

# Association of TCF7L2 Gene Polymorphisms with T2DM in the Population of Hyderabad, India

Kommoju Uma Jyothi<sup>1</sup>, Maruda Jayaraj<sup>1</sup>, Kadarkarai Samy Subburaj<sup>1</sup>, Kotla Jaya Prasad<sup>2</sup>, Irgam Kumuda<sup>1</sup>, Velaga Lakshmi<sup>3</sup>, Battini Mohan Reddy<sup>1\*</sup>

**1** Molecular Anthropology Group, Biological Anthropology Unit, Indian Statistical Institute, Hyderabad, India, **2** JP Endocrine Center, Ashok Nagar, Hyderabad, India, **3** Department of Human Genetics, Andhra University, Visakhapatnam, India

## Abstract

We attempt to evaluate the nature of association of TCF7L2 gene variants with T2DM, for the first time in the population of Hyderabad, which is considered to be diabetic capital of India. It is a case-control study of the three SNPs of TCF7L2, rs7903146, rs12255372 and rs11196205, genotyped on Sequenom Massarray platform, in a sample of 758 patients and 621 controls. The risk allele frequency of the three SNPs was found to be significantly higher in the T2DM cases than controls, implicating susceptibility for diabetes ( $p < 0.01$ ). The greatest risk of developing the disease was conferred by rs7903146. Further, the logistic regression of genotypes of each SNP under log additive model, and the haplotypes constituted by at least one of the three risk alleles also show significantly greater risk of developing T2DM when compared to the wild type haplotype. Further, BMI and WHR emerge as significant covariates with confounding effects. The strong association of the TCF7L2 SNPs with T2DM is consistent with the findings among other Indian and Non-Indian populations, suggesting universal phenomena of its association across ethnic groups globally, both within and outside the Indian subcontinent, albeit the functional relevance of these SNPs needs yet to be established.

**Citation:** Uma Jyothi K, Jayaraj M, Subburaj KS, Prasad KJ, Kumuda I, et al. (2013) Association of TCF7L2 Gene Polymorphisms with T2DM in the Population of Hyderabad, India. PLoS ONE 8(4): e60212. doi:10.1371/journal.pone.0060212

**Editor:** Qingyang Huang, Central China Normal University, China

**Received:** November 27, 2012; **Accepted:** February 22, 2013; **Published:** April 5, 2013

**Copyright:** © 2013 Uma Jyothi et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by the Indian Statistical Institute. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: bmrissi@gmail.com

## Introduction

Type 2 diabetes mellitus (T2DM) is the most common form of diabetes characterized by hyperglycemia, which is caused by impairment in both insulin secretion and action. It is a chronic disease leading to various complications such as coronary heart disease, diabetic nephropathy, neuropathy and retinopathy. It is becoming an epidemic with increasing prevalence throughout the world. Both genetic and environmental factors play a strong role in the manifestation of this complex genetic disorder. Unique in genetic pre-disposition of its population to diabetes, coupled with rapid urbanization, India is one of the leading countries presenting with largest number of diabetics (62.4 million). This number is increasing at an alarming rate and expected to touch 100 million by the year 2030 [1,2]. The development of high throughput micro-array platforms unraveled a large number of genes involved in the etiology of T2DM. Among them, TCF7L2 (*Transcription factor 7 like 2*) is considered as one of the most important candidate genes which plays a major role in blood-glucose homeostasis and beta cell function. Strong association of TCF7L2 with T2DM was initially found in Icelandic population which has been subsequently replicated in Danish and U.S populations [3]. The three TCF7L2 SNPs (rs7390146, rs12255372 and rs11196205) that were strongly associated with T2DM in the above study were subsequently replicated, along with other SNPs of TCF7L2, in a huge meta-analysis [4] prompting their inclusion in any future replication effort.

The first GWAS study on T2DM in the French population showed strong signal for TCF7L2 [5]. Its association, along with the other T2DM genes, was also confirmed in subsequent GWAS studies. The consistency in the findings of its association with T2DM observed among many studies of diverse ethnic groups was considered indicative of a universal contribution of this gene to T2DM [6], albeit a few studies showed weak or no association with T2DM [7–12], which can be attributed to the extremely low frequency of the risk alleles and inadequate sample size, hence lack of power of the study.

The replicative studies of this candidate gene among a couple of Indian populations also showed strong association of TCF7L2 with T2DM but given enormous ethnic, cultural, geographic and genetic heterogeneity, large population size and relatively high prevalence of diabetes, very small and insignificant number/proportion of Indian populations were hitherto studied [13–17]. Further, these studies were confined to select regions like Pune, Punjab, Haryana, Himachal Pradesh, Delhi, Jammu & Kashmir and Chennai. Most of these populations, except that of Chennai come from northern parts of India and belong to the Indo-European linguistic family and considered to be quite distinct from those of Southern India- ethnically, linguistically and genetically [18]. Therefore, there is an urgent need to screen populations of the other Indian regions in order to project a representative scenario for the Indian subcontinent as a whole. It is also necessary to study the effect of environment and ethnicity on the T2DM susceptible genes to better appreciate the ‘Asian Indian phenotype’

that is uniquely predisposed to develop the disease. The Hyderabad city, in particular, which is considered as diabetic capital of India with high prevalence of T2DM, has not been explored and to the best of our knowledge, this is the first genetic study on T2DM in the population of Hyderabad or Andhra Pradesh. We present here the results of our analysis of the three SNPs of TCF7L2 (rs7903146 - intron 3, rs12255372 and rs11196205 - intron 4) for their association with T2DM.

These three SNPs, individually and/or as a group, were found to be associated with T2DM in a wide spectrum of world populations: Finnish [19], Amish [20], U.S [21], Polish [22], Scandinavian [23], Chinese [24], Brazil [25], Japanese [26–28], Swedish [29], U.K [30], Dutch [31], Palestinian [32], Iranian [33,34], Arab [35], Tunisian [36], Persian [37] as well as Indian populations [13–17]. Of the three SNPs, rs7903146 is the most consistent in several GWAS studies [5,38–45] as well as most replicated variant of TCF7L2. We present here the results of our analyses of association of the above three SNPs of TCF7L2 gene with T2DM in a large sample of patients and controls from the population of Hyderabad, which is considered as the diabetic capital of India.

## Materials and Methods

This study is based on a total sample of 1379 subjects (758 T2DM cases and 621 controls) from Hyderabad, Andhra Pradesh. The T2DM patients were recruited from J.P Endocrine Center, Hyderabad, based on their medical record and adhering to ADA 2010 criteria [46], which stipulates fasting plasma glucose  $\geq 126$  mg/dl and/or 2 hr plasma glucose  $\geq 200$  mg/dl or casual plasma glucose (random blood sugar)  $\geq 200$  mg/dl. We conducted a free diabetic camp in different organizations in Hyderabad and individuals above the age of 40 years were tested for random blood sugar (where fasting and 2 hr plasma glucose were not feasible) and those with random blood sugar values  $< 140$  mg/dl were recruited as controls. We collected background data on age at onset of T2DM, duration of T2DM since diagnosis, family history, and anthropometric measurements like height, weight, waist and hip circumferences and other parameters such as blood pressure of the cases. The above four anthropometric measurements were taken for both the cases and controls in addition to the family history, lifestyle parameters and health status.

About 3–5 ml blood was collected from both T2DM cases and controls. We isolated DNA using the phenol-chloroform method [47], which was quantified with the help of NanoDrop. The three SNPs of the TCF7L2 viz., rs7903146, rs12255372 and rs11196205, were genotyped on Sequenom Massarray platform [48] at The Centre for Genomic Application (TCGA), Delhi, during the year 2011. The raw data files generated by Mass Array Sequenom were analysed for the intensity peaks of calibrant to ascertain the quality of the data. An overall call rate of  $> 95\%$  was maintained. For every 96 samples (a quadrant of Sequenom chip), 4 samples were duplicated and the call rates were checked for concordance. The calls in the negative control (no DNA) was also monitored in all the runs.

Statistical analyses were performed with the help of SPSS statistical software (version 18.0, IBM SPSS). Data for continuous variables were expressed as mean  $\pm$  SD, median and range. Chi-square test was used to compare the allele and the genotype frequencies between the cases and the controls. Using ‘SNP association’ package of R PROGRAM, logistic regression analysis was done under the assumption of different genetic models viz., co-dominant, dominant, over-dominant, recessive and log additive, to explore the most parsimonious one that fits our data in case

of each of the three SNPs. The above logistic regression procedure was repeated after adding BMI (Body Mass Index) and WHR (Waist to Hip Ratio) as covariates. Post-hoc power of the study was calculated using G\*Power software (version 3.1). The Hardy-Weinberg equilibrium was determined through goodness of fit  $\chi^2$  test using Pypop software. Haploview and THESIAS softwares were used to estimate LD and generate haplotype frequencies.

## Ethics Statement

The study was approved by the Indian Statistical Institute Review Committee for Protection of Research Risks to Humans and written informed consent was obtained from all the participants.

## Results

The clinical profile of the subjects, represented by mean, median and range of different variables such as age of the subjects, age at onset of the disease, BMI, WHR, FPG, PPG etc is presented in Table S1. Over all, the sample, albeit not purposive, represents a relatively greater proportion of male subjects as compared to females, in the ratio of  $\sim 3:2$ . The case and control samples represent similar age range, mean and median values. However, the mean and median values of BMI, WHR and RBG of the cases were much higher than the controls and highly significantly different ( $p < 0.01$ ) from them.

Tables 1 and 2 show allele and genotype frequencies of the three SNPs of TCF7L2 considered for this study. Using PyPop software, the observed genotype counts were compared with those expected under Hardy-Weinberg proportions (HWP) and the  $\chi^2$  test suggest that all the three SNP loci conform to the Hardy-Weinberg equilibrium in both case and control samples. The risk allele frequency for each of the three SNPs – **T** of rs7903146 (C/T), **C** of rs11196205 (G/C) and **T** of rs12255372 (G/T) - was found to be significantly higher in the T2DM cases than the normal controls, suggesting that they may confer risk for diabetes. The heterozygote and the risk homozygote frequencies were also higher in T2DM cases when compared to the controls.

The logistic regression yielded significant odds ratio (Table 1) suggesting that the risk allele confers significant risk for developing T2DM with reference to each of the three SNPs: rs 7903146 (OR 1.88,  $p < 0.001$ ); rs11196205 (OR 1.23,  $p = 0.011$ ) and rs12255372 (OR 1.50,  $p < 0.001$ ). The genotypic OR from Logistic regression under additive model also showed significant association (Table 2) of rs7903146 and rs12255372 (OR = 1.86,  $p = < 0.001$  and OR = 1.54,  $p = < 0.001$ , respectively) with T2DM. Although the second SNP (rs11196205) also showed significant association (OR = 1.24,  $p = 0.001$ ), it was not as strong as in case of the other two SNPs.

Logistic regression analysis of T2DM on the genotypes of each of the three SNPs was repeated by using BMI and WHR as covariates under additive model to explore if the observed association of the SNPs with T2DM is not due to the confounding nature of association of the covariates with the genotypes as well as the disease phenotype. The results of this analysis (Table 2) suggest that, after adjusting for covariates, there has been virtually no change in the status of association, especially in case of SNP genotypes pertaining to rs7903146 and rs11196205 implying that the genetic association observed with reference to TCF7L2 SNP genotypes is independent of the effect of covariates. However, after removing the covariate effects, the strength of association was diminished in case of the third SNP rs12255372. Given large sample size, our study shows high statistical power ( $1 - \beta$  error probability) of 99–100%. However, the mean BMI did not

**Table 1.** Allelic frequency distribution and allelic OR from the Logistic regression of T2DM on the risk allele frequency.

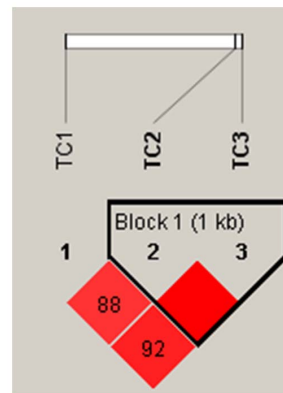
SNP	Allele	Cases(N= 758)	Controls(N= 621)	$\chi^2$	O.R (95% C.I)	p-value
rs7903146	C	0.67	0.79	46.9	<b>1.89 (1.57–2.26)</b>	<0.001
	T	<b>0.33</b>	<b>0.21</b>			
rs11196205	G	0.61	0.65	6.51	<b>1.23 (1.05–1.45)</b>	<b>0.011</b>
	C	<b>0.39</b>	<b>0.35</b>			
rs12255372	G	0.76	0.82	17.93	<b>1.51 (1.25–1.82)</b>	<0.001
	T	<b>0.24</b>	<b>0.18</b>			

doi:10.1371/journal.pone.0060212.t001

significantly vary among genotypes, either in cases or control presented in Table S2 albeit, as expected, significant mean difference in BMI was evident ( $p < 0.01$ ) between cases and controls with reference to each genotype of the respective SNPs. We did not find significant association of age and gender as covariates with T2DM. In an effort to validate these results in the same population and to check internal consistency of our results we have split our sample into 30%, 50% and 70% random subsets of cases and controls and repeated the analysis (i.e. allele, genotype frequencies, Logistic regression of respective alleles and logistic regression of genotypes under additive model of all the three SNPs). The results showed similar association pattern with T2DM with reference to both alleles (Table S3) and genotypes (Table S4) in each of the three subsets as compared to the total sample, showing internal consistency.

On the other hand, while rs11196205 and rs12255372 are in perfect LD with  $D' = 1$ ,  $r^2 = 0.46$  the other two combinations of rs7390146: rs11196205 and rs7390146: rs12255372 also show strong LD with  $D'$  of 0.88 and 0.92 (Figure 1) and with  $r^2$  value of 0.53 and 0.58, respectively. The 121, 221 and 222 haplotypes, with risk alleles, show relatively greater difference in frequency between cases and controls as compared to the others (Tables 3&4). The logistic regression analysis, using the most prevalent wild haplotype (111 i.e. CGG) as the reference, revealed significant OR for 121, 221 and 222 haplotypes (Table 3); where as 121 is protective, 221 and 222 are risk conferring in nature. However, after Bonferroni correction for multiple testing, only the two risk conferring haplotypes (222 and 221) were retained as significantly associated with T2DM. It is indeed pertinent to note that the most

strongly associated and risk conferring haplotype, 222, is a combination of only risk alleles of the three SNPs, the haplotype 221 represents risk alleles of two of the three SNPs (rs7390146 and rs11196205). A similar pattern of association was apparent when logistic regression was performed using BMI and WHR as covariates (Table 4), albeit with marginally increased strength in case of all risk alleles 222 and the protective 121 haplotypes, which retain significance even after Bonferroni correction.



**Figure 1.** Linkage disequilibrium plot for TCF7L2 SNPs in the pooled sample of T2DM cases and controls. (\*TC1 = rs7903146; TC2 = rs11196205, TC3 = rs12255372)  $D'$  values are mentioned in the LD blocks.

doi:10.1371/journal.pone.0060212.g001

**Table 2.** Genotypic frequency distribution of the three SNPs of TCF7L2 in T2DM cases and controls.

SNP	Genotype	Cases (N = 758)	Controls (N = 621)	$\chi^2$ p-value	*OR	95%C.I	p-value	adjusted for BMI and WHR		
								*OR	95%C.I	p-value
rs7903146	CC	0.45	0.63	45.96(<0.001)	<b>1.86</b>	<b>1.54–2.23</b>	<0.001	<b>1.85</b>	<b>1.53–2.24</b>	<0.001
	CT	<b>0.43</b>	<b>0.31</b>							
	TT	<b>0.11</b>	<b>0.05</b>							
rs11196205	GG	0.36	0.43	6.755(0.034)	<b>1.24</b>	<b>1.05–1.46</b>	<b>0.01</b>	<b>1.23</b>	<b>1.03–1.45</b>	<b>0.02</b>
	GC	<b>0.49</b>	<b>0.46</b>							
	CC	<b>0.15</b>	<b>0.12</b>							
rs12255372	GG	0.57	0.68	18.20(<0.001)	<b>1.50</b>	<b>1.24–1.82</b>	<0.001	<b>1.48</b>	<b>1.22–1.81</b>	<0.001
	GT	<b>0.37</b>	<b>0.28</b>							
	TT	<b>0.05</b>	<b>0.04</b>							

\*Odds ratios from Logistic regression under additive model.

doi:10.1371/journal.pone.0060212.t002

**Table 3.** Haplotype frequencies in T2DM cases and controls and the results of haplotype based logistic regression analysis.

Haplotype	Haplotype code	Cases	Controls	#OR (95% C.I)	p≤
CGG	111	0.58	0.65	- -	-
CCG	121	<b>0.08</b>	<b>0.12</b>	<b>0.73 0.55-0.96</b>	<b>0.03</b>
CCT	122	0.01	0.01	0.62 0.29-1.34	0.23
TGG	211	0.02	0.01	1.98 1.00-3.91	0.05
TCG	221	<b>0.07</b>	<b>0.05</b>	<b>1.68 1.17-2.42</b>	<b>0.01*</b>
TCT	222	<b>0.24</b>	<b>0.16</b>	<b>1.62 1.32-1.99</b>	<b>0.001*</b>

#Haplotypic OR is in comparison to the reference haplotype 111, which is constituted by the wild type alleles of the three SNPs and shows in maximum frequency.

\*Significant even after Bonferroni Correction.

doi:10.1371/journal.pone.0060212.t003

The risk allele frequency of the three SNPs, rs7903146, rs11196205 and rs12255372 observed in our population and the range of allele frequency for other Indian and Non-Indian populations are presented in Table S5. Broadly speaking, it is within the range observed for Indian populations albeit at the lower end. Further, the range of risk allele frequency falls within the range for non-Indian populations. Although the very few studies hitherto conducted on vastly heterogeneous population of India precludes scope for any meaningful meta-analysis of the data generated on the three SNPs, we have compiled population and region specific data on allele frequency and odds ratio to gauge the possible heterogeneity in the allele frequency and the nature and magnitude of its association with T2DM (Table 5). Unfortunately, while 6 populations were studied for rs7903146 (two from south [17], two from North [14,15] and two independent populations from Pune [13,14]), only one more population (besides ours from Hyderabad) from North [15] for rs11196205 and three more for rs12255372 were screened representing south [17], north [15] and Pune [13] from mid-west India. Nevertheless, the test of homogeneity of allele frequency across Indian populations so far studied suggest significant heterogeneity of allele frequency for each of the three SNPs in the pooled sample of cases and controls (rs7903146:  $\chi^2 = 62.01$ ,  $df = 5$ ,  $p < 0.0001$ ; rs11196205:  $\chi^2 = 44.5$   $df = 1$ ,  $p < 0.0001$ ; rs12255372:  $\chi^2 = 62.4$   $df = 3$ ,  $p < 0.0001$ ). The heterogeneity observed is, to a large extent, because of the differences in allele frequency between the ethnically, linguistically

as well as geographically distinct northern and southern populations (Table 5). These results were qualitatively similar when cases and controls were separately analyzed. This would suggest the nature of genetic heterogeneity implicit in the complex Indian population underscoring the importance of exploring a number of geographically and ethnically heterogeneous populations before one can project representative picture on the genetic susceptibility of Indian population as a whole.

**Discussion**

TCF7L2 gene which spans about 215.9 Kb with 17 exons is a high mobility group box containing transcription factor involved in the Wnt signaling pathway, playing a key role in cell development and regulatory mechanisms. WNT activity is important for lipid and glucose metabolism, pancreatic beta cell proliferation and function and for production of incretin hormone GLP-1 [49]. WNT signaling is also critical for glucagon like peptide-1 (GLP-1) secretion by the intestinal endocrine L-cells [50]. Thus, alteration in this pathway could lead to reduced secretion of GLP-1 which, in turn, could have effect on insulin secretion. The GLP-1, in concert with insulin, plays a critical role in blood glucose homeostasis. It has been postulated that TCF7L2 gene variants may confer susceptibility to T2DM indirectly by altering GLP-1 levels [51]. It was predicted that in the carriers of TCF7L2 risk alleles, the Wnt signaling is potentially increased and thereby the pre-adipocytes are not differentiated into the mature adipocytes and probably influence the growth and development of adipose tissue and alter BMI [51,52]. Thus TCF7L2 might play some role in the adipogenesis resulting in the deposition of triglycerides in peripheral tissues leading to insulin resistance [20].

Consistent with the functional relevance of this gene, the results of our study are confirmatory to the earlier findings among the other Indian [13-17] as well as non-Indian populations [3,19-31] in that they show strong association of TCF7L2 variants with increased risk of T2DM. In general, the association of TCF7L2 with T2DM is reiterated in our population not only with reference to risk alleles of the SNPs studied but also in case of genotypes of each SNP (under additive model) and haplotypes derived from the three SNPs which are in strong LD. Thus our results reiterate TCF7L2 as the most promising T2DM susceptible gene that has been most universally replicated. While the risk allele of each of the three TCF7L2 SNPs was highly significantly associated with T2DM, the greatest risk of developing the disease was conferred by rs7903146, which is consistent with the findings from other

**Table 4.** Haplotype frequency of T2DM Cases and Controls and the results of logistic regression of T2DM on the haplotypes, using BMI and WHR as covariates.

Haplotype	Haplotype Code	Covariate	Cases	Controls	#OR (95% C.I)	p value
		<b>BMI</b>			<b>1.10 1.08-1.12</b>	<b>&lt;0.01</b>
		<b>WHR</b>			<b>0.01 0.002-0.014</b>	<b>&lt;0.01</b>
CGG	111		0.58	0.65	- -	-
CCG	121		<b>0.07</b>	<b>0.12</b>	<b>0.68 0.50-0.92</b>	<b>0.01*</b>
CCT	122		0.01	0.01	0.62 0.29-1.32	0.21
TGG	211		0.02	0.01	2.09 0.99-4.41	0.05
TCG	221		<b>0.07</b>	<b>0.05</b>	<b>1.54 1.04-2.26</b>	<b>0.03</b>
TCT	222		<b>0.23</b>	<b>0.15</b>	<b>1.60 1.29-1.98</b>	<b>0.000021*</b>

#Haplotypic OR is in comparison to the reference haplotype 111, which is constituted by the wild type alleles of the three SNPs and shows in maximum frequency.

\*Significant even after Bonferroni Correction.

doi:10.1371/journal.pone.0060212.t004

**Table 5.** Comparative presentation of Risk allele frequency (RAF), OR and C.I of the three SNPs of TCF7L2 gene among Indian populations.

Reference	Indian Populations	Sample size		rs7903146(C/*T)			rs1196205(G/C)			rs12255372(G/T)		
		Cases	Controls	RAF	OR(p-value)	C.I	RAF	OR(p value)	C.I	RAF	OR(p value)	C.I
<i>Our study</i>	<b>Hyderabad</b>	758	621	0.33	1.89(<0.001)	1.57–2.26	0.39	1.23(0.01)	1.05–1.45	0.24	1.51(<0.001)	1.25–1.82
<i>Chandak et al (2007) [13]</i>	<b>Pune</b>	955	399	0.37	1.46(3 × 10 <sup>-5</sup> )	1.22–1.75	-	-	-	0.30	1.50(4 × 10 <sup>-5</sup> )	1.24–1.82
<i>Bodhini et al (2007) [17]</i>	<b>Chennai</b>	1031	1038	0.33	1.29(0.0001)	1.13–1.48	-	-	-	0.23	1.30(0.001)	1.11–1.51
<i>Sanghera et al (2008) [15]</i>	<b>Khatri Sikhs - North India (Punjab, Haryana, Himachal Pradesh, Delhi, Jammu &amp; Kashmir)</b>	556	536	0.41	— <sup>†</sup> 1.39(0.010)	— <sup>†</sup> 1.08–1.79	0.52	— <sup>†</sup> 1.44(0.01)	— <sup>†</sup> 1.09–1.92	0.37	#NS <sup>‡</sup> 1.19	0.93–1.52
<i>Chauhan et al (2010) [14]</i>	<b>Delhi</b>	1019	1006	0.40	1.67(1.7 × 10 <sup>-13</sup> )	1.46–1.92	-	-	-	-	-	-
	<b>Pune</b>	1467	1672	0.37	2.10(1.7 × 10 <sup>-17</sup> )	1.77–2.49	-	-	-	-	-	-
	<b>combined (Delhi+Pune)</b>	2486	2678	0.38	1.89(4 × 10 <sup>-34</sup> )	1.71–2.09	-	-	-	-	-	-

\*Risk allele highlighted in bold.

<sup>†</sup>Non-significant.

<sup>‡</sup>Genotypic adjusted OR,C.I and p value under dominant model only for this study as allelic odds ratio was not available.

doi:10.1371/journal.pone.0060212.t005

Indian studies [13–17] as well as migrant Indians living in other countries [53]. Further, the risk conferred by homozygotes is much higher than the heterozygote carriers and it is marked in case of rs7903146. Thus our study not only shows consistency with the earlier findings in case of the association of risk alleles but also conforms to the pattern of multiplicative model of inheritance hypothesized by Grant et al (2006) in which the risk conferred by homozygotes was higher than that of the heterozygotes for all the SNPs. In somewhat a concurrent pattern, the three SNPs of TCF7L2, which are located on the chromosome 10q25.3, show strong LD and the logistic regression analysis based on haplotype frequency suggest that haplotype 222 which carries all the three risk alleles of the SNPs (rs7903146, rs1196205 and rs12255372) is most significantly associated with T2DM risk as compared to the others with either one or two of the risk alleles.

The confounding nature of the influence of environmental factors in the manifestation of T2DM is confirmed by the significant association of BMI and WHR as covariates in the Logistic Regression Analysis of T2DM on genetic variants of TCF7L2, albeit a highly significant association was evident for rs7903146 and rs1196205 even after adjusting for the effect of covariates. However, the other SNP (rs12255372) shows somewhat decreased risk, after adjusting for covariates, probably suggesting the confounding effect of environmental factors to be relatively greater with reference to this gene. The mean difference of BMI between cases and controls was highly significant for each genotype of respective SNPs, indirectly suggesting greater body fat of T2DM patients, in general, albeit inter genotype differences in mean BMI was not evident for any of the three SNPs. Despite TCF7L2 being the most promising T2DM susceptible gene, the precise mechanism of action of TCF7L2 SNPs in the etiology of T2DM is still unclear as all the SNPs of TCF7L2 identified so far are found in the intronic regions. It is necessary to understand how these intronic SNPs affect the expression of TCF7L2. Interestingly, none of the variants found in the exonic regions were associated with the T2DM. Earlier studies focusing on the function of TCF7L2 could trace expression to the gut rather than to the pancreatic islets. Although Lyssenko et al (2007) showed an increased gene expression of TCF7L2 in risk allele carriers in the pancreatic islets, certain controversies still exist with the increased or decreased expression of this gene. Therefore, research focusing on the functional relevance of this gene is of importance and more extensive studies are needed to study whether the variants of TCF7L2 are involved in the alternative splicing, gene expression or protein structure.

Given that the population of this region was not genetically explored hitherto for any of the complex genetic disorders, T2DM in particular, it is necessary to assess the role of different candidate genes in the etiology of T2DM in Andhra Pradesh and other regions of India with high prevalence of T2DM, to gauge the heterogeneity in the susceptible genetic profile of the Indian population, in general, which is characterized by enormous geographic, ethnic, cultural and genetic heterogeneity. We consider that our study contributes to this end significantly, even if it appears to only reiterate the oft observed strong association of the TCF7L2 SNPs with the risk of developing T2DM in the diverse populations of the world, reiterating universality of its association across ethnic groups. Further, our study could have also helped in understanding the effect of ethnicity, if existent, on the T2DM susceptible genes.

## Supporting Information

**Table S1 Clinical profile of the T2DM cases and controls.** Footnote: \*Numbers in the parentheses are the SDs. NA = Not Applicable and #NI = No information (DOC)

**Table S2 Mean BMI according to genotypes of TCF7L2 SNPs (rs7903146, rs11196205, rs12255372) and the values of t test for mean difference between cases and controls.** Footnote: P values <0.01, \*\*0.001 and \*\*\*<0.001. F values are not significant for all the SNPs (not mentioned in the table) (DOC)

**Table S3 Subset analysis (30%,50%,70%) of T2DM cases and controls for the TCF7L2 SNPs (rs7903146, rs11196205, rs12255372) showing allelic frequency distribution and OR using logistic regression.** (DOCX)

**Table S4 Subset analysis (30%,50%,70%) of T2DM cases and controls for the three SNP's (rs7903146, rs11196205, rs12255372) showing genotype frequencies and genotypic OR using logistic regression under additive model.**

(DOCX)

**Table S5 Risk Allele frequency of SNPs rs7903146, rs11196205 and rs12255372 of TCF7L2 in the population of A.P and among other Indian and Non-Indian populations.** Footnote: Risk Allele frequency is presented as range for the other Indian and Non-Indian populations. The RAF which is unusually low in the ##Japanese and ##Chinese samples is not included in the range. \*Genomic position of the respective SNPs on chromosome 10 and allelic nomenclature is according to NCBI dbSNP Build 37. \*\*Risk Alleles marked as bold. \*\*\*Odds ratio range specified for only Indian populations (DOC)

## Acknowledgments

The authors thank Director, Indian Statistical Institute for logistic support. We also thank all the subjects who participated in the study.

## Author Contributions

Conceived and designed the experiments: KUJ BMR. Performed the experiments: KUJ KSSR MJR IK. Analyzed the data: KUJ KSSR MJR. Contributed reagents/materials/analysis tools: JPK BMR. Wrote the paper: KUJ BMR.

## References

- Anjana RM, Pradeepa R, Deepa M, Datta M, Sudha V, et al. (2011) Prevalence of diabetes and prediabetes (impaired fasting glucose and/or impaired glucose tolerance) in urban and rural India: phase I results of the Indian Council of Medical Research-India DIABetes (ICMR-INDIAB) study. *Diabetologia* 54: 3022–3027.
- Shetty P (2012) India's diabetes time bomb. *Nature* 485: s14–16.
- Grant SF, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, et al. (2006) Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet* 38: 320–323.
- Tong Y, Lin Y, Zhang Y, Yang J, Zhang Y, et al. (2009) Association between TCF7L2 gene polymorphisms and susceptibility to type 2 diabetes mellitus: a large Human Genome Epidemiology (HuGE) review and meta-analysis. *BMC Med Genet* 10: 15.
- Sladek R, Rocheleau G, Rung J, Dina C, Shen L, et al. (2007) A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 445: 881–885.
- Cauchi S, El Achhab Y, Choquet H, Dina C, Krempler F, et al. (2007) TCF7L2 is reproducibly associated with type 2 diabetes in various ethnic groups: a global meta-analysis. *J Mol Med* 85: 777–782.
- Chang YC, Chang TJ, Jiang YD, Kuo SS, Lee KC, et al. (2007) Association study of the genetic polymorphisms of the transcription factor 7-like 2 (TCF7L2) gene and type 2 diabetes in the Chinese population. *Diabetes* 56: 2631–2637.
- Guo T, Hanson RL, Traurig M, Muller YL, Ma L, et al. (2007) TCF7L2 is not a major susceptibility gene for type 2 diabetes in Pima Indians: analysis of 3,501 individuals. *Diabetes* 56: 3082–3088.
- Ren Q, Han XY, Wang F, Zhang XY, Han LC, et al. (2008) Exon sequencing and association analysis of polymorphisms in TCF7L2 with type 2 diabetes in a Chinese population. *Diabetologia* 51: 1146–1152.
- Kifagi C, Makni K, Boudawara M, Mnif F, Hamza N, et al. (2011) Association of Genetic Variations in TCF7L2, SLC30A8, HHEX, LOC387761, and EXT2 with Type 2 Diabetes Mellitus in Tunisia. *Genet Test Mol Biomarkers* 15: 399–405.
- Zheng X, Ren W, Zhang S, Liu J, LIS, et al. (2012) Association of type 2 diabetes susceptibility genes (TCF7L2, SLC30A8, PCSK1 and PCSK2) and proinsulin conversion in a Chinese population. *Mol Biol Rep* 39: 17–23.
- Park SE, Lee WY, Oh KW, Baek KH, Yoon KH, et al. (2012) Impact of common type 2 diabetes risk gene variants on future type 2 diabetes in the non-diabetic population in Korea. *J Hum Genet* 57: 265–268.
- Chandak GR, Janipalli CS, Bhaskar S, Kulkarni SR, Mohankrishna P, et al. (2007) Common variants in the TCF7L2 gene are strongly associated with type 2 diabetes mellitus in the Indian population. *Diabetologia* 50: 63–67.
- Chauhan G, Spurgeon CJ, Tabassum R, Bhaskar S, Kulkarni SR, et al. (2010) 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.
- Sanghera DK, Nath SK, Ortega L, Gambarelli M, Kim-Howard X, et al. (2008) TCF7L2 polymorphisms are associated with type 2 diabetes in Khatri Sikhs from North India: genetic variation affects lipid levels. *Ann Hum Genet* 72: 499–509.
- Sanghera DK, Ortega L, Han S, Singh J, Ralhan SK, et al. (2008) Impact of nine common type 2 diabetes risk polymorphisms in Asian Indian Sikhs: PPARG2 Pro12Ala, IGF2BP2, TCF7L2 and FTO variants confer a significant risk. *BMC Med Genet* 9: 59.
- Bodhini D, Radha V, Dhar M, Narayani N, Mohan V (2007) The rs12255372(G/T) and rs7903146(C/T) polymorphisms of the TCF7L2 gene are associated with type 2 diabetes mellitus in Asian Indians. *Metabolism* 56: 1174–1178.
- Reich D, Thangaraj K, Patterson N, Price AL, Singh L (2009). Reconstructing Indian population history. *Nature* 461: 48 9–494.
- Scott LJ, Bonnycastle C, Willer CJ, Sprau AG, Jackson AU, et al. (2006) Association of transcription factor 7-like 2 (TCF7L2) variants with type 2 diabetes in a Finnish sample. *Diabetes* 55: 2649–2653.
- Damcott CM, Pollin TI, Reinhard LJ, Ott SH, Shen H, et al. (2006) Polymorphisms in the transcription factor 7-like 2 (TCF7L2) gene are associated with type 2 diabetes in the Amish: replication and evidence for a role in both insulin secretion and insulin resistance. *Diabetes* 55: 2654–2659.
- Zhang C, Qi L, Hunter DJ, Meigs JB, Manson JE, et al. (2006) Variant of transcription factor 7-like 2 (TCF7L2) gene and the risk of type 2 diabetes in large cohorts of U.S. women and men. *Diabetes* 55: 2645–2648.
- Moczulski D, Gawlik B, August R, Strojek K, Grzeszczak W. (2007) TCF7L2 gene is associated with type 2 diabetes in Polish population. *Exp Clin Diabetologia* 7: 109–111.
- Lyssenko V, Lupi R, Marchetti P, Del Guerra S, Orho-Melander M, et al. (2007) Mechanisms by which common variants in the TCF7L2 gene increase risk of type 2 diabetes. *J Clin Invest* 117: 2155–2163.
- Ng MC, Tam CHT, Lam VK, So WY, Ma RC, et al. (2007) Replication and identification of novel variants at TCF7L2 associated with type 2 diabetes in Hong Kong Chinese. *J Clin Endocrinol Metab* 92: 3733–3737.
- Marquezine GF, Pereira AC, Sousa AG, Mill JG, Hueb WA, et al. (2008). TCF7L2 variant genotypes and type 2 diabetes risk in Brazil: significant association, but not a significant tool for risk stratification in the general population. *BMC Med Genet* 9: 106.
- Hayashi T, Iwamoto Y, Kaku K, Hirose H, Maeda S. (2007) Replication study for the association of TCF7L2 with susceptibility to type 2 diabetes in a Japanese population. *Diabetologia* 50: 980–984.
- Horikoshi M, Hara K, Ito C, Nagai R, Froguel P, et al. (2007) A genetic variation of the transcription factor 7-like 2 gene is associated with risk of type 2 diabetes in the Japanese population. *Diabetologia* 50: 747–751.
- Kunika K, Tanahashi T, Numata S, Ueno S, Ohmori T, et al. (2008) Common coding variant in the TCF7L2 gene and study of the association with type 2 diabetes in Japanese subjects. *J Hum Genet* 53: 972–982.
- Mayans S, Lackovic K, Lindgren P, Ruikka K, Agren A, et al. (2007) TCF7L2 polymorphisms are associated with type 2 diabetes in northern Sweden. *Eur J Hum Genet* 15: 342–346.
- De Silva NM, Steele A, Shields B, Knight B, Parnell K, et al. (2007) The transcription factor 7-like 2 (TCF7L2) gene is associated with Type 2 diabetes in UK community-based cases, but the risk allele frequency is reduced compared with UK cases selected for genetic studies. *Diabet Med* 24: 1067–1072.

31. van Vliet-Ostapchouk JV, Shiri-Sverdlow R, Zhernakova A, Strengman E, van Haefen TW et al. (2007) Association of variants of transcription factor 7-like 2 (TCF7L2) with susceptibility to type 2 diabetes in the Dutch Breda cohort. *Diabetologia* 50: 59–62.
32. Ereqat S, Nasereddin A, Cauchi S, Azmi K, Abdeen Z, et al. (2010) Association of a common variant in TCF7L2 gene with type 2 diabetes mellitus in the Palestinian population. *Acta Diabetol* 47(Suppl): 195–198.
33. Amoli MM, Amiri P, Tavakkoly-Bazzaz J, Charmchi E, Hafeziyeh J, et al. (2010) Replication of TCF7L2 rs7903146 association with type 2 diabetes in an Iranian population. *Genet Mol Bio* 33: 449–451.
34. Alami FM, Ahmadi M, Bazrafshan H, Tabarraei A, Khosravi A, et al. (2012) Association of the TCF7L2 rs12255372 (G/T) variant with type 2 diabetes mellitus in an Iranian population. *Genet Mol Bio* 35: 413–417.
35. Mtiraoui N, Turki A, Nemr R, Ehtay A, Izzidi I, et al. (2012) Contribution of common variants of ENPP1, IGF2BP2, KCNJ11, MLXIPL, PPAR $\gamma$ , SLC30A8 and TCF7L2 to the risk of type 2 diabetes in Lebanese and Tunisian Arabs. *Diabetes Metab* 38: 444–449.
36. Ezzidi I, Mtiraoui N, Cauchi S, Vaillant E, Dechaume A, et al. (2009) Contribution of type 2 diabetes associated loci in the Arabic population from Tunisia: a case-control study. *BMC Med Genet* 10: 33.
37. Palizban A, Nikpour M, Salehi R, Maracy MR. (2012) Association of a common variant in TCF7L2 gene with type 2 diabetes mellitus in a Persian population. *Clin Exp Med* 12: 115–119.
38. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, et al. (2007) Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* 316: 1336–1341.
39. Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, et al. (2007) Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 316: 1331–1336.
40. Steinthorsdottir V, Thorleifsson G, Reynisdottir I, Benediktsson R, Jonsdottir T, et al. (2007) A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nature Genetics* 39: 770–775.
41. Salonen JT, Uimari P, Aalto JM, Pirskanen M, Kaikkonen J, et al. (2007) A Type 2 diabetes whole-genome association study in four populations: The DiaGen consortium. *Am J Hum Genet* 81: 338–345.
42. Timpson NJ, Lindgren CM, Weedon MN, Randall J, Ouwehand WH, et al. (2009) Adiposity-related heterogeneity in patterns of type 2 diabetes susceptibility observed in genome-wide association data. *Diabetes* 58: 505–510.
43. Takeuchi F, Serizawa M, Yamamoto K, Fujisawa T, Nakashima E, et al. (2009) Confirmation of Multiple Risk Loci and Genetic Impacts by a Genome-Wide Association Study of Type 2 Diabetes in the Japanese Population. *Diabetes* 58: 1690–1699.
44. Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, et al. (2010) Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet* 42: 579–589.
45. Kho AN, Hayes MG, Rasmussen-Torvik L, Pacheco JA, Thompson WK, et al. (2012) Use of diverse electronic medical record systems to identify genetic risk for type 2 diabetes within a genome-wide association study. *J Am Med Inform Assoc* 19: 212–218.
46. American Diabetes Association (ADA). (2010) Diagnosis and classification of Diabetes mellitus. *Diabetes Care* 33 (Suppl 1): 62–69.
47. Sambrook J, Fritschi EF, Maniatis T. (1989) *Molecular cloning: a laboratory manual*. New York: Cold Spring Harbor Laboratory Press.
48. Gabriel S, Ziaugra L, Tabbaa D. (2009) SNP genotyping using the Sequenom Massarray iPLEX Platform. *Curr Protoc Hum Genet* 60: 2.12.1–2.12.16.
49. T Jin. (2008) The WNT signalling pathway and diabetes mellitus. *Diabetologia* 51: 1771–1780.
50. Yi F, Brubaker PL, Jin T. (2005) TCF-4 mediates cell type-specific regulation of proglucagon gene expression by beta-catenin and glycogen synthase kinase-3 $\beta$ . *J Biol Chem* 280: 1457–1464.
51. Smith U. (2007) TCF7L2 and type 2 diabetes—weWNT to know. *Diabetologia* 50: 5–7.
52. Gupta V, Khadgawat R, Saraswathy KN, Sachdeva MP, Kalla AK. (2008) Emergence of TCF7L2 as a most promising gene in predisposition of Diabetes Type II. *Int J Hum Genet*, 8: 199–215.
53. Rees SD, Bellary S, Britten AC, O'Hare JP, Kumar S, et al. (2008) Common variants of the TCF7L2 gene are associated with increased risk of type 2 diabetes mellitus in a UK-resident South Asian population. *BMC Med Genet* 9: 8.