriate. Only nondiabetic control subjects from stages 1 and 2 were considered for association analysis with clinical traits. Pairwise LD between SNPs at the associated loci was determined using Haploview 4.0 software (21). Haplotype association analyses were adjusted for age and sex and assessed at 10,000 permutations using PLINK v1.07.

Imputation analysis. SNPs within a 2-Mb region around four lead signals were imputed in final post-QC genotype data using a 1000 Genomes Phase 1 (interim) panel that consists of combined phased haplotypes of 1,094 individuals from Africa, Asia, Europe, and America (build 37) using IMPUTE v2.1.2. (https://mathgen.stats.ox.ac.uk/impute/impute_v2.html). Association was tested using SNPTEST v2.2.0 (https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html) adjusting for age, sex, PC1, and PC2. Association analysis was done for SNPs with a call rate >99%, threshold posterior probability of 0.90, and proper_info >0.50.

Genotyping and analysis of the of the top imputed SNP. The SNP rs998451 that showed most significant association after imputation was also confirmed by genotyping. Genotyping was performed in a total of 5,262 subjects (2,538 T2D patients and 2,724 nondiabetic control subjects) that included 2,036 subjects (1,087 T2D patients and 949 nondiabetic control subjects) of stage 1 and 3,226 subjects (1,451 T2D patients and 1,775 nondiabetic control subjects) of stage 2. Genotyping was carried out by single-nucleotide extension of SNP-specific probes using a SNaPshot ddNTP primer extension kit. The extended probes were then electrophoresed on an ABI Prism 3100 genetic analyzer (Applied Biosystems, Foster City, CA), and genotypes were determined using GeneMapper Software v4.0. Call rate of the SNP was >99%, and genotypes did not deviate from HWE ($P > 0.05$ in control subjects) in stage 1 or stage 2 study populations. Association analyses were performed as mentioned above for stage 1 and stage 2.

In silico replication. Associations of four top loci obtained in meta-analysis of Indo-European samples were tested in Europeans by in silico replication using GWAS phase data from the DIAGRAM+ study. DIAGRAM+ comprises GWAS phase data of 8,130 individuals with T2D and 38,987 control subjects of European descent with an effective sample size of 22,570 subjects (22). Meta-analysis by combining summary statistics of results across the studies was performed by inverse variance method under fixed-effect model using METAL.

Calculation of phenotypic variance explained by selective SNPs. The GCTA tool (23) was used for calculating the variance explained by combination of 56 SNPs from the 56 known loci of T2D and the SNP rs6723108 near *TMEM163* identified by the current study. The analysis was performed assuming a disease prevalence of 10% and adjusting for sex, age, PC1, and PC2. Assuming a disease prevalence of 10% would mean the rest of the subjects (90%) are control subjects and should represent the remaining liability distribution, which is not true in the current study, as control subjects were defined based on certain other criteria in addition to being unaffected with T2D. Thus, results of this analysis are approximate.

Power calculation. Statistical power of the study was calculated for a range of allele frequencies from 0.01 to 0.50, with odds ratios (ORs) ranging from 1.10 to 1.40, using Quanto software (http://hydra.usc.edu/gxe/), assuming a log-additive model of inheritance and 10% prevalence of disease at $\alpha = 0.05$ (Supplementary Fig. 3).

RESULTS

After stringent QC, a successfully genotyped dataset for 1,101 T2D patients and 1,027 control subjects at 536,420 autosomal SNPs was obtained in stage 1 (Supplementary Fig. 1). The quantile–quantile plot showed good agreement with null distribution (Supplementary Fig. 4). The genomic control inflation factor (λ) was 1.09. Consistent with previous reports, *TCF7L2* was identified as the strongest signal in Indians as well (OR 1.56; $P = 5.01 \times 10^{-11}$) (Fig. 2). In stage 2, 708 SNPs (stage 1, $P < 1.0 \times 10^{-3}$) not reported for association with T2D by previous studies and 25 SNPs of previously reported T2D-associated loci were genotyped. QC filtering yielded data for 653 SNPs in 6,552 Indo-European subjects and 635 SNPs in 2,256 Dravidians. We tested association of SNPs with T2D separately in Indo-Europeans and Dravidians and combined the results of two stages for Indo-European subjects by meta-analysis using the fixed-effect inverse variance method (20).

Combined analysis of Indo-European samples revealed strongest signal at rs7903146 of *TCF7L2* (OR = 1.51 [95% CI 1.42–1.62; $P = 1.2 \times 10^{-35}$) (Table 1). We identified an unreported locus on 2q21 as a susceptibility region for T2D at genome-wide significance ($P < 5.0 \times 10^{-8}$). On 2q21, two SNPs: rs6723108 (OR 1.31 [95% CI 1.20–1.44]; $P = 3.3 \times 10^{-9}$) and rs7570971 (OR 1.25 [95% CI 1.15–1.35]; $P = 2.0 \times 10^{-8}$) reached genome-wide significance (Table 1), whereas rs621341 (OR 1.25 [95% CI 1.14–1.36]; $P = 6.9 \times 10^{-7}$) and rs624817 (OR 1.17 [95% CI 1.08–1.26]; $P = 3.4 \times 10^{-5}$) showed suggestive association (Supplementary Table 2). These four SNPs were in moderate LD with each other ($r^2 = 0.33$–0.54) (Supplementary Fig. 5). Conditional analysis led to substantial reduction in association of other SNPs in the region when conditioned for rs6723108 (Fig. 3 and Supplementary Table 3). Haplotype analysis suggested a stronger signal within 2q21 region (OR 0.69; $P = 4.0 \times 10^{-11}$ for TGC haplotype encompassing nonrisk alleles of SNPs rs621341, rs6723108, and rs7570971, respectively) (Supplementary Table 4). Hence, we performed in silico fine mapping by imputation within a 2-Mb region of rs6723108 in stage 1 samples (Fig. 3) using the 1000 Genomes reference panel. The concordance rate of genotypes pre- and postmasking 10% of the genotyped SNPs was >98%. Imputation revealed a stronger signal at rs998451 (OR 1.61; $P = 9.2 \times 10^{-7}$) that lies in intron 2 of *TMEM163* ($r^2 = 0.84$ with rs6723108) than any directly

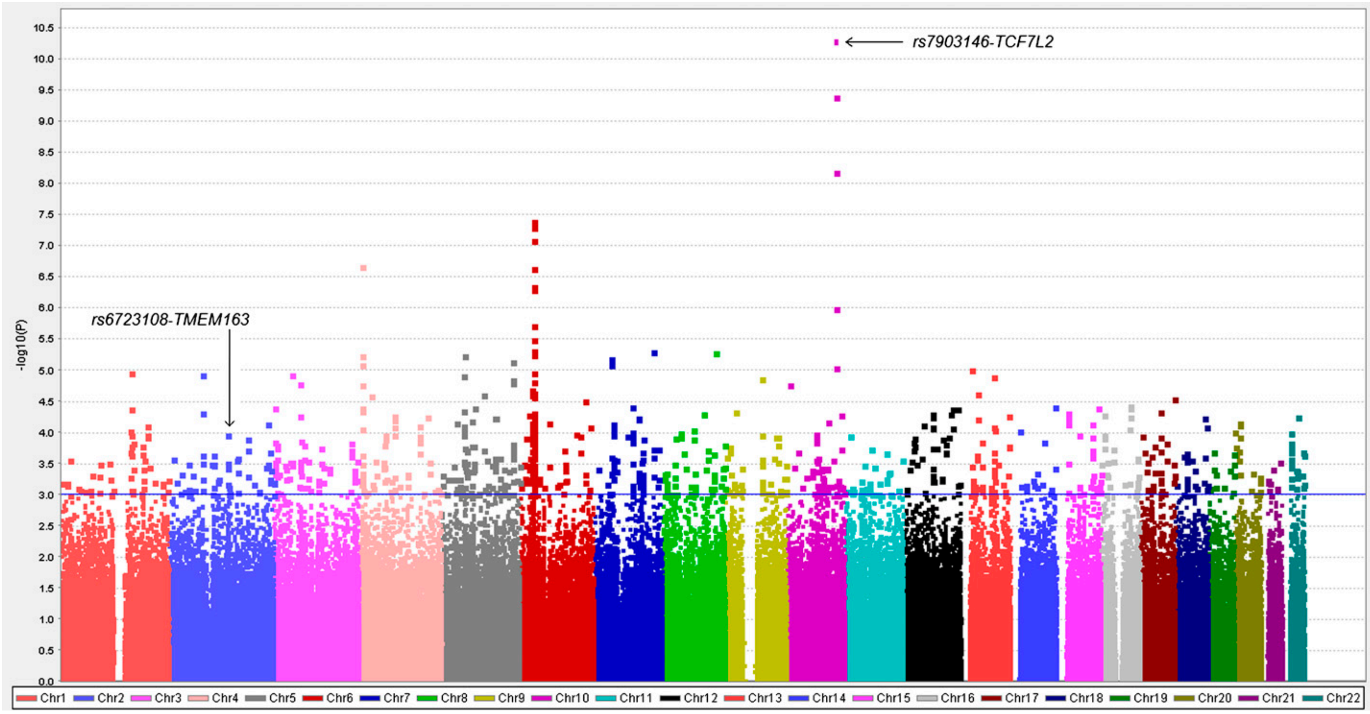

FIG. 2. Manhattan plot of the association *P* values for T2D. The −log10 *P* values for association of directly genotyped SNPs are plotted as a function of genomic position (National Center for Biotechnology Information Build 37). The *P* values were determined using logistic regression adjusted for age, sex, PC1, and PC2 in stage 1 analysis. Each chromosome (Chr) has been represented with a unique color.

genotyped SNP of phase 1 (Supplementary Table 5). The association results of imputation were also confirmed by genotyping rs998451 in a subset of samples of stage 1 (N = 2,036) and stage 2 (N = 3,226). The concordance rate of imputed genotypes and directly genotyped genotypes was 94%, suggesting successful imputation using the 1000 Genomes panel as a reference panel. The variant rs998451 was associated in both stage 1 (OR 1.43; $P = 3.0 \times 10^{-4}$) and stage 2 (OR 1.66; $P = 2.7 \times 10^{-9}$) and crossed genome-wide significance (OR 1.56; $P = 6.3 \times 10^{-12}$) after meta-analysis (Table 1).

Replication analysis of 708 previously unreported SNPs in the independent study population of Dravidians showed association of 19 of them ($P < 0.05$) with consistent direction with Indo-Europeans (Supplementary Table 6). Among the top four loci in Indo-Europeans ($P < 5.0 \times 10^{-6}$; Table 1), only rs1929752 (*FLJ35379*) was nominally associated, but with opposite effect. SNPs rs10461617 (*MAP3K1*) and rs11165354 (*TGFBR3*) showed consistency in direction but were not associated. The observed discrepancies in association results might be due to differences in allele frequencies or effect sizes of these variants between Indo-Europeans and Dravidians (Supplementary Table 6); however, a larger study population of Dravidian origin would be required to state anything conclusively. We carried out in silico replication of top four loci—*TMEM163*, *MAP3K1*, *TGFBR31*, and *FLJ35379* in DIAGRAM+—and performed meta-analysis (Supplementary Fig. 6). The SNP rs6723108 (*TMEM163*) had the same direction of effect in DIAGRAM+; however, the association was not statistically significant. At the *MAP3K1* locus (rs10461617), suggestive association with consistent direction was observed in DIAGRAM+. Nominal association of rs11165354 of *TGFBR3* in DIAGRAM+ was found (P = 0.036), but there was inconsistency in direction of effect. The observed differences in association results might be due to differences in allele frequencies and LD structures in Europeans and Indians.

Further, we investigated association of 56 T2D loci identified by previous GWAS and assessed heterogeneity in effect of the present and previous studies (Tables 2 and 3). A total of 49 loci showed consistency in direction of association with similar effect sizes with previous reports (P-heterogeneity >0.009, P value corrected for multiple testing), and association for 25 of these loci was also observed ($P < 0.05$). The seven loci *HNF1B*, *DGKB-TMEM195*, *SRR*, *KCNQ1*, *GLIS3*, *GCC1-PAX4*, and *PSMD6* were not consistent in direction, two of which (*HNF1B* and *SRR*) also showed significant heterogeneity of effect even when our study was sufficiently powered (>98%) to detect the associations. In Dravidians, association analysis of 22 known loci that were genotyped in stage 2 (Supplementary Table 7) showed that 17 loci have consistency in direction and similar effect size, with 7 of these attaining statistical significance ($P < 0.05$).

To explore the possible functional mechanisms of the identified loci, we performed association analysis with quantitative traits. The risk allele of rs6723108 was associated with decrease in fasting plasma insulin (β= −5.71 pmol/L; $P = 3.7 \times 10^{-3}$) and HOMA-IR (β= −0.21; $P = 7.4 \times 10^{-3}$) in Indo-Europeans (Supplementary Table 8). Similarly, rs998451 was also associated with decrease in fasting plasma insulin (β= −5.54 pmol/L; P = 0.02) and HOMA-IR (β= −0.20; P = 0.04). This suggests that *TMEM163* variants might modulate T2D susceptibility by affecting insulin secretion. Previous studies have shown that at this 2q21 region, rs6723108 and a variant in *RAB3GAP1* (rs6730157) are associated with waist circumference (24) ($P = 6.2 \times 10^{-5}$) and total cholesterol ($P = 8.5 \times 10^{-8}$) (25), respectively. However, variants from

TABLE 1
SNPs showing association with T2D at $P < 5.0 \times 10^{-6}$

					Stage 1				Stage 2				Meta-analysis		
SNP	Gene	Chromosome	Position	RA/ OA	RAF case subjects	RAF control subjects	OR (95% CI)	*P* value	RAF case subjects	RAF control subjects	OR (95% CI)	*P* value	OR (95% CI)	*P* value	Het-*P*
rs7903146	*TCF7L2*	10	114758349	T/C	0.39	0.30	1.56 (1.36–1.77)	5.01×10^{-11}	0.38	0.29	1.50 (1.39–1.62)	3.53×10^{-26}	1.51 (1.42–1.62)	1.2×10^{-35}	0.65
rs6723108	*TMEM163*	2	135479980	T/G	0.89	0.85	1.43 (1.19–1.71)	1.10×10^{-4}	0.89	0.86	1.28 (1.15–1.42)	4.31×10^{-6}	1.31 (1.20–1.44)	3.32×10^{-9}	0.29
rs998451*	*TMEM163*	2	135429288	G/A	0.91	0.86	1.43 (1.18–1.74)	3.00×10^{-4}	0.92	0.86	1.66 (1.41–1.96)	2.70×10^{-9}	1.56 (1.38–1.77)	6.30×10^{-12}	0.26
rs10461617	*MAP3K1*	5	56104308	A/G	0.25	0.20	1.30 (1.13–1.51)	4.01×10^{-4}	0.24	0.21	1.17 (1.07–1.27)	3.66×10^{-4}	1.20 (1.11–1.29)	1.23×10^{-6}	0.20
rs11165554	*TGFBR3*	1	92194322	A/C	0.82	0.78	1.26 (1.08–1.47)	2.59×10^{-3}	0.81	0.78	1.18 (1.08–1.28)	1.53×10^{-4}	1.20 (1.11–1.29)	1.8×10^{-6}	0.45
rs1929752	*FLJ35379*	13	76812217	G/T	0.75	0.70	1.31 (1.15–1.50)	8.84×10^{-5}	0.72	0.70	1.13 (1.05–1.22)	1.63×10^{-3}	1.17 (1.10–1.25)	3.1×10^{-6}	0.06

All results presented are for subjects of Indo-European ethnicity. Chromosomal positions of SNPs are based on National Center for Biotechnology Information genome build 37. Alleles presented are indexed to the positive strand. ORs were calculated with respect to the risk alleles. Het-*P*, *P* value for heterogeneity in effect sizes between stage 1 and stage 2; OA, other allele; RA, risk allele; RAF, risk allele frequency in control subjects. *The SNP was identified through imputation and confirmed by genotyping. Association results presented for rs998451 were obtained from genotyped data in 2,036 subjects from stage 1 and 3,226 subjects from stage 2.

this locus were not associated with these traits in our study, which could be due to the small effect size of these variants for these traits in the current study population or lower power of the current study to detect them.

DISCUSSION

GWAS has enormously contributed to the identification of susceptibility genes for T2D and many other complex disorders. The discovery of new susceptibility loci in GWAS of different ethnic groups emphasizes the need for conducting more GWAS in populations of diverse ethnic groups. Our study also shows the potential of identifying new signals through GWAS in population of different ethnicity. In this study, we performed a GWAS in Indians and identified a new signal for T2D on chromosome 2q21.

The strongest signal on newly identified locus 2q21 mapped to *TMEM163*. The association of *TMEM163* variants (rs6723108 and rs998451) with reduced fasting plasma insulin levels and HOMA-IR indicates that it might modulate susceptibility to T2D by affecting insulin secretion. Previously, rs6723108 has also been suggested to be associated with waist circumference. These data indicate the biological relevance of the identified locus with T2D. *TMEM163* encodes a synaptic vesicle membrane protein with six putative transmembrane helices that may function as a vesicular transporter. *TMEM163* is expressed in select brain regions and contained in subpopulations of nerve terminals (26,27). This region also harbors *RAB3GAP1* and *ACMSD*, which are shown to be involved in neurologic disorders (28). *RAB3GAP1* encodes a catalytic subunit of a Rab GTPase-activating protein involved in regulated exocytosis of neurotransmitters and hormones. Mutations in *RAB3GAP1* are associated with neurodevelopmental abnormalities known as Warburg micro syndrome (OMIM 600118). *ACMSD* codes for an enzyme involved in picolinic and quinolinic acid homeostasis and NAD biosynthesis (29,30). ACMSD activity and expression have been shown to be elevated in streptozotocin-induced diabetic rats and suppressed by insulin injection and is suggested as a potential therapeutic avenue for diabetes treatment (30). Thus, our results suggest the involvement of neurologic component in the etiology of T2D, which needs to be confirmed through further investigations.

The signal on 2q21 observed in Indo-Europeans was not detected in Dravidians and Europeans from DIAGRAM+. The observed differences in association results might be due to differences in allele frequencies and LD structures or statistical power to detect the association. In Dravidians, the minor allele frequency of the index variant rs6723108 is very low (0.03) and differs significantly from that of Indo-Europeans, and hence, to capture the association of this variant, if any, a larger sample size is required. Thus, the statistical power of the Dravidian study population might not be sufficient to detect the observed association in Indo-Europeans (Supplementary Fig. 3). Among Europeans from DIAGRAM+ study, direction of effect was same as found in Indo-Europeans, but the association could not reach statistical significance.

MAP3K1 was identified as a promising candidate for T2D, which has not been reported earlier, but needs further investigation. *MAP3K1* codes for serine/threonine protein kinase and has been shown to have an essential role in stress-induced β-cell death (31) and inhibits

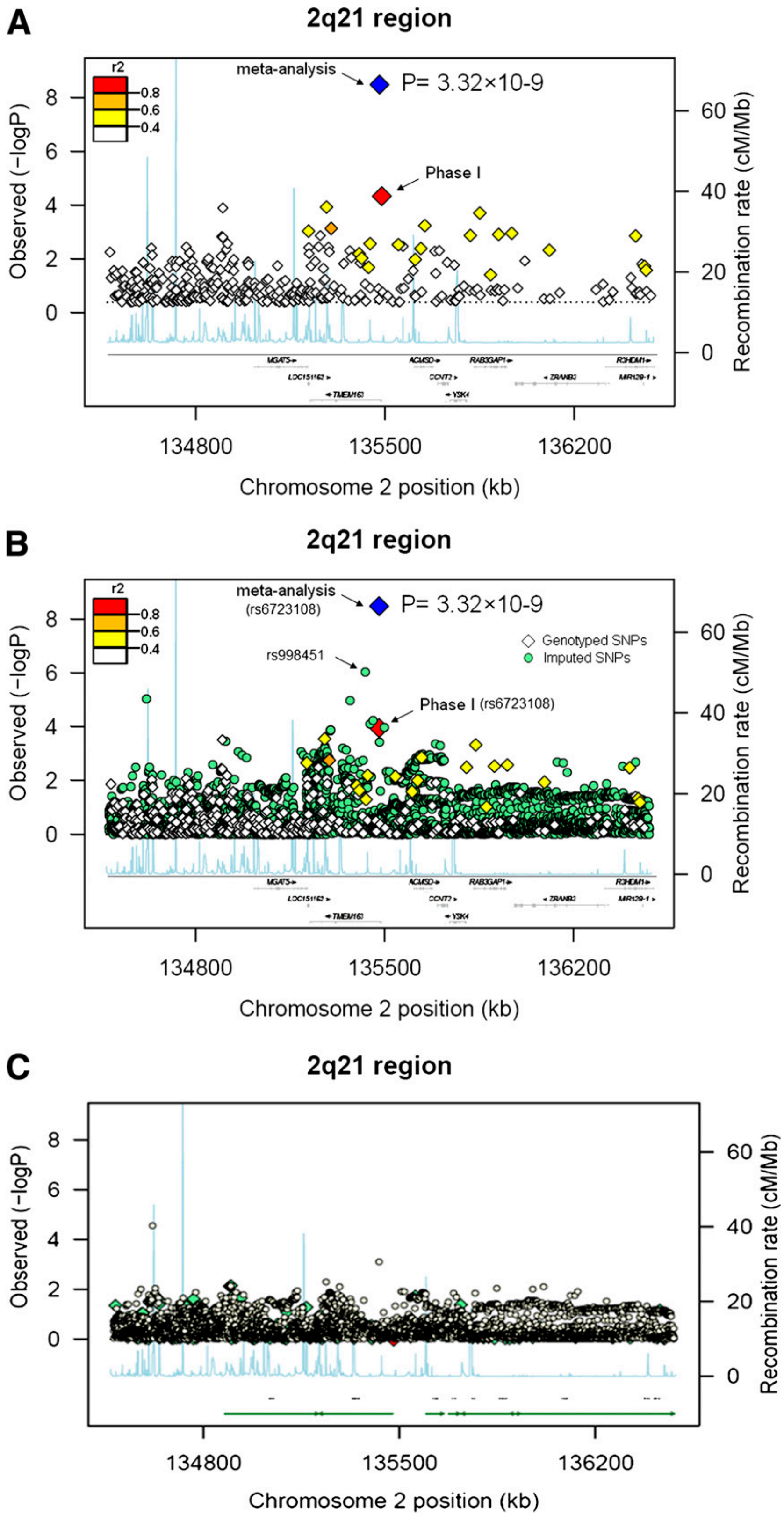

FIG. 3. Regional plot of 2q21. Regional association plot depicting *P* value distribution of all genotyped and imputed SNPs in a 2-Mb region flanking the SNP rs6723108 (1 Mb each side) at chromosome 2q21. *A*: Stage 1 genotyped SNPs. *B*: Stage 1 genotyped and imputed SNPs. *C*: Stage 1 genotyped and imputed SNPs conditioned for rs6723108. Strength of pairwise LD between all SNPs and rs6723108 has been color-coded and, based on r^2, computed from present study data (stage 1). Recombination rates were estimated from HapMap2 combined panel.

TABLE 2
Association of previously reported SNPs with T2D (genotyped in stage 1 and stage 2)

Index SNP	Nearest gene	Source (first author, year)	Index/proxy SNP in present study	r^2	Risk allele¶	Source RAF	Present study RAF	Source OR (95% CI)	Source P value	Stage 1 OR (95% CI)	Stage 1 P value	Stage 2 OR (95% CI)	Stage 2 P value	Meta-analysis OR (95% CI)	Meta-analysis P value	Dir.*	Het. P§	Power‖
rs1801282	*PPARG*	Saxena et al., 2007 (34)	rs6802898	1.00#	C	0.86	0.85	1.14 (1.08–1.20)	1.7×10^{-6}	1.25 (1.05–1.49)	0.01	1.16 (1.05–1.29)	3.8×10^{-3}	1.18 (1.08–1.29)	0.0002	+	0.51	0.85
rs5219	*KCNJ11*	Saxena et al., 2007 (34)	rs5219		T	0.47	0.36	1.14 (1.10–1.19)	6.7×10^{-11}	1.22 (1.07–1.38)	2.7×10^{-3}	1.07 (1.00–1.15)	0.06	1.11 (1.04–1.18)	0.002	+	0.33	0.99
rs7903146	*TCF7L2*	Saxena et al., 2007 (34)	rs7903146		T	0.26	0.30	1.37 (1.31–1.43)	1.0×10^{-48}	1.56 (1.36–1.77)	5.0×10^{-11}	1.50 (1.39–1.62)	3.5×10^{-26}	1.51 (1.42–1.62)	1.2×10^{-35}	+	0.02	0.99
rs4402960	*IGF2BP2*	Saxena et al., 2007 (34)	rs4402960		T	0.29	0.41	1.14 (1.11–1.18)	8.9×10^{-36}	1.20 (1.06–1.36)	3.8×10^{-3}	1.13 (1.05–1.21)	7.3×10^{-4}	1.15 (1.08–1.22)	1.2×10^{-5}	+	0.80	0.99
rs10946398	*CDKAL1*	Saxena et al., 2007 (34)	rs4712523	0.91#	C	0.31	0.30	1.12 (1.08–1.16)	4.1×10^{-11}	1.12 (0.98–1.27)	0.10	1.14 (1.05–1.23)	9.8×10^{-4}	1.13 (1.06–1.21)	0.0002	+	0.82	0.93
rs13266634	*SLC30A8*	Saxena et al., 2007 (34)	rs13266634†		C	0.65	0.74	1.12 (1.07–1.16)	5.3×10^{-8}	1.21 (1.05–1.39)	0.01	NA	NA	NA	NA	+	0.31	0.36
rs10811661	*CDKN2A*	Saxena et al., 2007 (34)	rs2383208	0.87#	T	0.83	0.85	1.20 (1.14–1.25)	7.8×10^{-25}	1.47 (1.22–1.76)	4.6×10^{-5}	1.17 (1.05–1.29)	2.8×10^{-3}	1.23 (1.13–1.34)	3.3×10^{-6}	+	0.64	0.99
rs1111875	*HHEX*	Saxena et al., 2007 (34)	rs1111875		C	0.53	0.42	1.13 (1.09–1.17)	5.7×10^{-10}	1.24 (1.09–1.40)	7.2×10^{-4}	1.07 (1.00–1.15)	0.06	1.11 (1.04–1.18)	0.0009	+	0.62	0.98
rs8050136	*FTO*	Scott et al., 2007 (35)	rs8050136		A	0.38	0.34	1.17 (1.12–1.22)	1.3×10^{-12}	1.09 (0.96–1.24)	0.18	1.18 (1.10–1.27)	7.0×10^{-6}	1.16 (1.09–1.24)	6.1×10^{-6}	+	0.83	0.99
rs10010131	*WFS1*	Sandhu et al., 2007 (36)	rs10012946	1.00#	G	0.60	0.70	1.11 (1.08–1.16)	1.4×10^{-7}	1.15 (1.01–1.31)	0.04	1.14 (1.06–1.24)	9.0×10^{-4}	1.14 (1.07–1.22)	0.0001	+	0.47	0.87
rs4430796	*HNF1B*	Gudmundsson et al., 2007 (37)	rs4430796		A	0.49	0.65	1.22 (1.15–1.30)	1.4×10^{-11}	0.94 (0.83–1.08)	0.38	0.95 (0.88–1.02)	0.20	0.95 (0.89–1.01)	0.11	−	1.1×10^{-8}	0.99
rs10923931	*NOTCH2*	Zeggini et al., 2008 (38)	rs2641348	0.90#	T	0.11	0.21	1.13 (1.08–1.17)	4.1×10^{-8}	1.10 (0.95–1.28)	0.19	1.12 (1.03–1.22)	7.5×10^{-3}	1.12 (1.04–1.20)	0.003	+	0.84	0.98
rs7578597	*THADA*	Zeggini et al., 2008 (38)	rs7578597		T	0.90	0.87	1.15 (1.10–1.20)	1.1×10^{-9}	1.01 (0.84–1.20)	0.93	1.05 (0.95–1.16)	0.35	1.04 (0.95–1.14)	0.38	+	0.05	0.86
rs4607103	*ADAMTS9*	Zeggini et al., 2008 (38)	rs4411878†	0.89#	C	0.12	0.52	1.09 (1.06–1.12)	1.2×10^{-8}	1.09 (0.97–1.23)	0.15	NA	NA	NA	NA	+	1.00	0.29
rs864745	*JAZF1*	Zeggini et al., 2008 (38)	rs1635852	0.94#	T	0.501	0.69	1.10 (1.07–1.13)	5.0×10^{-14}	1.12 (0.98–1.28)	0.10	1.12 (1.03–1.21)	5.4×10^{-3}	1.12 (1.04–1.19)	0.0009	+	0.63	0.82
rs12779790	*CDC123−CAMK1D*	Zeggini et al., 2008 (38)	rs11257622	0.69€	G	0.18	0.11	1.11 (1.07–1.14)	1.2×10^{-10}	1.15 (0.96–1.39)	0.14	1.09 (1.00–1.20)	0.06	1.10 (1.02–1.20)	0.02	+	0.85	0.58
rs7961581	*TSPAN8−LGR5*	Zeggini et al., 2008 (38)	rs1353362	0.95#	C	0.27	0.36	1.09 (1.06–1.12)	1.1×10^{-9}	1.04 (0.92–1.18)	0.56	1.08 (1.00–1.16)	0.04	1.07 (1.00–1.14)	0.04	+	0.60	0.78
rs2237892	*KCNQ1*	Yasuda et al., 2008 (39)	rs2237892		C	0.45	0.98	1.40 (1.34–1.47)	1.7×10^{-42}	1.45 (0.88–2.39)	0.14	1.57 (1.19–2.07)	1.6×10^{-3}	1.54 (1.21–1.96)	0.0005	+	0.45	0.82
rs10830963	*MTNR1B*	Prokopenko et al., 2009 (40)	rs1387153	0.56€	G	0.31	0.37	1.09 (1.05–1.12)	3.3×10^{-7}	1.12 (0.99–1.27)	0.06	1.10 (1.03–1.18)	6.3×10^{-3}	1.11 (1.04–1.18)	0.001	+	0.80	0.78
rs2943641	*IRS1*	Rung et al., 2009 (41)	rs2943641		C	0.63	0.76	1.19 (1.13–1.25)	9.3×10^{-12}	1.15 (1.00–1.33)	0.06	1.06 (0.98–1.15)	0.14	1.08 (1.01–1.16)	0.03	+	0.03	0.99
rs340874	*PROX1*	Dupuis et al., 2010 (42)	rs340874		C	0.42	0.55	1.07 (1.05–1.09)	7.2×10^{-10}	1.03 (0.91–1.16)	0.67	1.05 (0.98–1.12)	0.20	1.04 (0.98–1.11)	0.15	+	0.56	0.59
rs780094	*GCKR*	Dupuis et al., 2010 (42)	rs780094		C	0.62	0.77	1.06 (1.04–1.08)	1.3×10^{-9}	1.03 (0.89–1.19)	0.71	1.12 (1.03–1.22)	5.6×10^{-3}	1.10 (1.02–1.18)	0.01	+	0.33	0.36
rs11708067	*ADCY5*	Dupuis et al., 2010 (42)	rs6798189†	0.94#	A	0.78	0.75	1.12 (1.09–1.15)	9.9×10^{-21}	1.17 (1.02–1.35)	0.03	NA	NA	NA	NA	+	0.56	0.35
rs2191349	*DGKB−TMEM195*	Dupuis et al., 2010 (42)	rs10244051	1.00€	T	0.52	0.37	1.06 (1.04–1.08)	1.1×10^{-8}	0.95 (0.84–1.09)	0.45	0.99 (0.92–1.06)	0.75	0.98 (0.92–1.04)	0.55	−	0.02	0.46
rs4607517	*GCK*	Dupuis et al., 2010 (42)	rs4607517		A	0.16	0.13	1.07 (1.05–1.10)	5.0×10^{-8}	1.02 (0.85–1.21)	0.87	1.08 (0.97–1.20)	0.15	1.06 (0.97–1.16)	0.17	+	0.84	0.33

Effect sizes of the source study and current study (Indo-Europeans) are presented with respect to the risk allele in the source study. Dir., direction; Het., heterogeneity; NA, not available; RAF, risk allele frequency in control. ¶Risk allele as in the source study (indexed to the positive strand of National Center for Biotechnology Information build 37). *Direction was + if there was concordance between the source and present study and − if there was discordance. §P for difference in OR of the source and present study. ‖Power of the current study based on its sample size and allele frequency to detect the previously reported effect sizes under a log-additive model assuming 10% prevalence of T2D. †SNPs were attempted for genotyping in stage 2 but failed in assay, hence data for only stage 1 is presented. #r^2 values from HapMap data for GIH. €r^2 values from the 1000 Genomes project for CEU.

TABLE 3
Association of previously reported SNPs with T2D (genotyped in stage 1 only)

Index SNP	Nearest gene	Source (first author, year)	Index/proxy SNP in present study	r^2	Risk allele¶	Source RAF	Present Study RAF	Source		Present study, stage 1				
								OR (95% CI)	P value	OR (95% CI)	P value	Dir.*	Het. P§	Power‖
rs17584499	*PTPRD*	Tsai et al., 2010 (43)	rs17584499		T	0.07	0.24	1.57 (1.36–1.82)	8.5×10^{-10}	1.14 (0.99–1.32)	0.07	+	0.002	0.99
rs391300	*SRR*	Tsai et al., 2010 (43)	rs391300		C	0.62	0.49	1.28 (1.18–1.39)	3.1×10^{-9}	0.99 (0.88–1.12)	0.99	−	0.0009	0.98
rs7593730	*RBMS−ITGB6(2q24)*	Qi et al., 2010 (44)	rs4077463	1.00€	C	0.78	0.79	1.11 (1.08–1.16)	3.7×10^{-28}	1.02 (0.88–1.18)	0.81	+	0.28	0.28
rs243021	*BCL11A*	Voight et al., 2010 (22)	rs243021		A	0.46	0.49	1.08 (1.06–1.10)	2.9×10^{-15}	1.10 (0.98–1.25)	0.10	+	0.77	0.24
rs4457053	*ZBED3*	Voight et al., 2010 (22)	rs7708285	0.93€	G	0.26	0.19	1.08 (1.06–1.11)	2.8×10^{-12}	1.10 (0.94–1.29)	0.22	+	0.82	0.17
rs972283	*KLF14*	Voight et al., 2010 (22)	rs13234407	0.97€	G	0.55	0.59	1.07 (1.05–1.10)	2.2×10^{-10}	1.03 (0.91–1.17)	0.66	+	0.56	0.19
rs896854	*TP53INP1*	Voight et al., 2010 (22)	rs896854		T	0.48	0.41	1.06 (1.04–1.09)	9.9×10^{-10}	1.06 (0.94–1.20)	0.36	+	1.00	0.15
rs13292136	*CHCHD9*	Voight et al., 2010 (22)	rs4295736	1.00€	C	0.93	0.88	1.11 (1.07–1.15)	2.8×10^{-8}	1.04 (0.86–1.25)	0.70	+	0.50	0.19
rs231362	*KCNQ1*	Voight et al., 2010 (22)	rs231358	0.53€	G	0.52	0.16	1.08 (1.06–1.10)	2.8×10^{-13}	0.93 (0.78–1.10)	0.36	−	0.08	0.15
rs1552224	*CENTD2*	Voight et al., 2010 (22)	rs1552224		A	0.88	0.84	1.14 (1.11–1.17)	1.4×10^{-22}	1.11 (0.94–1.31)	0.22	+	0.76	0.37
rs1531343	*HMGA2*	Voight et al., 2010 (22)	rs2612067	1.00#	C	0.10	0.19	1.10 (1.07–1.14)	3.6×10^{-9}	1.02 (0.87–1.19)	0.83	+	0.35	0.23
rs7957197	*HNF1A*	Voight et al., 2010 (22)	rs7965349	0.87€	T	0.85	0.95	1.07 (1.05–1.10)	2.4×10^{-8}	1.30 (0.96–1.76)	0.09	+	0.21	0.08
rs11634397	*ZFAND6*	Voight et al., 2010 (22)	rs4778582	0.74#	G	0.60	0.55	1.06 (1.04–1.08)	2.4×10^{-9}	1.09 (0.97–1.23)	0.16	+	0.66	0.16
rs8042680	*PRC1*	Voight et al., 2010 (22)	rs8042680		A	0.22	0.62	1.07 (1.05–1.09)	2.4×10^{-10}	1.11 (0.98–1.26)	0.10	+	0.57	0.19
rs5945326	*DUSP9*	Voight et al., 2010 (22)	rs5945326		A	0.79	0.58	1.27 (1.18–1.37)	3.0×10^{-10}	1.28 (1.10–1.48)	0.001	+	0.93	0.97
rs7612463	*UBE2E2*	Yamauchi et al., 2010 (5)	rs7612463		C	0.82	0.77	1.19 (1.12–1.26)	2.27×10^{-9}	1.19 (1.03–1.37)	0.02	+	1.00	0.65
rs7172432	*C2CD4A−C2CD4B*	Yamauchi et al., 2010 (5)	rs7172432		A	0.56	0.60	1.11 (1.08–1.14)	8.8×10^{-14}	1.13 (1.00–1.28)	0.04	+	0.78	0.38
rs1359790	*SPRY2*	Shu et al., 2010 (6)	rs1215451	1.00#	G	0.71	0.80	1.15 (1.10–1.20)	6.5×10^{-9}	1.26 (1.07–1.48)	0.005	+	0.28	0.43
rs3923113	*GRB14*	Kooner et al., 2011 (7)	rs3923113		A	0.74	0.76	1.09 (1.06–1.13)	1.0×10^{-8}	1.15 (0.99–1.33)	0.07	+	0.48	0.22
rs16861329	*ST6GAL1*	Kooner et al., 2011 (7)	rs16861329		C	0.75	0.75	1.09 (1.06–1.12)	3.4×10^{-8}	1.14 (0.99–1.31)	0.07	+	0.54	0.23
rs1802295	*VPS26A*	Kooner et al., 2011 (7)	rs1802295		T	0.26	0.27	1.08 (1.05–1.12)	4.1×10^{-8}	1.09 (0.95–1.25)	0.24	+	0.90	0.20
rs7178572	*HMG20A*	Kooner et al., 2011 (7)	rs7178572		G	0.52	0.51	1.09 (1.06–1.12)	7.1×10^{-11}	1.15 (1.02–1.30)	0.02	+	0.40	0.29
rs2028299	*AP3S2*	Kooner et al., 2011 (7)	rs12594808	0.95#	C	0.31	0.30	1.10 (1.07–1.13)	1.9×10^{-11}	1.10 (0.96–1.25)	0.17	+	1.00	0.30
rs4812829	*HNF4A*	Kooner et al., 2011 (7)	rs4812829		A	0.29	0.28	1.09 (1.06–1.12)	2.6×10^{-10}	1.22 (1.07–1.39)	0.004	+	0.10	0.25
rs6815464	*MAEA*	Cho et al., 2012 (8)	rs6599300	1.00¥	C	0.58	0.84	1.13 (1.10–1.16)	1.6×10^{-20}	1.18 (0.99–1.39)	0.06	+	0.62	0.78
rs7041847	*GLIS3*	Cho et al., 2012 (8)	rs7041847		A	0.41	0.57	1.10 (1.07–1.13)	1.9×10^{-14}	0.99 (0.87–1.11)	0.82	−	0.10	0.57
rs6467136	*GCC1-PAX4*	Cho et al., 2012 (8)	rs6467136		G	0.79	0.56	1.11 (1.07–1.14)	4.9×10^{-11}	0.88 (0.78–1.00)	0.05	−	4×10^{-4}	0.48
rs831571	*PSMD6*	Cho et al., 2012 (8)	rs831571		C	0.61	0.78	1.09 (1.06–1.12)	8.4×10^{-11}	0.90 (0.77–1.04)	0.16	−	0.01	0.47
rs9470794	*ZFAND3*	Cho et al., 2012 (8)	rs9470794		C	0.27	0.12	1.12 (1.08–1.16)	2.1×10^{-10}	1.11 (0.92–1.33)	0.27	+	0.93	0.63
rs3786897	*PEPD*	Cho et al., 2012 (8)	rs7258031	0.83#	A	0.56	0.74	1.10 (1.07–1.14)	1.3×10^{-8}	1.05 (0.91–1.20)	0.53	+	0.52	0.57
rs1535500	*KCNK16*	Cho et al., 2012 (8)	rs3734618	1.00€	T	0.42	0.47	1.08 (1.05–1.11)	2.3×10^{-8}	1.08 (0.96–1.22)	0.20	+	1.00	0.41

Effect sizes of the source study and stage 1 of the current study (Indo-Europeans) are presented with respect to the risk allele in the source study. Dir., direction; Het., heterogeneity; NA, not available; RAF, risk allele frequency in control subjects. ¶Risk allele as in the source study (indexed to the positive strand of National Center for Biotechnology Information build 37). *Direction was + if there was concordance between the source and present study and − if there was discordance. §P for difference in OR of the source and present study. ‖Power of the current study based on its sample size and allele frequency to detect the previously reported effect sizes under a log-additive model assuming 10% prevalence of T2D. #r^2 values from HapMap data for GIH. €r^2 values from the 1000 Genomes project for CEU. ¥r^2 values from the 1000 Genomes project for CHB+JPT.

transcription of insulin (32). We observed suggestive association at rs10461617 in the 5′ flank of *MAP3K1*. The rs10461617 was also observed to have nominal association with T2D in Dravidians and DIAGRAM+ with consistency in direction of effect. Although the association of *MAP3K1* variant could not reach genome-wide significance, these results and biological roles of *MAP3K1* suggest it as an important candidate for T2D.

Most of the previously reported T2D-associated loci from earlier GWAS showed consistency in direction of association with similar effect sizes in Indo-European subjects of the present study, of which 25 were also associated at $P < 0.05$. Association at other known loci could not be detected mainly due to the low power of the current study to detect those associations, as most of these loci were identified through meta-analysis of GWAS, including large study populations. Only at loci *PTPRD*, *HNF1B*, *THADA*, and *SRR*, we did not observe association even when our study was sufficiently powered (>98%) to detect the previously reported effect sizes. The effect sizes at these loci in our study population could be very small; thus, a larger study population would be required to confirm their association status in our population. We also found consistent results as in the recent GWAS in South Asians. Of the six loci identified in the recent GWAS in South Asians (7), *HNF4A* has been shown to be associated in our previous study (33) involving study subjects from the current study and was also associated in the current study ($P = 0.004$). The loci *HMG20A* was also associated in the current study ($P = 0.02$) while the other four loci (*GRB14*, *ST6GAL1*, *VPS26A*, and *AP3S2*) showed consistency in direction and effect size. All of the previously associated loci in Indo-Europeans accounted for 7.17% of the variance in the risk of T2D in our study population, which is similar to the observations in previous studies (5 to 10%). The newly identified 2q21 locus alone contributed 0.48% of the variance, and addition of 2q21 locus to the list increased the total variance explained to 7.65%. Thus, our analysis provides further evidence for the belief that underlying disease mechanisms involving common variants are largely similar among populations (11).

In conclusion, we identified a new locus associated with T2D on the 2q21 region. The identified region 2q21 harbors the genes *TMEM163*, *RAB3GAP1*, and *ACMSDI*, and incidentally, all three are involved in neurologic processes, suggesting a neurologic component in the etiology of T2D. The findings provide new insights into the genetic architecture and pathophysiology of T2D; however, further studies exploring the functional and clinical significance of the identified locus are warranted.

ACKNOWLEDGMENTS

The major funding for this work comes from Council for Scientific and Industrial Research, Government of India, in the form of the grant “Diabetes mellitus—New drug discovery R&D, molecular mechanisms, and genetic and epidemiological factors” (NWP0032-19). R.T. received a postdoctoral fellowship from the Fogarty International Center and the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development at the National Institutes of Health (D43-HD-065249).

No potential conflicts of interest relevant to this article were reported.

R.T., G.C., and O.P.D. researched data, contributed to discussion, and wrote the manuscript. A.M., A.J., I.K., K.B., T.S., B.J.M., Y.P., M.I.M., R.V., and V.M. researched data and contributed to discussion. M.C., A.S., S.C., S.S., L.R., P.V., S.K.A., S.G., D.P., R.K.S., M.S., M.B., S.M., A.Bh., V.N.S., S.V.M., R.K.M., A.Ba., and V.S. researched data. N.T. and D.B. researched data, contributed to discussion, and reviewed and edited the manuscript. D.B. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

The authors thank all of the study subjects for participation. The authors also thank A.K. Sharma (Council for Scientific and Industrial Research-Institute of Genomics and Integrative Biology) for helping in sample collection and Prof. Greg Gibson (Georgia Institute of Technology, Atlanta, Georgia) for critical evaluation of the manuscript.

REFERENCES

1. Unwin N, Gan D, Whiting D. The IDF Diabetes Atlas: providing evidence, raising awareness and promoting action. Diabetes Res Clin Pract 2010;87:2–3
2. Ramachandran A, Ma RC, Snehalatha C. Diabetes in Asia. Lancet 2010;375: 408–418
3. Indian Genome Variation Consortium. Genetic landscape of the people of India: a canvas for disease gene exploration. J Genet 2008;87:3–20
4. McCarthy MI. Genomics, type 2 diabetes, and obesity. N Engl J Med 2010; 363:2339–2350
5. Yamauchi T, Hara K, Maeda S, et al. A genome-wide association study in the Japanese population identifies susceptibility loci for type 2 diabetes at UBE2E2 and C2CD4A-C2CD4B. Nat Genet 2010;42:864–868
6. Shu XO, Long J, Cai Q, et al. Identification of new genetic risk variants for type 2 diabetes. PLoS Genet 2010;6:6
7. Kooner JS, Saleheen D, Sim X, et al.; DIAGRAM; MuTHER. Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. Nat Genet 2011;43:984–989
8. Cho YS, Chen CH, Hu C, et al.; DIAGRAM Consortium; MuTHER Consortium. Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in east Asians. Nat Genet 2012;44:67–72
9. Bustamante CD, Burchard EG, De la Vega FM. Genomics for the world. Nature 2011;475:163–165
10. Sim X, Ong RT, Suo C, et al. Transferability of type 2 diabetes implicated loci in multi-ethnic cohorts from Southeast Asia. PLoS Genet 2011;7: e1001363
11. McCarthy MI. The importance of global studies of the genetics of type 2 diabetes. Diabetes Metab J 2011;35:91–100
12. Reich D, Thangaraj K, Patterson N, Price AL, Singh L. Reconstructing Indian population history. Nature 2009;461:489–494
13. INdian DIabetes COnsortium. INDICO: the development of a resource for epigenomic study of Indians undergoing socioeconomic transition. The HUGO journal 2011;5:65–69
14. Indian Genome Variation Consortium. The Indian Genome Variation database (IGVdb): a project overview. Hum Genet 2005;118:1–11
15. 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 2003;26(Suppl. 1):S5–S20
16. Deepa M, Pradeepa R, Rema M, et al. The Chennai Urban Rural Epidemiology Study (CURES)—study design and methodology (urban component) (CURES-I). J Assoc Physicians India 2003;51:863–870
17. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia 1985;28:412–419
18. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–575
19. Price AL, Weale ME, Patterson N, et al. Long-range LD can confound genome scans in admixed populations. Am J Hum Genet 2008;83:132–135; author reply 135–139
20. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010;26:2190–2191
21. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21:263–265
22. Voight BF, Scott LJ, Steinthorsdottir V, et al.; MAGIC investigators; GIANT Consortium. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet 2010;42:579–589

23. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 2011;88:76–82
24. Heard-Costa NL, Zillikens MC, Monda KL, et al. NRXN3 is a novel locus for waist circumference: a genome-wide association study from the CHARGE Consortium. PLoS Genet 2009;5:e1000539
25. Ma L, Yang J, Runesha HB, et al. Genome-wide association analysis of total cholesterol and high-density lipoprotein cholesterol levels using the Framingham heart study data. BMC Med Genet 2010;11:55
26. Barth J, Zimmermann H, Volknandt W. SV31 is a Zn2+-binding synaptic vesicle protein. J Neurochem 2011;118:558–570
27. Burré J, Zimmermann H, Volknandt W. Identification and characterization of SV31, a novel synaptic vesicle membrane protein and potential transporter. J Neurochem 2007;103:276–287
28. Aligianis IA, Johnson CA, Gissen P, et al. Mutations of the catalytic subunit of RAB3GAP cause Warburg Micro syndrome. Nat Genet 2005;37:221–223
29. Tanabe A, Egashira Y, Fukuoka S, Shibata K, Sanada H. Expression of rat hepatic 2-amino-3-carboxymuconate-6-semialdehyde decarboxylase is affected by a high protein diet and by streptozotocin-induced diabetes. J Nutr 2002;132:1153–1159
30. Garavaglia S, Perozzi S, Galeazzi L, Raffaelli N, Rizzi M. The crystal structure of human alpha-amino-beta-carboxymuconate-epsilon-semialdehyde decarboxylase in complex with 1,3-dihydroxyacetonephosphate suggests a regulatory link between NAD synthesis and glycolysis. FEBS J 2009;276:6615–6623
31. Mokhtari D, Myers JW, Welsh N. The MAPK kinase kinase-1 is essential for stress-induced pancreatic islet cell death. Endocrinology 2008;149:3046–3053
32. Oetjen E, Blume R, Cierny I, et al. Inhibition of MafA transcriptional activity and human insulin gene transcription by interleukin-1beta and mitogen-activated protein kinase kinase kinase in pancreatic islet beta cells. Diabetologia 2007;50:1678–1687
33. Chavali S, Mahajan A, Tabassum R, et al. Association of variants in genes involved in pancreatic β-cell development and function with type 2 diabetes in North Indians. J Hum Genet 2011;56:695–700
34. Saxena R, Voight BF, Lyssenko V, et al. Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, and Novartis Institutes of BioMedical Research. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 2007;316: 1331–1336
35. Scott LJ, Mohlke KL, Bonnycastle LL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 2007;316:1341–1345
36. Sandhu MS, Weedon MN, Fawcett KA, et al. Common variants in WFS1 confer risk of type 2 diabetes. Nat Genet 2007;39:951–953
37. Gudmundsson J, Sulem P, Steinthorsdottir V, et al. Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. Nat Genet 2007;39:977–983
38. Zeggini E, Scott LJ, Saxena R, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat Genet 2008;40:638–645
39. Yasuda K, Miyake K, Horikawa Y, et al. Variants in KCNQ1 are associated with susceptibility to type 2 diabetes mellitus. Nat Genet 2008;40:1092–1097
40. Prokopenko I, Langenberg C, Florez JC, et al. Variants in MTNR1B influence fasting glucose levels. Nat Genet 2009;41:77–81
41. Rung J, Cauchi S, Albrechtsen A, et al. Genetic variant near IRS1 is associated with type 2 diabetes, insulin resistance and hyperinsulinemia. Nat Genet 2009;41:1110–1115
42. Dupuis J, Langenberg C, Prokopenko I, et al. New genetic loci implicated in fasting glucose homeostasis and their im- pact on type 2 diabetes risk. Nat Genet 2010;42:105–116
43. Tsai FJ, Yang CF, Chen CC, et al. A Genome-Wide Association Study Identifies Susceptibility Variants for Type 2 Diabetes in Han Chinese. PLoS Genet 2010;6:e1000847
44. Qi L, Cornelis MC, Kraft P, et al. Diabetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium: Genetic variants at 2q24 are associated with susceptibility to type 2 diabetes. Hum Mol Genet 2010;19:2706–2715