ORIGINAL ARTICLE

Genetic variants of *FOXA2*: risk of type 2 diabetes and effect on metabolic traits in North Indians

Rubina Tabassum · Sreenivas Chavali · Om Prakash Dwivedi · Nikhil Tandon · Dwaipayan Bharadwaj

Abstract Here, we examined the association of genetic variants of *FOXA2*, an upstream activator of the β-cell transcription factor network, with type 2 diabetes and related phenotypes in North India. We genotyped three SNPs (rs1212275, rs1055080, rs6048205) and the $(TCC)_n$ repeat polymorphism in 1,656 participants comprising 1,031 patients with type 2 diabetes and 625 controls. SNPs rs1212275 and rs6048205 were uncommon (MAF $< 5\%$) with similar distribution among patients and controls. We found a strong association of $(TCC)_n$ common allele A5 with type 2 diabetes [OR = 1.66 (95% CI 1.36–2.04, $p = 5.9 \times 10^{-7}$) for A5 homozygotes]. Obese individuals with A5A5 genotype had enhanced risk when segregated from normal-weight subjects [OR = 1.92 (95% CI 1.47–2.51), $p = 1.6 \times 10^{-6}$]. A5 was also nominally associated with higher fasting glucose ($p = 0.02$) and lower fasting insulin ($p = 0.0028$) and C-peptide ($p = 0.036$) levels among controls. At the rs1055080 locus, GG was found to provide reduced risk among normal-weight subjects [OR = 0.59 (95% CI 0.40–0.88), $p = 0.011$]. Combination of protective GG and non-risk genotypes of $(TCC)_n$ showed reduced risk of type 2 diabetes both among normal-weight [OR = 0.43 (95% CI 0.29–0.65), $p = 1.2 \times 10^{-6}$] and obese individuals [0.47 (95% CI 0.34–0.64), $p = 4.3 \times 10^{-5}$]. For the first time we demonstrated that *FOXA2* variants may affect risk of type 2 diabetes and metabolic traits in North India, however replication analyses in other cohorts are required to confirm the findings.

Keywords *FOXA2* · Polymorphisms · Type 2 diabetes · Association · North Indians

Introduction

The prevalence of type 2 diabetes has reached epidemic proportions worldwide and has become a major health-care burden, more so because of the associated vascular complications leading to high mortality. India is estimated to show the greatest increase in the prevalence and incidence of type 2 diabetes, and it has been estimated that the number of individuals with diabetes will reach 79 million by the year 2030 (Wild et al. 2000). However, factors affecting susceptibility to diabetes and associated complications in this high-risk group are not clearly understood. Moreover, there is a very slow progress in the identification of genetic determinants of type 2 diabetes, especially in Asian Indians.

Type 2 diabetes is mainly characterized as a state of hyperglycemia resulting from defects in insulin action and β-cell dysfunction. Normal levels of blood glucose are maintained through integrated mechanisms of glucose sensing, insulin production, and glucose utilization. Hence, pancreatic β-cells occupy the central role in maintaining glucose homeostasis, and defects in its development and function can result in imbalances in glucose metabolism and diabetes. A complex network of transcription factors is believed to be involved in pancreas development and function of adult β-cells (St-Onge et al. 1999; Sander et al. 2000). The winged helix/forkhead transcription factor

R. Tabassum · S. Chavali · O. P. Dwivedi · D. Bharadwaj (✉)
Functional Genomics Unit,
Institute of Genomics and Integrative Biology,
CSIR, Mall Road, Delhi 110 007, India
e-mail: db@igib.res.in

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

Foxa2, an upstream activator of the β-cell transcription factor network involving Nkx6.1, Pdx1, HNF1α, HNF1β and HNF4α, might be a key player in the manifestation of type 2 diabetes (Duncan et al. 1998; Wu et al. 1997; Kaestner 2000). Foxa2 is involved in multiple pathways of insulin secretion and plays a vital role in regulating the expression of genes such as SUR1 and Kir6.2, which are essential for maintaining β-cell glucose sensing and glucose homeostasis (Lantz et al. 2004; Wang et al. 2002). Further evidence that it is required for normal β-cell function is provided by β-cell specific ablation of Foxa2 that results in hyperinsulinemic hypoglycemia (Sund et al. 2001). Foxa2 has also been identified as a novel transcriptional regulator of insulin sensitivity, modulating the expression of genes involved in glucose and lipid metabolic pathways that are dysregulated in insulin resistance and type 2 diabetes, hence improving insulin resistance in peripheral tissues (Wolfrum et al. 2004). Additionally, Foxa2 has been suggested to be a crucial inhibitor of adipocyte differentiation and an activator of insulin-sensitizing genes in adipocytes, thus counteracting adipogenesis (Wolfrum et al. 2003).

Despite being an important biological candidate for type 2 diabetes, genetic analysis of *FOXA2* variants has not been conducted extensively in any population. Although mutations in *FOXA2* have been identified in maturity-onset diabetes of the young, these are rare in patients with type 2 diabetes and only the A86T mutation has been suggested as having a possible role in the manifestation of this disease (Yamada et al. 2000; Navas et al. 2000; Zhu et al. 2000). Owing to vital roles of Foxa2 in regulation of insulin secretion and glucose homeostasis, we hypothesized that genetic variants of *FOXA2* may affect metabolic traits and contribute to the manifestation of type 2 diabetes. We therefore examined the association of polymorphisms in *FOXA2* with type 2 diabetes and its related phenotypes in the North Indian population.

Patients and methods

Study population

The study population consisted of a cohort of 1,656 unrelated subjects, comprising 1,031 patients with type 2 diabetes as cases and 625 healthy individuals as controls, from North India belonging to Indo-European ethnicity. Cases included consecutive patients who attended the Endocrinology clinic of the All-India Institute of Medical Sciences, New Delhi, from 2003 to 2007. Type 2 diabetes was diagnosed in accordance with World Health Organization criteria (Expert Committee 2003). The inclusion criteria for control group were:

1. ≥ 40 years of age;
2. HbA1c level $\leq 6.0\%$;
3. fasting glucose level < 110 mg/dL;
4. no family history of diabetes in first and/or second degree relatives; and
5. urban dwellers of Indo-European ethnicity.

Further, at the time of enrolment 30% of the subjects were selected to undergo a 75 g oral glucose-tolerance test (OGTT) at 0 and 120 min to confirm their glucose-tolerance status. The subjects were considered for sampling in the control group only if they had no personal or family history of diabetes or glucose intolerance and OGTT was performed only for those subjects who had no symptoms suggestive of uncontrolled hyperglycemia such as excessive thirst, urination, and hunger. We observed that of all the subjects with normal levels of HbA1c and fasting plasma glucose that underwent OGTT, only 8% of the individuals had impaired glucose tolerance (IGT), and were excluded from the study, and none met diagnostic criteria for type 2 diabetes mellitus. Informed consent was obtained from all the participants. The study was approved by the Ethics Committees of the participating institutions and was in accordance with the principles of the Helsinki Declaration.

Anthropometric and biochemical measurements

Height and weight were measured in light clothes and without shoes. Body mass index (BMI) was calculated as weight in kilograms divided by height in meters squared. Waist circumference was measured in standing position midway between iliac crest and lower costal margin, and hip circumference was measured at its maximum. Waist-to-hip ratio (WHR) was calculated using waist and hip circumferences. Systolic and diastolic blood pressures were measured, using a standard sphygmomanometer, twice in the right arm in the sitting position, after resting for at least 5 min, and the average of the two readings was used.

Venous blood samples for biochemical assessments were obtained from the subjects after 12 h of overnight fasting. Serum levels of glucose, total cholesterol, triglycerides (TG), high-density lipoprotein (HDL), low-density lipoprotein (LDL), urea, uric acid, and creatinine were measured by spectrophotometric methods using the Cobas Integra 400 Plus automated clinical chemistry analyzer (Roche Diagnostics, Mannheim, Germany). Glycated hemoglobin A1c (HbA1c) was determined by low-pressure liquid chromatography (LPLC) on DiaSTAT analyzer (Bio-Rad Laboratories, Richmond, CA, USA). Measurements of plasma levels of insulin and C-peptide were performed by electro-chemiluminescence immunoassay (ECLIA) on an Elecsys 2010 automated immunoanalyzer

(Roche Diagnostics). The homeostasis model assessment of insulin resistance (HOMA-IR) index was calculated as the product of fasting plasma insulin in microunits per milliliter and fasting plasma glucose in millimoles per liter, divided by 22.5 (Matthews et al. 1985).

Identification of polymorphisms

We searched NCBI Variation Database (dbSNP; http://www.ncbi.nlm.nih.gov/SNP) for the identification of polymorphisms in *FOXA2* with minor allele frequency >5% at least in two ethnically different populations. We identified three SNPs based on the above mentioned criteria: SNP1 (G/A, rs1212275) leading to synonymous variation (Q396Q) in exon3, SNP2 (G/A, rs1055080) in 3′UTR, and SNP3 (A/G, rs6048205) in the 3′ flank region. To identify novel base changes, if any, in the North Indian population, we re-sequenced around 5-kb region of *FOXA2* including all the exons, exon–intron boundaries, and the 3′ and 5′ flanking regions in 20 individuals (ten diabetic and ten non-diabetic subjects) using CEQ 8000 Genetic Analysis System (Beckman Coulter, Fullerton, CA, USA). In the sequenced samples, we captured SNPs rs1212275 and rs1055080; however, no additional or novel polymorphisms were identified in our cohort. *FOXA2* also contains a $(TCC)_n$ tri-nucleotide repeat polymorphism in intron1. The $(TCC)_n$ repeat polymorphism and SNPs 1–3 were therefore selected for further studies.

Genotyping

Genomic DNA was isolated from leukocytes by the salting out method. For genotyping of the repeat locus, polymerase chain reaction (PCR) was carried out using a Hot Star Taq polymerase kit (Qiagen) containing Q solution and PCR buffer by following the manufacturer's procedure with 6-carboxyfluorescein (FAM)-labeled forward primer and unlabeled reverse primer. The fluorescence-labeled amplified products were electrophoresed on a ABI Prism 3100 genetic analyzer (Applied Biosystems, Foster City, CA, USA) and genotypes were determined using GeneMapper Software v4.0. Ten percent of samples were randomly selected and re-genotyped as a quality control for genotyping.

Genotyping of the identified SNPs was carried out by single-nucleotide extension of SNP-specific probes using SNaPshot ddNTP Primer Extension Kit. The extended probes were then electrophoresed on an ABI Prism 3100 genetic analyzer and genotypes were determined using GeneMapper. Ten percent of samples were randomly selected and re-genotyped as a quality control for genotyping. The genotype calls were validated by sequencing at least five samples for each genotype. All the PCR primers used for genotyping were designed by Primer3 software (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi) because this has already been shown to provide better primer design for amplification (Chavali et al. 2005).

Statistical analysis

The genotypic distributions at each polymorphic locus in cases and controls were tested for Hardy–Weinberg equilibrium using Genepop software (http://wbiomed.curtin.edu.au/genepop). Linkage disequilibrium between SNP and repeat locus was determined and haplotype analysis was carried out using Haploview 4.0 software (Barrett et al. 2005). Differences in mean and median values of continuous variables (clinical traits), as appropriate, were compared using Student's t test and the Mann–Whitney U test, respectively. Fisher's exact test and χ^2 analyses were employed as appropriate to examine the differences in allelic and genotypic frequencies. In addition, logistic regression analysis was also carried out to analyze the genotype distributions, and odd ratios were calculated after adjusting for age, sex, and BMI. Also, the effects of age, sex, and BMI on the association of the polymorphisms with disease was assessed by the Wald test, as reported earlier (Sladek et al. 2007). Bonferroni correction was applied to correct for the number of statistical tests performed for each allele ($n = 3$) and the number of alleles tested ($n = 2$). Hence, a p value of <0.008 (Bonferroni adjusted $\alpha = 0.05/6$) was considered significant after correcting for multiple comparisons. The uncorrected p values are provided in the text. The statistical power of the study was estimated by using PS power and sample size program (Dupont and Plummer 1997). The statistical analyses were mainly performed using the statistical package SPSS version 15.0 (SPSS, Chicago, IL, USA).

Results

Clinical characteristics of the study population

The descriptive data and comparison of anthropometric and clinical characteristics of the study groups are presented in Table 1. BMI, WHR, and BP were significantly elevated in patients compared with controls (all $p < 0.05$). Significantly higher levels of serum triglycerides, creatinine, urea, and uric acid, and significantly lower levels of total cholesterol and LDL were found in diabetic subjects compared with control subjects (all $p < 0.05$). Low levels of total cholesterol and LDL among diabetic subjects may be attributed to the treatment regime as patients were undergoing drug treatment, especially with statins, at the time of

Table 1 Anthropometric and clinical characteristics of the study cohort

Characteristics	Type 2 diabetic patients	Control subjects	p Value
N (men/women)	1,031 (575/456)	625 (337/288)	
Age (years)	55 (47–63)	50 (43–60)	3.5×10^{-7}
Age at diagnosis of diabetes (years)	45 (39–52)	–	
Duration of diabetes (years)	7 (3–14)	–	
BMI (kg/m^2)			
Men	23.8 (22.0–25.8)	22.7 (19.4–25.3)	1.8×10^{-7}
Women	26.7 (24.1–29.3)	24.2 (20.4–27.8)	9.5×10^{-11}
WHR			
Men	1.0 (0.97–1.03)	0.95 (0.91–0.95)	7.9×10^{-28}
Women	1.0 (0.96–1.03)	0.86 (0.82–0.91)	4.6×10^{-46}
Systolic BP (mmHg)	130 (130–140)	130 (120–140)	3.1×10^{-12}
Diastolic BP (mmHg)	80 (78–90)	80 (70–88)	5.0×10^{-4}
HbA1c (%)	7.7 (6.5–9.1)	5.2 (4.8–5.5)	8.0×10^{-193}
Fasting glucose (mmol/L)	7.7 (6.3–10.0)	4.9 (4.3–5.4)	9.4×10^{-163}
Fasting insulin (μU/mL)	13.7 (7.1–27.1)	6.7 (3.6–11.8)	4.3×10^{-36}
HOMA-IR	5.2 (2.4–9.6)	1.4 (0.6–2.5)	2.7×10^{-71}
C-peptide (ng/mL)	2.7 (1.7–4.1)	1.6 (1.1–2.4)	2.7×10^{-32}
Total cholesterol (mg/dL)	166 (140–197)	176 (149–204)	3.0×10^{-3}
LDL (mg/dL)	101 (81–129)	112 (93–134)	5.5×10^{-7}
HDL (mg/dL)	43 (37–48)	43 (35–52)	0.471
Triglycerides (mg/dL)	135 (96–194)	122 (88–167)	4.9×10^{-4}
Urea (mg/dL)	27.0 (20.5–34.0)	24.6 (19.4–30.0)	4.3×10^{-6}
Uric acid (mg/dL)	4.8 (3.9–6.0)	4.7 (3.8–5.6)	0.038
Creatinine (mg/dL)	0.9 (0.7–1.1)	0.8 (0.7–0.9)	5.2×10^{-28}

N represents the number of individuals. Data are presented as median (interquartile ranges)

BP blood pressure

p values were calculated using the Mann–Whitney U test

recruitment. Patients with type 2 diabetes had higher levels of Insulin and C-peptide and higher index of insulin resistance, as calculated by HOMA-IR, compared with controls (all $p < 0.05$).

Polymorphisms in *FOXA2*

As mentioned earlier, we could not identify any novel polymorphism in the North Indian population upon re-sequencing. At the $(TCC)_n$ locus, we identified 11 alleles, designated A1 to A11, ranging from 193 bp through 223 bp of PCR products. The A5 (205 bp) allele was found to be most common allele in our population with a frequency of 72.8 and 79.4% in controls and patients, respectively. After genotyping 285 cases and 250 controls we found that SNP1 (minor allele frequency, $A = 4.5$ and 3.1% in controls and patients, respectively) and SNP3 (minor allele frequency, $G = 1.2$ and 2.2% in controls and patients, respectively) are uncommon in our population with similar genotype distribution among patients and controls. Hence, these SNPs were not genotyped in the entire cohort and were excluded from further analyses. The minor allele frequencies for SNP2 were 11.0 and 12.7% in controls and diabetic subjects, respectively. For linkage disequilibrium (LD) analysis, we divided the $(TCC)_n$

alleles into two groups, designated A5 and non-A5 alleles (any other allele except A5). We found coexistence of the G allele of SNP2 with the A5 allele of the $(TCC)_n$ locus, suggesting modest linkage disequilibrium between these two loci with D' and r^2 values of 0.80 and 0.29, respectively.

FOXA2 polymorphisms and type 2 diabetes

SNP2 and $(TCC)_n$ repeat polymorphisms were genotyped in the entire cohort of 1,656 individuals. The genotypic distributions of these polymorphisms among patients and control groups were in Hardy–Weinberg equilibrium ($p > 0.05$). Table 2 provides the allelic and genotypic frequencies of *FOXA2* variants. Comparison of allelic frequencies at the $(TCC)_n$ locus showed that allele A5 is significantly more prevalent in diabetic patients (79.4%) than in control subjects (72.8%) with odds ratio of 1.44 (95% CI 1.22–1.70, $p = 1.5 \times 10^{-5}$). With the given frequencies and odds ratio of $(TCC)_n$ alleles, our study had statistical power of 86% to detect an association of $(TCC)_n$ variants with type 2 diabetes. A5 homozygote individuals were found to be more likely to develop type 2 diabetes than other genotypes [OR = 1.66 (95% CI 1.36–2.04), $p = 5.9 \times 10^{-7}$]. In addition, application of logistic

Table 2 Allelic and genotypic distribution of *FOXA2* polymorphisms and their association with type 2 diabetes in the North Indian population

	All diabetic patients	All controls	Normal-weight patients	Normal-weight controls	Obese patients	Obese controls
(TCC)$_n$						
Allele	2n = 2,026	2n = 1,234	2n = 566	2n = 568	2n = 1,390	2n = 648
A1	3 (0.1)	1 (0.1)	3 (0.5)	1 (0.2)	0 (0.0)	0 (0.0)
A2	4 (0.2)	2 (0.2)	0 (0.0)	1 (0.2)	4 (0.3)	1 (0.2)
A3	5 (0.2)	2 (0.2)	2 (0.4)	0 (0.0)	2 (0.1)	2 (0.3)
A4	23 (1.2)	77 (6.2)	3 (0.5)	37 (6.5)	18 (1.3)	39 (6.0)
A5	1,608 (79.4)	898 (72.8)	434 (76.7)	416 (73.2)	1,119 (80.5)	470 (72.5)
A6	108 (5.3)	84 (6.8)	26 (4.6)	38 (6.7)	80 (5.7)	43 (6.6)
A7	43 (2.1)	28 (2.3)	16 (2.8)	13 (2.3)	26 (1.9)	14 (2.2)
A8	127 (6.3)	77 (6.2)	41 (7.2)	30 (5.3)	78 (5.6)	46 (7.1)
A9	48 (2.4)	28 (2.3)	14 (2.5)	15 (2.6)	33 (2.4)	13 (2.0)
A10	47 (2.3)	27 (2.2)	21 (3.7)	11 (1.9)	26 (1.9)	16 (2.5)
A11	10 (0.5)	10 (0.8)	6 (1.1)	6 (1.1)	4 (0.3)	4 (0.6)
OR (95% CI)[b]	1.44 (1.22–1.70)		1.20 (0.92–1.57)		1.56 (1.26–1.94)	
p value	1.5×10^{-5}		0.205		5.2×10^{-5}	
Genotype	n = 1,013	n = 617	n = 283	n = 284	n = 695	n = 324
A5/A5	644 (63.6)	316 (51.2)	167 (59.0)	152 (53.5)	455 (65.5)	161 (49.7)
A5/−[a]	320 (31.6)	266 (43.1)	100 (35.3)	112 (39.4)	210 (30.2)	148 (45.7)
−/−	49 (4.8)	35 (5.7)	16 (5.7)	20 (7.1)	30 (4.3)	15 (4.6)
OR (95% CI)[c]	1.66 (1.36–2.04)		1.25 (0.90–1.74)		1.92 (1.47–2.51)	
p value	5.9×10^{-7}		0.218		1.6×10^{-6}	
SNP2						
Allele	2n = 2,024	2n = 1,242	2n = 560	2n = 572	2n = 1,394	2n = 652
G	1,766 (87.3)	1,105 (89.0)	476 (85.0)	515 (90.0)	1,231 (88.3)	573 (87.9)
A	258 (12.7)	137 (11.0)	84 (15.0)	57 (10.0)	163 (11.7)	79 (12.1)
OR (95% CI)[b]	0.85 (0.68–1.06)		0.63 (0.44–0.90)		0.96 (0.72–1.28)	
p value	0.160		0.011		0.839	
Genotype	n = 1,012	n = 621	n = 280	n = 286	n = 697	n = 326
GG	776 (76.7)	495 (79.7)	206 (73.6)	236 (82.5)	545 (78.2)	251 (77.0)
GA	214 (21.1)	115 (18.5)	64 (22.9)	43 (15.0)	141 (20.2)	71 (21.8)
AA	22 (2.2)	11 (1.8)	10 (3.5)	7 (2.5)	11 (1.6)	4 (1.2)
OR (95% CI)[c]	0.84 (0.66–1.07)		0.59 (0.40–0.88)		1.07 (0.78–1.47)	
p value	0.171		0.011		0.687	

p values presented are uncorrected for multiple testing (Bonferroni adjusted *p* = 0.008). All the comparisons and *p* value calculations presented here were performed using Fisher's exact or χ^2 test as appropriate. Logistic regression was also applied to adjust for age, sex, and BMI but association remained significant after adjustment

n and 2*n* represent the number of individuals and alleles, respectively

In the allele section data are number of alleles (%) and in the genotype section data are number of individuals with a given genotype (%)

[a] − indicates the presence of any allele other than the one examined

[b] Odds ratio (95% confidence interval) of common allele compared with other allele/s

[c] Odds ratio (95% confidence interval) of homozygotes of common allele compared with all other genotypes

regression adjusting for age, sex, and BMI, revealed consistency in the association [OR$_{adj}$ = 1.65 (95% CI 1.34–2.04), *p* = 3.3×10^{-6}]. The distribution of alleles at the SNP2 locus did not differ significantly among patients and controls.

We then segregated the subjects into two groups on the basis of BMI: normal-weight (BMI < 23 kg/m^2) and obese (BMI ≥ 23 kg/m^2) (WHO Expert Consultation 2004). Interestingly, comparison of normal-weight cases vs. normal-weight controls revealed that the distribution

of the susceptible allele A5 does not differ significantly in these two groups ($p = 0.21$), whereas the A5 allele was found to be significantly over-represented among obese patients (80.5%) compared with obese controls (72.5%) with odds ratio of 1.56 (95% CI 1.26–1.94, $p = 5.2 \times 10^{-5}$). Obese individuals with A5A5 genotype had an odds ratio of 1.92 (95% CI 1.47–2.51, $p = 1.6 \times 10^{-6}$) to develop type 2 diabetes compared with other genotypes. Intriguingly, adjusting for age, sex, and BMI, led to an increase in the odds ratio ($OR_{adj} = 1.98$ (95% CI 1.51–2.60, $p = 7.9 \times 10^{-7}$). When we compared the distribution of SNP2, G allele was found to be less frequent among normal-weight patients (85.0%) compared with normal-weight controls (90.0%) with odds ratio 0.63 (95% CI 0.44–0.90, $p = 0.011$). With the given allele frequencies and odds ratio of SNP2, our study had statistical power of 83% to detect an association of SNP2 with type 2 diabetes. Moreover, normal-weight individuals with GG genotype had reduced risk of type 2 diabetes, with odds ratio 0.59 (95% CI 0.40–0.88, $p = 0.011$) compared with other genotypes. Stronger association was observed after adjusting for age, sex, and BMI as covariates with OR_{adj} of 0.50 (95% CI 0.32–0.79, $p = 0.003$). Hence, SNP2 is associated with type 2 diabetes among normal-weight individuals. In contrast, genotypic and allelic distributions of SNP2 were comparable among obese patients and obese controls (78.2% and 77.0%, respectively, for GG; $p = 0.687$). These data suggest the probable role of GG genotype in protection against diabetes among normal-weight individuals.

FOXA2 polymorphisms and quantitative traits

Further, to investigate the effect of *FOXA2* polymorphisms on quantitative metabolic traits, clinical variables of patients and control subjects were compared across the genotypes. We found a nominal association of the $(TCC)_n$ repeat with BMI [median(IQR) = 24.9(22.7–27.9) vs. 24.2(22.2–27.1) kg/m² for A5/A5 vs. A5/− and −/−, $p = 0.024$] in the diabetes group, but no significant differences were observed in other clinical parameters. Among control subjects, association of the $(TCC)_n$ repeat with the levels of fasting glucose [median(IQR) = 5.01(4.41–5.38) vs. 4.84(4.21–5.34) mmol/L, $p = 0.02$], fasting insulin [median(IQR) = 6.1(3.4–10.0) vs. 7.3(3.8–14.5) μU/mL, $p = 0.0028$], and C-peptide [median(IQR) = 1.56(1.05–2.26) vs. 1.80(1.15–2.71) ng/mL, $p = 0.036$] was found when A5/A5 were compared with other genotypes. For SNP2, differences in BMI [median(IQR) = 25.0(22.8–27.8) vs. 24.2(22.0–26.9) kg/m² for GG vs. GA + AA, $p = 0.0058$] among diabetes patients

and in WHR [median(IQR) = 0.90(0.85–0.96) for GG vs. 0.93(0.87–0.97) for GA + AA, $p = 0.013$] and HbA1c levels [median(IQR) = 5.20(4.8–5.5) for GG vs. 5.32(4.9–5.7) for GA + AA, $p = 0.012$] among control subjects were observed.

Haplotype analysis

Haplotype analysis of SNP2 and $(TCC)_n$ revealed that G-A5 haplotype conferred risk (haplotype frequency 0.77 vs. 0.71 for diabetes vs. control, $p = 2.0 \times 10^{-4}$) whereas a G-nonA5 haplotype consisting of G allele of SNP2 and non-A5 alleles of $(TCC)_n$ conferred protection (haplotype frequency 0.10 vs. 0.18 for diabetes vs. control, $p = 4.1 \times 10^{-10}$) for type 2 diabetes. Among obese individuals, the association remained significant for both the haplotypes ($p = 5.0 \times 10^{-4}$ and 1.5×10^{-5}, respectively); among normal-weight subjects, however, only the G-nonA5 haplotype was found to be associated with reduced risk ($p = 5.7 \times 10^{-5}$).

As two polymorphisms were in LD, we also evaluated the primary effects of SNP2 and $(TCC)_n$ independent of each other; these were, interestingly, found to be higher after controlling for the other. The ORs of the GG genotype of SNP2 and the A5A5 genotype of $(TCC)_n$ were 0.52 (95% CI 0.39–0.70, $p = 1.1 \times 10^{-5}$) and 2.16 (95% CI 1.7–2.7, $p = 2.6 \times 10^{-10}$), respectively. The higher independent effects of polymorphisms can be explained by the opposite effects, with SNP2 providing protection and $(TCC)_n$ conferring susceptibility.

Genotype combinations of SNP2 and $(TCC)_n$ repeat and type 2 diabetes

Next, we assessed the additive effect of combined genotypes of SNP2 and $(TCC)_n$ repeat on the risk of type 2 diabetes. Because our analysis showed that the A5A5 genotype of $(TCC)_n$ repeat is susceptible and the GG genotype of SNP2 is protective, we evaluated the effect of combined genotype (GG + A−/−) including protective GG and non-risk genotypes of $(TCC)_n$ (all other genotypes except A5A5) on the risk of type 2 diabetes (Table 3). We found that individuals with this combined genotype had reduced risk of developing type 2 diabetes [OR = 0.44 (95% CI 0.35–0.56, $p = 1.6 \times 10^{-11}$]. Reduced risk of type 2 diabetes was also observed among normal-weight and obese individuals harboring combined genotypes, with odds ratio of 0.43 (95% CI 0.29–0.65, $p = 1.2 \times 10^{-6}$) and 0.47 (95% CI 0.34–0.64, $p = 4.3 \times 10^{-5}$), respectively. The association of the combined genotypes remained significant after adjustment for age, sex, and BMI (Table 3).

Table 3 Frequencies of genotype combination of SNP2 and (TCC)$_n$ repeat and risk of type 2 diabetes in different groups considered in the study

Subjects	Genotype combination		OR (95% CI)[b]	OR$_{adj}$ (95% CI)[c]	p Value	p$_{adj}$ Value
	GG + A −/−[a]	Other combinations				
All diabetics	164 (16.5)	830 (83.5)	0.44 (0.35–0.56)	0.47 (0.37–0.61)	1.6×10^{-11}	4.9×10^{-9}
All controls	189 (30.8)	424 (69.2)				
Normal-weight diabetics	44 (15.9)	233 (84.1)	0.43 (0.29–0.65)	0.36 (0.23–0.56)	1.2×10^{-6}	8.7×10^{-6}
Normal-weight controls	86 (30.5)	196 (69.5)				
Obese diabetics	116 (17.0)	566 (83.0)	0.47 (0.34–0.64)	0.45 (0.33–0.62)	4.3×10^{-5}	6.9×10^{-7}
Obese controls	98 (30.4)	224 (69.6)				

Data are presented as number of individuals, with percentages in parentheses

[a] −/− represents all the genotypes for (TCC)$_n$ other than A5 homozygous

[b] Unadjusted odds ratio (95% confidence interval) calculated using the chi-squared test

[c] Odds ratio adjusted for age, sex, and BMI using logistic regression

p values presented are uncorrected for multiple testing (Bonferroni adjusted $p = 0.008$)

Discussion

Foxa2 is a well known key regulator of pancreas development, adult β-cell function, and insulin sensitivity. Being an important biological candidate, it can be speculated that variations in the *FOXA2* gene may affect the risk of type 2 diabetes. However, very little has yet been explored in this regard. With the ever increasing incidence and prevalence of diabetes, Asian Indians are known to be the highest susceptibility group for type 2 diabetes. Moreover, there is a large pool of subjects with impaired glucose tolerance at a high risk of conversion to diabetes in India (Ramachandran et al. 2001).

A priori information from a public database identified three common SNPs in *FOXA2* (with MAF > 5%). After genotyping, however, two SNPs (rs1212275 and rs6048205) were found to be uncommon (MAF < 5%) in the North Indian population. Hence, we evaluated SNP rs1055080 along with the intronic (TCC)$_n$ repeat polymorphism for association with type 2 diabetes. We found strong association of the (TCC)$_n$ repeat polymorphism with type 2 diabetes in our cohort. Interestingly, we found that, upon segregation of participants into normal-weight and obese groups, obese individuals are at high risk of type 2 diabetes if they are homozygous for A5 whereas it has no effect among normal-weight individuals. Moreover, homozygotes for the A5 allele had higher BMI in type 2 diabetic patients, hence conferring risk for obesity also. Our finding that among control subjects, A5 is associated with higher levels of fasting glucose and lower levels of fasting insulin and C-peptide further suggests the high susceptibility of A5 homozygous individuals to develop type 2 diabetes. In contrast with the (TCC)$_n$ repeat polymorphism, the GG genotype of SNP2 (rs1055080) had protective effect against type 2 diabetes among normal-weight subjects. On the other hand, no difference in the

prevalence of the protective genotype was observed among obese individuals. It is intriguing that although both polymorphisms are in the same gene they have absolutely opposing effects, with the common allele of the (TCC)$_n$ locus increasing risk whereas the common genotype of SNP2 confers protection. Interactive analysis of the protective genotype at SNP2 (GG) and of non-risk genotypes at the (TCC)$_n$ locus (all the genotypes except A5A5 homozygous) revealed reduced risk of type 2 diabetes. Hence, our data suggest that the presence of the protective genotype and the absence of the risk genotype of *FOXA2* reduces susceptibility to type 2 diabetes.

Because of the diversity and heterogeneity of the Indian population, production of false positives in association studies because of population stratification might be plausible. With this in mind, we recruited case and control subjects of Indo-European ethnicity residing in the urban region of North India forming a homogenous cluster in accordance with a recent report of the genetic landscape of the people of India (Indian Genome Variation Consortium 2008). The pattern of clustering of the Indian population groups suggests that the effects of population stratification in disease-association studies may be small, if cases and controls are both drawn from the same cluster. Earlier, a study analyzing Indian genetic variation and diversity also suggested that the effects of population heterogeneity on the production of false positives in association studies might be smaller in Indians than might be expected for such a geographically and linguistically diverse subset of the human population (Rosenberg et al. 2006). However, the possibility of genetic heterogeneity cannot be completely ruled out and hence replication analyses of these *FOXA2* variants in other populations are required to validate the association of the *FOXA2* gene with type 2 diabetes. Also, as revealed through the OGTT test, there may be small proportion of individuals with

impaired glucose tolerance that could have been misclassified as healthy individuals in the control group. Nevertheless it is noteworthy that such a plausible misclassification and inclusion of IGT subjects would actually result in underestimation of the effect of polymorphisms on risk of type 2 diabetes, which implies that we would have observed stronger association if any of those misclassified subjects had been excluded.

This is perhaps the first report suggesting the association of polymorphisms of *FOXA2* with type 2 diabetes. We conclude that the rs1055080 and (TCC)$_n$ repeat polymorphisms of *FOXA2* independently influence the risk of type 2 diabetes and affect metabolic traits, with opposing effects, in North India. Hence, these polymorphisms can serve as predisposition markers for type 2 diabetes in North India for people of Indo-European descent. It is well known that rare variants are of considerable relevance to the type 2 diabetes-allelic spectrum (Sharma et al. 2005). We note that although we have made substantial efforts in this study to identify common and rare variants of *FOXA2* and analyzed for their genetic association, the possibility of rare polymorphisms in this gene affecting the risk of type 2 diabetes cannot be completely ruled out. Therefore, future genetic and functional studies evaluating the role of other genetic variants of the *FOXA2* gene and deciphering the physiological effects and their mechanisms are warranted to further explore the role of *FOXA2* in type 2 diabetes.

References

Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263–265

Chavali S, Mahajan A, Tabassum R, Maiti S, Bharadwaj D (2005) Oligonucleotide properties determination and primer designing: a critical examination of predictions. Bioinformatics 21:3918–3925

Duncan SA, Navas MA, Dufort D, Rossant J, Stoffel M (1998) Regulation of a transcription factor network required for differentiation and metabolism. Science 281:692–695

Dupont WD, Plummer WD (1997) PS power and sample size program available for free on the Internet. Control Clin Trials 18:274

Expert Committee on the Diagnosis and Classification of Diabetes Mellitus (2003) Report of the expert committee on the diagnosis and classification of diabetes mellitus. Diabetes Care 26:S5–S20

Indian Genome Variation Consortium (2008) Genetic landscape of the people of India: a canvas for disease gene exploration. J Genet 87:3–20

Kaestner KH (2000) The hepatocyte nuclear factor 3 (HNF3 or FOXA) family in metabolism. Trends Endocrinol Metab 11:281–285

Lantz KA, Vatamaniuk MZ, Brestelli JE, Friedman JR, Matschinsky FM, Kaestner KH (2004) Foxa2 regulates multiple pathways of insulin secretion. J Clin Invest 114:512–520

Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC (1985) Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia 28:412–419

Navas MA, Vaisse C, Boger S, Heimesaat M, Kollee LA, Stoffel M (2000) The human HNF-3 genes: cloning, partial sequence and mutation screening in patients with impaired glucose homeostasis. Hum Hered 50:370–381

Ramachandran A, Snehalatha C, Kapur A, Vijay V, Mohan V, Das AK, Rao PV, Yajnik CS, Prasanna Kumar KM, Nair JD (2001) High prevalence of diabetes and impaired glucose tolerance in India: National Urban Diabetes Survey. Diabetologia 44:1094–1101

Rosenberg NA, Mahajan S, Gonzalez-Quevedo C, Blum MG, Nino-Rosales L, Ninis V, Das P, Hegde M, Molinari L, Zapata G, Weber JL, Belmont JW, Patel PI (2006) Low levels of genetic divergence across geographically and linguistically diverse populations from India. PLoS Genet 2:e215

Sander M, Sussel L, Conners J, Scheel D, Kalamaras J, Dela Cruz F, Schwitzgebel V, Hayes-Jordan A, German M (2000) Homeobox gene Nkx6.1 lies downstream of Nkx2.2 in the major pathway of beta-cell formation in the pancreas. Development 127:5533–5540

Sharma A, Chavali S, Mahajan A, Tabassum R, Banerjee V, Tandon N, Bharadwaj D (2005) Genetic association, post-translational modification, and protein–protein interactions in type 2 diabetes mellitus. Mol Cell Proteomics 4:1029–1037

Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P, Vincent D, Belisle A, Hadjadj S, Balkau B, Heude B, Charpentier G, Hudson TJ, Montpetit A, Pshezhetsky AV, Prentki M, Posner BI, Balding DJ, Meyre D, Polychronakos C, Froguel P (2007) A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 22:881–885

St-Onge L, Wehr R, Gruss P (1999) Pancreas development and diabetes. Curr Opin Genet Dev 9:295–300

Sund NJ, Vatamaniuk MZ, Casey M, Ang SL, Magnuson MA, Stoffers DA, Matschinsky FM, Kaestner KH (2001) Tissue-specific deletion of Foxa2 in pancreatic beta cells results in hyperinsulinemic hypoglycemia. Genes Dev 15:1706–1715

Wang H, Gauthier BR, Hagenfeldt-Johansson KA, Iezzi M, Wollheim CB (2002) Foxa2 (HNF3-β) controls multiple genes implicated in metabolism-secretion coupling of glucose-induced insulin release. J Biol Chem 277:17564–17570

WHO Expert Consultation (2004) Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet 363:157–163

Wild S, Roglic G, Green A, Sicree R, King H (2000) Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. Diabetes Care 27:1047–1053

Wolfrum C, Asilmaz E, Luca E, Friedman JM, Stoffel M (2004) Foxa2 regulates lipid metabolism and ketogenesis in the liver during fasting and in diabetes. Nature 432:1027–1032

Wolfrum C, Shih DQ, Kuwajima S, Norris AW, Kahn CR, Stoffel M (2003) Role of Foxa2 in adipocyte metabolism and differentiation. J Clin Invest 112:345–356

Wu KL, Gannon M, Peshavaria M, Offield MF, Henderson E, Ray M, Marks A, Gamer LW, Wright CV, Stein R (1997) Hepatocyte

nuclear factor 3 beta is involved in pancreatic beta-cell specific transcription of the pdx-1 gene. Mol Cell Biol 17:6002–6013

Yamada S, Zhu Q, Aihara Y, Onda H, Zhang Z, Yu L, Jin L, Si YJ, Nishigori H, Tomura H, Inoue I, Morikawa A, Yamagata K, Hanafusa T, Matsuzawa Y, Takeda J (2000) Cloning of cDNA and the gene encoding human hepatocyte nuclear factor (HNF)-3

beta and mutation screening in Japanese subjects with maturity-onset diabetes of the young. Diabetologia 43:121–124

Zhu Q, Yamagata K, Yu L, Tomura H, Yamada S, Yang Q, Yoshiuchi I, Sumi S, Miyagawa J, Takeda J, Hanafusa T, Matsuzawa Y (2000) Identification of missense mutations in the hepatocyte nuclear factor-3β gene in Japanese subjects with late-onset type II diabetes mellitus. Diabetologia 43:1197–1200