# **ORIGINAL ARTICLE**

# Association of variants in genes involved in pancreatic β-cell development and function with type 2 diabetes in North Indians

Sreenivas Chavali<sup>1,4</sup>, Anubha Mahajan<sup>1,4</sup>, Rubina Tabassum<sup>1</sup>, Om Prakash Dwivedi<sup>1</sup>, Ganesh Chauhan<sup>1</sup>, Saurabh Ghosh<sup>2</sup>, Nikhil Tandon<sup>3</sup> and Dwaipayan Bharadwaj<sup>1</sup>

Variants in genes involved in pancreatic  $\beta$ -cell development and function are known to cause monogenic forms of type 2 diabetes and are also associated with complex form. In this study, we studied the genetic association of polymorphisms in such important genes with type 2 diabetes in the high-risk Indians. We genotyped 91 polymorphisms in 19 genes (*ABCC8, HNF1A, HNF1B, HNF4A, INS, INSM1, ISL1, KCNJ11, MAFA, MNX1, NEUROD1, NEUROG3, NKX2.2, NKX6.1, PAX4, PAX6, PDX1, USF1* and *WFS1*) in 2025 unrelated North Indians of Indo-European ethnicity comprising of 1019 diabetic and 1006 non-diabetic subjects. *HNF4A* promoter P2 polymorphisms rs1884613 and rs2144908, which are in high linkage disequilibrium, showed significant association with type 2 diabetes (odds ratio (OR)=1.37 (95% confidence interval (CI) 1.19–1.57),  $P=9.4 \times 10^{-6}$  for rs1884613 and OR=1.37 (95%CI 1.20–1.57),  $P=6.0 \times 10^{-6}$  for rs2144908), as previously shown in other populations. We observed body mass index-dependent association of these variants with type 2 diabetes in normal-weight/lean subjects. Variants in *USF1, ABCC8, ISL1* and *KCNJ11* showed nominal association, while haplotypes in these genes were significantly associated. rs3812704 upstream of *NEUROG3* significantly increased risk for type 2 diabetes in normal-weight/lean subjects (OR=1.68 (95%CI 1.25–2.24),  $P=4.9 \times 10^{-4}$ ). Thus, pancreatic  $\beta$ -cell development and function genes contribute to susceptibility to type 2 diabetes in North Indians.

Journal of Human Genetics (2011) 56, 695–700; doi:10.1038/jhg.2011.83; published online 4 August 2011

Keywords: association; β-cell; HNF4A; NEUROG3; obesity; type 2 diabetes

## INTRODUCTION

Type 2 diabetes (T2D) is characterized by impaired glucose tolerance and defects in insulin function and/or secretion resulting in insulin resistance. Studies of maturity-onset diabetes of the young in humans have shown that mutations in transcription factors required for development, differentiation and maintenance of pancreatic  $\beta$ -cells lead to  $\beta$ -cell dysfunction and hence, T2D.<sup>1</sup> Several studies in various ethnic populations have shown association of polymorphisms in genes involved in  $\beta$ -cell development and function with T2D and associated phenotypes.<sup>2–7</sup> Importantly, genome-wide association studies have also deciphered associations with loci predominantly involved in  $\beta$ -cell function.<sup>8</sup>

Indians have the highest incidence and prevalence of T2D worldwide.<sup>9</sup> Previous studies in Indians have suggested role of genetic variants in genes involved in  $\beta$ -cell development and function in T2D. We showed strong association of variants in *FOXA2*, an upstream activator of  $\beta$ -cell transcription factor network and *KCNJ11*, a regulator of insulin secretion with T2D, influenced by obesity in North Indians.<sup>3,4</sup> Another study in Dravidians from South India demonstrated association of polymorphisms in *NEUROD1*, *NEUROG3* and *HNF1A*, involved in  $\beta$ -cell development and function, with glucose intolerance.<sup>7</sup> Therefore, it is paramount to comprehensively assess the role of these genes in high-risk Indians. Here, we have evaluated genetic association of 91 single-nucleotide polymorphisms in 19 genes involved in pancreatic  $\beta$ -cell development and function with T2D and related metabolic traits in North Indians.

## MATERIALS AND METHODS

We analyzed a study population of 2025 unrelated North Indian subjects of Indo-European ethnicity comprising 1019 T2D patients (cases) and 1006 nondiabetic subjects (controls). Cases were recruited from the Endocrinology clinic of All India Institute of Medical Sciences, New Delhi, in accordance with WHO

<sup>&</sup>lt;sup>1</sup>Genomics and Molecular Medicine Unit, Institute of Genomics and Integrative Biology, CSIR, Delhi, India; <sup>2</sup>Human Genetics Unit, Indian Statistical Institute, Kolkata, West Bengal, India and <sup>3</sup>Department of Endocrinology, All India Institute of Medical Sciences, New Delhi, India <sup>4</sup>These authors contributed equally to this work.

Correspondence: Dr D Bharadwaj, Genomics and Molecular Medicine Unit, Institute of Genomics and Integrative Biology, CSIR, Mall Road, Delhi 110007, India.

E-mail: db@igib.res.in

or Professor N Tandon, Department of Endocrinology, All India Institute of Medical Sciences, Ansari Nagar, New Delhi 110 029, India.

E-mail: nikhil\_tandon@hotmail.com

Received 26 April 2011; revised 21 June 2011; accepted 23 June 2011; published online 4 August 2011

Expert Committee criteria.<sup>10</sup> Recruitment of cases and controls was carried out as previously described.<sup>3,11</sup> The study was carried out in accordance with the principles of Helsinki Declaration and approved by the ethics committee of both institutions. Informed written consent was obtained from all the participants.

Anthropometric measurements (height, weight, waist and hip circumferences) were carried out as per standardized protocols and body mass index (BMI) and waist hip ratio was calculated. Biochemical measurements including levels of glycated hemoglobin, fasting and 2-h postload plasma glucose, fasting plasma insulin, C-peptide, total cholesterol, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, triglyceride and calculation of HOMA-IR and HOMA-B were carried out as described previously.<sup>3,11,12</sup>

We selected 19 genes involved in pancreatic  $\beta$ -cell development and function including *ABCC8*, *HNF1A*, *HNF1B*, *HNF4A*, *INS*, *INSM1*, *ISL1*, *KCNJ11*, *MAFA*, *MNX1*, *NEUROD1*, *NEUROG3*, *NKX2.2*, *NKX6.1*, *PAX4*, *PAX6*, *PDX1*, *USF1* and *WFS1*. We selected 91 polymorphisms including those associated with any disease including T2D, located in the functional regions of genes, polymorphic in at least two HapMap II populations and/or tag SNPs from HapMap data from these genes (Supplementary Table 1).

Genotyping of 81 polymorphisms was carried out using Illumina GoldenGate assay and stringent quality control checks were applied.<sup>13</sup> Samples with a poor call rate (>10% missing) were excluded from further analysis. All the polymorphisms passed the quality check and had an average call rate of 98.2% and concordance rate >99% based on 7.2% duplication. Ten polymorphisms were genotyped using iPLEX (Sequenom, San Diego, CA, USA). The average genotyping success rate for these was 96.1%, with >99% consensus rate on a subset of 4% duplicate samples. Genotype frequencies for all the polymorphisms are provided in Supplementary Table 1. Loci that were monomorphic or had a minor allele frequency <5% in North Indians were not considered for association analysis.

We assessed deviation of genotypic frequencies from Hardy-Weinberg equilibrium (HWE) using  $\chi^2$ -test. Polymorphisms with deviations in HWE  $(P < 5.81 \times 10^{-4}; \alpha = 0.05/86)$  in controls (rs2071190 of *HNF1A* and rs6446482 of WFS1) were disregarded. Further analysis was carried out using 84 polymorphisms. Using additive model, we analyzed differences between genotype distribution frequencies and performed logistic regression to determine association with T2D after adjusting for age, sex and BMI. A P value  $< 5.9 \times 10^{-4}$ was considered significant after correcting for multiple testing (Bonferroni's correction;  $\alpha = 0.05/84$ ). The uncorrected P values are provided in the text. Power of the study was estimated using QUANTO version 1.2.14 Linear regression analysis was performed to test association of polymorphisms with natural log-transformed quantitative traits. Haplotype association analysis was carried out using Haploview 4.1.15 Only haplotypes with frequency >10% were used and statistical significance was assessed based on 10000 permutations. Meta-analysis was performed by combining summary estimates under random effect model. The forest plot for meta-analysis was drawn using STATA v. SE 10.1 (StataCorp, College Station, TX, USA). Statistical analyses were performed using SPSS version 17.0 (SPSS, Chicago, IL, USA) and PLINK.<sup>16</sup>

#### Table 1 Genotype and haplotype associations in the genes involved in pancreatic β-cell development and function with T2D in North Indians

Gene; polymorphism	MAF cases/ controls	OR <sup>a</sup> (95% CI); P-value <sup>a</sup>	<i>OR<sub>adj</sub><sup>b</sup> (95% CI</i> ); P <i>-value</i> <sub>adj</sub> <sup>b</sup>	Gene/s polymorphisms	Haplotypes; frequency in cases/ controls	OR; P-value	<i>OR<sub>adj</sub>b</i> P- <i>value</i> <sup>b</sup>
ABCC8; rs985136	0.40/0.44	0.86 (0.76–0.97);	0.86 (0.76–0.98);	ABCC8; rs1805036,	GT; 0.33/	0.87; 0.04	0.87; 0.03
		0.017	0.021	rs1799854	0.36		
ABCC8; rs1799854	0.33/0.36	0.87 (0.76–0.99);	0.86 (0.76–0.99);	<i>ABCC8;</i> rs1805036, rs1799854	GC; 0.61/ 0.58	1.17; 0.016	1.18; 0.013
ABCC8: rs757110	0.42/0.38	1.17 (1.03–1.33)	1.16 (1.02–1.32)	KCN111_ABCC8: rs5210	GCTGAG	1.17	1.15.0.03
,12000,10,0,110	0112/0100	0.017	0.026	rs5215_rs5219	0.39/0.35	0.017	1110, 0100
		0.017	0.020	rs5222, rs7124355, rs757110	0100/0100	0.017	
ABCC8; rs7124355	0.38/0.35	1.15 (1.01–1.31);	1.13 (0.99–1.29);	HNF4A; rs1884613,	GT; 0.33/	1.33;	1.35;
		0.03	0.06	rs2144908	0.27	$5.3 \times 10^{-5}$	$2.8 \times 10^{-5}$
HNF4A; rs1884613	0.34/0.27	1.35 (1.17–1.54);	1.37 (1.19–1.57);	HNF4A; rs1884613,	CC; 0.67/	0.73;	0.72;
		$1.9 \times 10^{-5}$	9.4×10 <sup>-6</sup>	rs2144908	0.73	8.2×10 <sup>-6</sup>	$3.5 \times 10^{-6}$
HNF4A; rs2144908	0.33/0.27	1.35 (1.18–1.54);	1.37 (1.20–1.57);	ISL1; rs2288468, rs3811911,	TCCA; 0.70/	1.24;	1.26;
		$1.5 \times 10^{-5}$	6.0×10 <sup>-6</sup>	rs991216, rs1017	0.65	$1.5 \times 10^{-3}$	$8.9 \times 10^{-4}$
ISL1; rs2288468	0.30/0.35	0.81 (0.71–0.92);	0.80 (0.70–0.91);	<i>ISL1;</i> rs2288468, rs3811911,	CTGT; 0.13/	0.77;	0.78;
		$1.7 \times 10^{-3}$	$1.0 \times 10^{-3}$	rs991216, rs1017	0.16	$4.3 \times 10^{-3}$	$5.1 \times 10^{-3}$
<i>ISL1;</i> rs3811911	0.13/0.16	0.77 (0.64–0.92);	0.77 (0.65–0.92);	USF1; rs3737787,	AGAACC;	0.76;	0.74;
		4.1×10 <sup>-3</sup>	4.8×10 <sup>-3</sup>	rs2774276, rs2073653, rs2516839, rs2516838,	0.11/0.14	3.4×10 <sup>-3</sup>	2.0×10 <sup>-3</sup>
ISI 1. rc001216	0 20/0 25	0.01 (0.71 0.02).	0 82 (0 70 0 02).	152510637			
1321;13991210	0.30/0.33	$3.1 \times 10^{-3}$	$1.9 \times 10^{-3}$				
<i>ISL1;</i> rs1017	0.29/0.34	0.81 (0.70-0.93);	0.80 (0.69–0.92);				
,		2.6×10 <sup>-3</sup>	1.4×10 <sup>-3</sup>				
KCNJ11; rs5215	0.39/0.35	1.19 (1.05–1.35);	1.17 (1.03–1.34);				
		8.8×10 <sup>-3</sup>	0.017				
KCNJ11; rs5219	0.39/0.35	1.19 (1.05–1.36);	1.17 (1.03–1.34);				
		8.6×10 <sup>-3</sup>	0.017				
USF1; rs3737787	0.11/0.14	0.75 (0.63–0.91); 3.3×10 <sup>-3</sup>	0.74 (0.61–0.89); 1.8×10 <sup>-3</sup>				

Bold letters in the genotype association columns indicate significant *P*-values after correcting for multiple comparisons.

For haplotype associations *P*-value estimates are based on 10000 permutations. <sup>a</sup>Based on additive model.

<sup>b</sup>Adjusted for age, sex and BMI.

RESULTS

The clinical and anthropometric details of the study population have been provided in Supplementary Table 2. Of the 84 variants in 19 genes analyzed, 13 in five genes showed significant association (P < 0.05) with T2D (Table 1). The *HNFA* promoter P2 polymorphisms rs1884613 and rs2144908 (Odds ratio (OR)=1.35 (95% confidence interval (CI) 1.17– 1.54)  $P=1.9 \times 10^{-5}$  and OR=1.35 (95% cI 1.18–1.54),  $P=1.5 \times 10^{-5}$ respectively) showed significant association after correcting for multiple testing. The association retained significance upon adjusting for age, sex and BMI (OR=1.37 (95%CI 1.19–1.57),  $P=9.4 \times 10^{-6}$  for rs2144908). The common variant rs5219 (E23 K) in *KCNJ11* associated with T2D through genome-wide association studies, also showed nominal association in North Indians. *WFS1* polymorphism rs10010131, which was previously associated with T2D in other ethnic groups did not show any association in North Indians. The major haplotype (CC) at the promoter locus of *HNF4A* involving the polymorphisms rs1884613 and rs2144908 increased risk for T2D while the minor GT haplotype conferred protection (Table 1). Similarly, the major haplotype TCCA in *ISL1* involving rs2288468, rs3811911, rs991216 and rs1017 polymorphisms conferred susceptibility to T2D, while the less frequent CTGT haplotype conferred protection. The AGAACC haplotype of *USF1* was significantly decreased in the cases conferring protection from T2D.

Our previous studies demonstrated an influence of obesity on the association of genetic variants with T2D in North Indians.<sup>3,17</sup> Here, we investigated the risk for T2D conferred by these genetic variants by segregating the subjects into obese (BMI  $\ge 25 \text{ kg m}^{-2}$ ) and non-obese (BMI  $< 25 \text{ kg m}^{-2}$ ) based on WHO criteria for Asian population.<sup>18</sup> Polymorphisms in *HNF1A*, *HNF4A*, *ISL1* and *USF1* showed nominal association with T2D in obese subjects (Table 2). Haplotype associations were observed with variants across *KCNJ11* and *ABCC8* and in

Table 2 Genotype association of polymorphisms with T2D in different study groups determined based on	on Bl	l o	sed	ba	эd	ine	eterm	s d	group	Jy .	stud	erent	diff	) in	T2D	with	phisms	ymor	i poly	۱ of	ciation	as	type	Geno	le 2	Tab
--	-------	-----	-----	----	----	-----	-------	-----	-------	------	------	-------	------	------	-----	------	--------	------	--------	------	---------	----	------	------	------	-----

Gene; polymorphism	OR (95% CI); P-value	<i>OR<sub>adj</sub>a (95% CI);</i> P- <i>value<sub>adj</sub>a</i>	Gene/s; polymorphisms	Haplotypes	OR; P-value	<i>OR<sub>adj</sub>a</i> ; P <i>-value<sub>adj</sub>a</i>
Obese subjects (cases=47	76; controls=388)					
HNF1A; rs7953249	0.82 (0.67–0.99); 0.05	0.81 (0.67–0.99); 0.04	HNF1A; rs7953249, rs1920792, rs1169288	AGT	0.80; 0.03	0.80; 0.03
<i>HNF4A;</i> rs6093978	0.80 (0.66–0.98); 0.03	0.81 (0.67–0.99); 0.04	<i>KCNJ11, ABCC8</i> ; rs1002227, rs5210, rs5215, rs5219, rs5222, rs7124355, rs757110	CGTCGGT	0.78; 0.03	0.78; 0.03
<i>ISL1;</i> rs2288468	0.79 (0.64–0.97); 0.03	0.79 (0.64–0.97); 0.03	<i>ISL1;</i> rs2288468, rs3811911, rs991216, rs1017	TCCA	1.27; 0.03	1.27; 0.03
<i>ISL1;</i> rs991216	0.80 (0.64–0.98); 0.03	0.79 (0.64–0.98); 0.03	USF1; rs3737787, rs2774276, rs2073653, rs2516839, rs2516838, rs2516837	AGAACC	0.71; 0.01	0.68; 6.3×10 <sup>-3</sup>
<i>ISL1;</i> rs1017	0.78 (0.63–0.97); 0.03	0.78 (0.63–0.97); 0.02				
USF1; rs3737787	0.70 (0.54–0.93); 0.012	0.68 (0.51–0.89); 5.6×10 <sup>-3</sup>				
Non-obese subjects (cases	=509; controls=610)					
ABCC8; rs722341	0.74 (0.57–0.95); 0.02	0.72 (0.55–0.94); 0.02	ABCC8; rs722341, rs3819523	AG	0.74; 0.03	0.71; 0.02
ABCC8; rs757110	1.21 (1.02–1.44); 0.03	1.11 (0.92–1.33); 0.27	ABCC8; rs722341, rs3819523	GG	1.22; 0.04	1.2; 0.06
ABCC8; rs7124355	1.21 (1.01–1.44); 0.03	1.11 (0.92–1.33); 0.28	HNF4A; rs1884613, rs2144908	GT	1.45; 7.61×10 <sup>-5</sup>	1.42; 3.86×10 <sup>-4</sup>
HNF4A; rs1884613	1.48 (1.23–1.77); 2.96×10 <sup>−5</sup>	1.44 (1.19–1.75); 1.8×10 <sup>−4</sup>	HNF4A; rs1884613, rs2144908	CC	0.66; 9.88×10 <sup>-6</sup>	0.68; 6.56×10 <sup>-5</sup>
HNF4A; rs2144908	1.48 (1.23–1.77); 2.84×10 <sup>−5</sup>	1.45 (1.20–1.76); 1.4×10 <sup>−4</sup>	KCNJ11; rs1002227, rs5210	CG	1.22; 0.03	1.17; 0.11
INSM1; rs10470054	1.33 (1.01–1.75); 0.04	1.40 (1.05–1.88); 0.02	KCNJ11, ABCC8; rs5215, rs5219, rs5222, rs7124355, rs757110	CTGAG	1.25; 0.01	1.14; 0.16
<i>ISL1;</i> rs2288468	0.80 (0.67–0.96); 0.02	0.79 (0.65–0.96); 0.02	<i>ISL1;</i> rs2288468, rs3811911, rs991216, rs1017	CTGT	0.75; 0.02	0.73; 0.02
<i>ISL1;</i> rs3811911	0.74 (0.58–0.95); 0.02	0.72 (0.56–0.93); 0.01	<i>ISL1;</i> rs2288468, rs3811911, rs991216, rs1017	TCCA	1.25; 0.02	1.27; 0.01
<i>ISL1;</i> rs991216	0.80 (0.67–0.97); 0.02	0.79 (0.65–0.97); 0.02	<i>NEUROG3;</i> rs3812704, rs2812545	AG	1.48; 9.3×10 <sup>-4</sup>	0.74; 3.6×10 <sup>-3</sup>
<i>ISL1;</i> rs1017	0.80 (0.67–0.97); 0.02	0.79 (0.64–0.96); 0.02	<i>NEUROG3;</i> rs3812704, rs2812545	GG	0.77; 6.9×10 <sup>-3</sup>	0.74; 3.6×10 <sup>-3</sup>
KCNJ11; rs5219	1.3 (1.05–1.49); 0.01	1.14 (0.95–1.37); 0.16				
KCNJ11; rs5215	1.24 (1.04–1.50); 0.02	1.14 (0.94–1.34); 0.18				

697

npg 698

# Table 2 (Continued)

Gene; polymorphism	OR (95% CI); P-value	<i>OR<sub>adj</sub><sup>a</sup> (95% CI);</i> P- <i>value<sub>adj</sub><sup>a</sup></i>	Gene/s; polymorphisms	Haplotypes	OR; P-value	<i>OR<sub>adj</sub>a</i> ; P <i>-value<sub>adj</sub>a</i>
NEUROG3; rs3812704	1.48 (1.77–1.87);	1.54 (1.20–1.97);				
	8.3×10 <sup>-4</sup>	5.7×10 <sup>-4</sup>				
<i>PAX6;</i> rs628224	0.81 (0.65–1.00);	0.78 (0.62–0.98);				
	0.05	0.03				
Over-weight subjects (cases	=215; controls=174)					
HNF1B; rs12450628	1.38 (1.02- 1.85);	1.40 (1.03- 1.89);	HNF4A; rs1884613, rs2144908	GT	1.40; 0.04	1.37; 0.05
	0.04	0.03				
HNF4A; rs2144908	1.40 (1.02- 1.91);	1.39 (1.01–1.90);	HNF4A; rs1884613, rs2144908	CC	0.72; 0.04	0.73; 0.05
	0.04	0.04				
HNF4A; rs3212199	0.73 (0.55–0.96);	0.72 (0.54–0.95);	HNF4A; rs3212199, rs6031595	GA	1.35; 0.04	1.38; 0.03
	0.02	0.02				0.70.0.01
HNF4A; rs6031595	1.37 (1.03–1.82);	1.41 (1.06–1.87);	HNF4A; rs3212199, rs6031595	11	0.72; 0.02	0.70; 0.01
INS. rs3812718	0.05	0.02	INS, re3812753 re3812718	тс	0.64.0.04	0.64, 0.04
1110,133042740	0.03 (0.44-0.37),	0.00 (0.44–0.58),	1110, 1300427 33, 1300427 40	Ĩŭ	0.04, 0.04	0.04, 0.04
NKX6.1. rs2278671	1.39 (1.02–1.89)	1.39 (1.02–1.89)	NFUROG3: rs3812704	GG	0.71.0.03	0.70.0.02
11110.1, 13227 007 1	0.04	0.04	rs2812545	dd	0.7 1, 0.00	0.70, 0.02
NKX6.1: rs2278672	1.41 (1.04–1.93):	1.41 (1.04–1.93):	<i>NKX6.1:</i> rs2278672, rs2278671.	СТСС	0.73: 0.04	0.73: 0.05
	0.03	0.03	rs3806750, rs1867544		,	,
Normal weight/lean subject	s (cases-291, controls-13	26)				
ABCC8 rs985136	0 78 (0 63–0 97)	0 74 (0 59–0 93)	ABCC8: rs722341 rs3819523	AGG	0 70. 0 04	0.62.
,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	0.02	0.01	rs1048099	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	011 0, 010 1	$8.4 \times 10^{-3}$
ABCC8; rs1048099	0.75 (0.60-0.92);	0.68 (0.54–0.86);	ABCC8; rs722341, rs3819523,	GGA	1.35;	1.48;
,	7.1×10 <sup>-3</sup>	1.0×10 <sup>-3</sup>	rs1048099		5.6×10 <sup>-3</sup>	8.0×10 <sup>-4</sup>
ABCC8; rs722341	0.70 (0.50–0.97);	0.61 (0.43–0.87);	HNF4A; rs1884613, rs2144908	GT	1.47;	1.42;
	0.03	6.2×10 <sup>-3</sup>			$1.1 \times 10^{-3}$	6.2×10 <sup>-3</sup>
ABCC8; rs1799858	1.27 (1.01–1.60);	1.28 (1.00–1.64);	HNF4A; rs1884613, rs2144908	CC	0.64;	0.66;
	0.04	0.05			$1.4 \times 10^{-4}$	9.8×10 <sup>-4</sup>
ABCC8; rs1799854	0.80 (0.64–0.99);	0.80 (0.63–1.00);	KCNJ11, ABCC8; rs5210, rs5215,	GCTGAG	1.33; 0.01	1.29; 0.04
	0.04	0.06	rs5219, rs5222, rs7124355,			
			rs757110			
ABCC8; rs7124355	1.29 (1.03–1.62);	1.24 (0.97–1.58);	NEUROG3; rs3812704,	AG	1.70;	1.82;
	0.03	0.08	rs2812545		3.9×10 <sup>-4</sup>	2.5×10 <sup>-4</sup>
HNF1B; rs1//051//	1.29 (1.03–1.61);	1.36 (1.07–1.73);	<i>PAX6;</i> rs1506, rs3026393,	IIAIA	0.66;	0.71; 0.04
	0.02	0.01	rs3026390, rs2239789,		5.9×10 <sup>-3</sup>	
HNF4A: rs1884613	1.52 (1.21-1.91)	1 47 (1 15–1 88)	PAX6: rs1506_rs3026393	ATATG	1.37.0.02	1.55
	3.1×10 <sup>-4</sup>	2.2×10 <sup>-3</sup>	rs3026390, rs2239789,	,	1107, 0102	3.2×10 <sup>-3</sup>
			rs628224			
HNF4A; rs2144908	1.50 (1.20–1.89);	1.46 (1.15–1.87);				
	<b>4.4</b> ×10 <sup>-4</sup>	2.3×10 <sup>-3</sup>				
INSM1; rs10470054	1.46 (1.04–2.04);	1.38 (0.95–1.99);				
	0.03	0.09				
KCNJ11; rs5219	1.34 (1.07–1.68);	1.28 (1.01–1.63);				
	0.01	0.04				
<i>KCNJ11</i> ; rs5215	1.34 (1.07–1.67);	1.28 (1.01–1.63);				
NEUDOC2	0.01	0.04				
NEURUGS; 153812/04	1.00 (1.20-2.24); 1.0~10-4	1.79 (1.30-2.45); 2.3 \. 10-4				
PAY6. rc628221	$4.9 \times 10^{-1}$	0.74 (0.54 + 1.01)				
1770, 13020224	0.01	0.04-1.01);				
PAX6: rs1506	0.75 (0.58–0.96)	0.81 (0.62–1.05)				
	0.02	0.11				
PDX1; rs2293941	1.42 (1.07–1.89);	1.37 (1.01–1.86):				
	0.01	0.04				

Bold letters indicate significant *P*-values after correcting for multiple comparisons; *P*-value estimates for haplotype associations are based on 10 000 permutations. <sup>a</sup>Adjusted for age, sex and BMI. genes HNF1A, ISL1 and USF1. In non-obese subjects, a strong association with T2D was observed for the HNF4A promoter variants rs1884613 (OR=1.48 (95%CI 1.23-1.77), P=2.96×10<sup>-5</sup>) and rs2144908 (OR=1.48 (95%CI 1.23-1.77), P=2.84×10<sup>-5</sup>) (Table 2). This association was statistically significant even after correcting for age, sex and BMI (OR=1.44 (95%CI 1.19-1.75), P=1.8×10<sup>-4</sup> and OR=1.45 (95%CI 1.20-1.76), P=1.4×10<sup>-4</sup> for rs1884613 and rs2144908 respectively). Strikingly, the polymorphism rs3812704 upstream of NEUROG3 showed suggestive association with T2D in the non-obese group (OR=1.48 (95%CI 1.77-1.87),  $P=8.3\times10^{-4}$ ). This association was statistically significant after adjusting for age, sex and BMI (OR=1.54 (95%CI 1.20-1.97), P=5.7×10<sup>-4</sup>). Apart from these, 11 other polymorphisms in five genes showed nominal association in the non-obese group. As observed earlier, GT and CC haplotypes of HNF4A promoter polymorphisms conferred strong susceptibility and protection, respectively, to T2D in the non-obese subjects (Table 2). A less frequent haplotype (AG) in NEUROG3 encompassing rs3812704, rs2812545 showed increased risk to T2D, while another less frequent GG haplotype conferred protection. Besides these, six haplotypes in three other genes also showed association in this group.

On further subdividing the non-obese group into over-weight (BMI  $\geq$  23 < 25 kg m<sup>-2</sup>) and normal-weight/lean (BMI < 23 kg m<sup>-2</sup>), we observed that significant association of polymorphisms rs1884613 and rs2144908 of *HNF4A* and rs3812704 of *NEUROG3* could be observed only in the normal-weight/lean subjects (OR=1.52 (95%CI 1.21–1.91), *P*=3.1×10<sup>-4</sup>; OR=1.50 (95%CI 1.20–1.89), *P*=4.4×10<sup>-4</sup>; OR=1.68 (95%CI 1.25–2.24), *P*=4.9×10<sup>-4</sup>) respectively). The

association was significant even after correcting for age, sex and BMI for rs3812704 of *NEUROG3* (OR=1.79 (95%CI 1.30– 2.45),  $P=3.3\times10^{-4}$ ). We did not observe any significant gene–gene interactions among the associations observed here. Further, nominal association was observed with 14 quantitative traits analyzed here (Supplementary Table 3).

# DISCUSSION

Here, we conducted an extensive association analysis of 91 polymorphisms in 19 genes involved in pancreatic B-cell development and function with T2D in North Indians. This is by far the largest study conducted in this population considering all these polymorphisms in these genes together using a pathway-based approach. We identified genotypic association in ABCC8, HNF4A, ISL1, KCNJ11 and USF1. Association of HNF4A polymorphisms rs1884613 and rs2144908 remained statistically significant even after correcting for multiple testing. This observation is consistent with observations in many populations emphasizing the role of variants in HNF4A in the manifestation of T2D (Figure 1).<sup>5,6,19-29</sup> Meta-analysis revealed that the trend for the elevated OR, as found here in North Indians, is seen in 10 other populations. Notably, rs1884613 has higher effect size in North Indians (OR=1.35 (95%CI 1.17-1.54)) compared with pooled Scandinavian subjects (OR=1.14 (95%CI 1.06-1.23)), lesser only to Ashkenazi population (OR=1.70 (95%CI 1.23-2.38)).<sup>5,6</sup> Similar results were obtained for rs2144908 that was in a very strong linkage disequilibrium with rs1884613 ( $r^2$ =0.95). BMI-dependent analysis showed that these variants were associated with T2D in non-obese individuals, especially in normal-weight/lean subjects.



**Figure 1** Meta-analysis of *HNF4A* promoter P2 polymorphisms (rs1884613/rs2144908) in 16098 cases and 19567 control subjects. As both the polymorphisms are in strong LD in all the populations, the observations for one polymorphism reflects that of the other. The % weight and OR (95%CI) for combined analysis were calculated using a random-effect model. Heterogeneity was observed between the samples (Q=54.651 on 22 degrees of freedom (P=<0.001). Moment-based estimate of between studies variance was 0.014.

Variants in ABCC8, ISL1, KCNJ11 and USF1 showed suggestive significance. The well-studied ABCC8- S1369A polymorphism (rs757110) and the common non-synonymous polymorphisms in KCNJ11-E23K (rs5219) and I337V (rs5215) showed only nominal association. The haplotype GCTGAG involving polymorphisms in both ABCC8 and KCNJ11 increased risk for T2D. Similarly, ISL1 polymorphisms individually showed nominal association, while their haplotypes showed significant association. This is perhaps the first study to demonstrate the role of ISL1 in T2D. Previous association studies of 3' UTR polymorphism rs3737787 of USF1 with T2D have vielded mixed results (Supplementary Table 4). Although studies in French population and in multi-ethnic groups did not show any association,<sup>30,31</sup> the common allele was associated with an increased risk for T2D in the Chinese population.<sup>32</sup> Similar to Chinese, polymorphisms and haplotypes in USF1 are associated with T2D in North Indians. Further investigations are required to ascertain the role of USF1 in T2D. We did not find association of previously associated WFS1 polymorphism rs10010131 with T2D in North Indians. This could have resulted from the limited power of the study. Our study has a power of 74% to detect associations of polymorphisms with a minor allele frequency of 10% and an OR of 1.3.

Most of the variants in the genes studied showed nominal association in a BMI-dependent analysis. Strikingly, we found statistically significant association of rs3812704 polymorphism of *NEUROG3* in the non-obese group. This association was stronger in normal-weight/ lean subjects, though we did not observe association in the whole study group. This is the first study to demonstrate the association of this variant with T2D in any population. Thus, the variants in *HNFA* and *NEUROG3* indicate a strong predisposition of normal-weight/lean subjects to T2D in North Indians. This suggests increased effect of variants in  $\beta$ -cell development and function genes in a BMI-dependent manner, warranting further studies. In conclusion, we demonstrate a significant role of *HNF4A* and other genes involved in  $\beta$ -cell development and function in the manifestation of T2D in North Indians.

# **CONFLICT OF INTEREST**

The authors declare no conflict of interest.

#### ACKNOWLEDGEMENTS

We thank all the study participants. We thank Mr AK Sharma for his help in sample collection. This study was supported by 'Diabetes mellitus-New drug discovery R&D, molecular mechanisms and genetic & epidemiological factors' (NWP0032-14) funded by Council of Scientific and Industrial Research (CSIR), Government of India.

- Fajans, S. S., Bell, G. I. & Polonsky, K. S. Molecular mechanisms and clinical pathophysiology of maturity-onset diabetes of the young. *N Engl J Med.* 345, 971–980 (2001).
- 2 Yokoi, N., Kanamori, M., Horikawa, Y., Takeda, J., Sanke, T., Furuta, H. et al. Association studies of variants in the genes involved in pancreatic beta-cell function in type 2 diabetes in Japanese subjects. *Diabetes.* **55**, 2379–2386 (2006).
- 3 Tabassum, R., Chavali, S., Dwivedi, O. P., Tandon, N. & Bharadwaj, D. Genetic variants of FOXA2: risk of type 2 diabetes and effect on metabolic traits in North Indians. *J Hum Genet.* 53, 957–965 (2008).
- 4 Chauhan, G., Spurgeon, C. J., Tabassum, R., Bhaskar, S., Kulkarni, S. R., Mahajan, A. et al. Impact of common variants of PPARG, KCNJ11, TCF7L2, SLC30A8, HHEX, CDKN2A, IGF2BP2, and CDKAL1 on the risk of type 2 diabetes in 5,164 Indians. *Diabetes.* 59, 2068–2074 (2010).
- 5 Johansson, S., Raeder, H., Eide, S. A., Midthjell, K., Hveem, K., Sovik, O. *et al.* Studies in 3523 Norwegians and meta-analysis in 11571 subjects indicate that variants in the hepatocyte nuclear factor 4 alpha (HNF4A) P2 region are associated with type 2 diabetes in Scandinavians. *Diabetes.* **56**, 3112–3117 (2007).

- 6 Barroso, I., Luan, J., Wheeler, E., Whittaker, P., Wasson, J., Zeggini, E. *et al.* Population-specific risk of type 2 diabetes conferred by HNF4A P2 promoter variants: a lesson for replication studies. *Diabetes.* **57**, 3161–3165 (2008).
- 7 Jackson, A. E., Cassell, P. G., North, B. V., Vijayaraghavan, S., Gelding, S. V., Ramachandran, A. *et al.* Polymorphic variations in the neurogenic differentiation-1, neurogenin-3, and hepatocyte nuclear factor-1alpha genes contribute to glucose intolerance in a South Indian population. *Diabetes.* **53**, 2122–2125 (2004).
- 8 Bonnefond, A., Froguel, P. & Vaxillaire, M. The emerging genetics of type 2 diabetes. *Trends Mol Med.* 16, 407–416 (2010).
- 9 Wild, S., Roglic, G., Green, A., Sicree, R. & King, H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care.* 27, 1047–1053 (2004).
- 10 American Diabetes Association. 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), S5–20 (2003).
- 11 Mahajan, A., Tabassum, R., Chavali, S., Dwivedi, O. P., Bharadwaj, M., Tandon, N. *et al.* High-sensitivity C-reactive protein levels and type 2 diabetes in urban North Indians. *J Clin Endocrinol Metab.* **94**, 2123–2127 (2009).
- 12 Matthews, D. R., Hosker, J. P., Rudenski, A. S., Naylor, B. A., Treacher, D. F. & Turner, R. C. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia*. 28, 412–419 (1985).
- 13 Shen, R., Fan, J. B., Campbell, D., Chang, W., Chen, J., Doucet, D. et al. Highthroughput SNP genotyping on universal bead arrays. *Mutat Res.* 573, 70–82 (2005).
- 14 Gaudeman, W. J. Candidate gene association analysis for a quantitative trait, using parent-offspring trios. *Genet Epidemiol.* 25, 327–338 (2003).
- 15 Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. **21**, 263–265 (2005).
- 16 Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 81, 559–575 (2007).
- 17 Mahajan, A., Tabassum, R., Chavali, S., Dwivedi, O. P., Chauhan, G., Tandon, N. et al. Obesity-dependent association of TNF-LTA locus with type 2 diabetes in North Indians. J Mol Med. 88, 515–522 (2010).
- 18 WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet* 363, 157–163 (2004).
- 19 Bagwell, A. M., Bento, J. L., Mychaleckyj, J. C., Freedman, B. I., Langefeld, C. D. & Bowden, D. W. Genetic analysis of HNF4A polymorphisms in Caucasian-American type 2 diabetes. *Diabetes.* 54, 1185–1190 (2005).
- 20 Love-Gregory, L. D., Wasson, J., Ma, J., Jin, C. H., Glaser, B., Suarez, B. K. et al. A common polymorphism in the upstream promoter region of the hepatocyte nuclear factor-4 alpha gene on chromosome 20q is associated with type 2 diabetes and appears to contribute to the evidence for linkage in an ashkenazi jewish population. *Diabetes.* 53, 1134–1140 (2004).
- 21 Winckler, W., Graham, R. R., de Bakker, P. I., Sun, M., Almgren, P., Tuomi, T. *et al.* Association testing of variants in the hepatocyte nuclear factor 4alpha gene with risk of type 2 diabetes in 7,883 people. *Diabetes*. **54**, 886–892 (2005).
- 22 Hansen, S. K., Rose, C. S., Glumer, C., Drivsholm, T., Borch-Johnsen, K., Jorgensen, T. et al. Variation near the hepatocyte nuclear factor (HNF)-4alpha gene associates with type 2 diabetes in the Danish population. *Diabetologia*. **48**, 452–458 (2005).
- 23 Bonnycastle, L. L., Willer, C. J., Conneely, K. N., Jackson, A. U., Burrill, C. P., Watanabe, R. M. et al. Common variants in maturity-onset diabetes of the young genes contribute to risk of type 2 diabetes in Finns. Diabetes. 55, 2534–2540 (2006).
- 24 Vaxillaire, M., Dina, C., Lobbens, S., Dechaume, A., Vasseur-Delannoy, V., Helbecque, N. et al. Effect of common polymorphisms in the HNF4alpha promoter on susceptibility to type 2 diabetes in the French Caucasian population. *Diabetologia*. 48, 440–444 (2005).
- 25 Wang, F., Han, X. Y., Ren, Q., Zhang, X. Y., Han, L. C., Luo, Y. Y. et al. Effect of genetic variants in KCNJ11, ABCC8, PPARG and HNF4A loci on the susceptibility of type 2 diabetes in Chinese Han population. *Chin Med J (Engl)*. **122**, 2477–2482 (2009).
- 26 Tanahashi, T., Osabe, D., Nomura, K., Shinohara, S., Kato, H., Ichiishi, E. *et al.* Association study on chromosome 20q11.21-13.13 locus and its contribution to type 2 diabetes susceptibility in Japanese. *Hum Genet.* **120**, 527–542 (2006).
- 27 Muller, Y. L., Infante, A. M., Hanson, R. L., Love-Gregory, L., Knowler, W., Bogardus, C. et al. Variants in hepatocyte nuclear factor 4alpha are modestly associated with type 2 diabetes in Pima Indians. *Diabetes.* 54, 3035–3039 (2005).
- 28 Wanic, K., Malecki, M. T., Wolkow, P. P., Klupa, T., Skupien, J., Bobrek, J. et al. Polymorphisms in the gene encoding hepatocyte nuclear factor-4alpha and susceptibility to type 2 diabetes in a Polish population. *Diabetes Metab.* 32, 86–88 (2006).
- 29 Weedon, M. N., Owen, K. R., Shields, B., Hitman, G., Walker, M., McCarthy, M. I. et al. Common variants of the hepatocyte nuclear factor-4alpha P2 promoter are associated with type 2 diabetes in the UK population. *Diabetes*. 53, 3002–3006 (2004).
- 30 Gibson, F., Hercberg, S. & Froguel, P. Common polymorphisms in the USF1 gene are not associated with type 2 diabetes in French Caucasians. *Diabetes*. 54, 3040–3042 (2005).
- 31 Zeggini, E., Damcott, C. M., Hanson, R. L., Karim, M. A., Rayner, N. W., Groves, C. J. et al. Variation within the gene encoding the upstream stimulatory factor 1 does not influence susceptibility to type 2 diabetes in samples from populations with replicated evidence of linkage to chromosome 1q. *Diabetes.* 55, 2541–2548 (2006).
- 32 Ng, M. C., Miyake, K., So, W. Y., Poon, E. W., Lam, V. K., Li, J. K. et al. The linkage and association of the gene encoding upstream stimulatory factor 1 with type 2 diabetes and metabolic syndrome in the Chinese population. *Diabetologia*. 48, 2018–2024 (2005).

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