SUMMARY

Type 2 Diabetes (T2D) is a major health concern in today's world with an inevitable fact that it predominates over the other types of diabetes, affecting up to 90-95% of diabetic individuals (Sierra et al., 2009). According to the International Diabetes Federation (IDF) 2011, there are at least 285 million people affected with T2D and number is expected to touch the figure of 438 million worldwide by the end of year 2030. The Asia alone accounts for 60% of the world’s diabetic population out of which India and China are two major contributors (Chan et al., 2009). Due to adaptation of western dietary habits and sedentary lifestyle, India is currently experiencing an epidemic of T2D and is often referred to as the diabetes capital of the world with more than 50.8 million diabetic subjects in India (IDF., 2011) and number is expected to rise to 69.9 million by the year 2025 (Mohan et al., 2009).

The pathophysiology of diabetes mellitus mainly revolves around two main hormones namely insulin and glucagon. The negative feedback mechanism of these hormones regulates the glucose level and thus maintains blood glucose homeostasis in the body (Scheen, 2003). The malfunction of these two hormones can lead to the impairment of different pathways involved in T2D. Apart from this, being a complex disease, T2D involves interplay of various factors namely anthropometric, physiometric, clinical, demographic parameters and genetic factors (Ramachandran et al., 2008). Genetics play a major role in complex disease like T2D and also provides an insight into the possible mechanisms involving the influence of genetic loci on disease susceptibility. The molecular genetics has revealed the role of various candidate genes causing T2D in various population using different approaches such as linkage studies, candidate gene approach, positional cloning, genome wide association studies (Permutt et al., 2005). Gloyn et al., 2001 were first to report the association of KCNJ11 gene with T2D by candidate gene approach in Caucasian population. It has been assumed that variants in KCNJ11 gene could result in abnormal expression and regulation of the KATP channel and thus affecting the first phase of insulin secretion. The polymorphism E23K (rs5219) in KCNJ11 has shown strong signal of association with T2D in almost every European population and less information about this polymorphism is available in Indian
population (Chauhan et al., 2010; Gupta et al., 2010). The other gene showing the strongest association so far with T2D is TCF7L2. TCF7L2 is involved in cell proliferation and in adipogenesis, myogenesis, and pancreatic islet development. It also activates the genes encoding intestinal proglucagon and glucagon-like peptides-1 and -2 via wnt signaling pathway which establishes the link of TCF7L2 with T2D. Variations in TCF7L2 are associated with impaired insulin secretion and increased hepatic glucose production. Association of TCF7L2 gene with T2D was first reported by Grant et al., (2006) in Iceland population by linkage study. There after this gene was replicated in almost every European population to find its role with T2D. Despite comprehensive genotyping efforts across TCF7L2 gene locus, intron 3 single nucleotide polymorphism rs7903146 T allele has been found consistently associated with T2D in individuals of European ancestry as reviewed by Dehwah et al., (2009) however, scanty information is available in Indian population (Bodhini et al., 2007; Chandak et al., 2007; Sanghera et al., 2008; Chauhan et al., 2010; Gupta et al., 2010).

The wave of GWAS and the shift from GWAS to GWAS meta-analysis has raised the number of independent loci showing significant associations with T2D from 19 to 44 in various populations across the world at the end of year 2011 (Wheeler and Barroso, 2011). With the advent of technology the genome wide association studies (GWAS) has subsequently documented success in finding novel loci associated with T2D (Saxena et al., 2012) but looking at the overall outcomes of all GWAS and GWAS meta-analysis only three genes TCF7L2, KCNJ11, PPARG (Visscher et al., 2012) belonging to insulin secreting pathway are believed to be strong putative candidate genes for T2D.

In addition to nuclear genome, mitochondrial genome also serves multiple essential cellular functions apart from energy generation via oxidative phosphorylation (OXPHOS) which support various metabolic reactions of human body and the hypothesis that prominent features of T2D are caused by mitochondrial dysfunction (Lowell and Shulman, 2005). Defects in mitochondria genome either through excessive ROS production or through the imbalance in the ATP/ADP ratio is also believed to cause T2D. The 10398 G>A polymorphism in MT-ND3 is believed to increase ROS
production in stress condition like hyperglycemic which could be one of the reasons in development of T2D (Bhat et al., 2007).

TCF7L2 and KCNJ11 gene has shown positive association with T2D in almost every European population (Gloyn et al., 2003; Cauchi et al., 2007). The variants in this gene contribute more powerfully to the risk of developing T2D than any other gene discovered (Zeggini and McCarthy, 2006).

Given the background of diversity and endogamous nature of the population and the scanty information of the polymorphisms, rs5219 of KCNJ11, rs7903146 of TCF7L2, rs2853826 of MT-ND3 in Indian population, the present case control study was focused on Punjab to

- Explore frequency of polymorphisms of KCNJ11 (rs5219), TCF7L2 (rs7903146), MT-ND3 (rs2853826) genes in cases and controls from the studied population groups.

- Analyse the association of various risk factors such as BMI, dietary pattern and physical activity with development of T2D.

- Apply statistical tools to find that, the above mentioned polymorphisms along with anthropometric parameters impart any risk towards development of T2D in endogamous groups of Punjab.

- Explore SNP–SNP interaction if any, to understand the broader aspect of T2D development.

The present case control study included 1813 subjects (859 cases and 954 controls) belonging to seven endogamous population groups (Bania, BC, Brahmin, Jat Sikh, Khatri, Rajput and SC). This sample size was attained by calculating power of this study (>95) by using PS software (www.vanderbilt.biostats.edu). Data collection was done by recording the details of demographic variables in proformas prepared after thorough perusal of literature. 3ml blood sample was collected in pre-labeled 0.5M EDTA vials with informed consent and after approval from ethical committee of Guru Nanak Dev University, Amritsar and that of Jawaharlal Nehru University, New-Delhi.
For candidate gene association analyses, samples were collected randomly for various regions of Punjab under supervision of resident doctors and diabetologist.

The criteria for selection of patients was based on American diabetes association (ADA), 2007 criteria (FPG $\geq$ 126mg/dL or OGTT $\geq$ 200mg/dL). T2D subjects with history of ketoacidosis / requiring continuous insulin treatment since diagnosis / having exocrine pancreatic disease / having exceptionally early age of onset, severe liver or renal dysfunction were excluded from the study.

For sampling of controls, the individuals with normal blood glucose levels (FPG $<$ 99mg/dl or OGTT $<$ 139mg/dl) and having no first degree relatives with positive family history of T2D were considered as controls. Only age, gender and ethnicity matched controls were included in the study.

After obtaining informed consent, the blood samples were collected randomly from hospitals, clinics and by arranging medical camps in various villages and towns with the help of resident doctors. 3ml of venous blood was collected from individuals and immediately transferred to prelabeled 0.5M EDTA (anti-coagulant) vials. All samples were transported from place of collection to the laboratory on ice to laboratory and stored at $-20^\circ$C for further processing. Anthropometric measurements were taken following standard procedures and protocols. Estimation of the caste and sub-castes was made by asking questions about surnames and sub-castes or gotras of all the subjects. Other parameters which were relevant for the studies such as family history of T2D, age of onset, duration of T2D and type of medication was also carefully noted. Genomic DNA was isolated from whole blood using standard phenol chloroform method with minor changes according to lab conditions (Kunkel et al., 1977). The isolated genomic DNA was quantified using dual beam UV spectrophotometer. Quantified DNA was then diluted to a concentration of 50ng/$\mu$l of each sample.

Screening of rs7903146 polymorphism in TCF7L2 gene was done by automated DNA sequencing technique, whereas, screening of rs5219 in KCNJ11 and rs2853826 in MT-ND3 gene were analyzed by restriction fragment length polymorphism (RFLP) and were confirmed by automated DNA sequencing. Data analysis was done by using
appropriate statistical tests (SPSS software package ver. 20). The statistical software package SPSS (IBM SPSS statistics version 20.0) was used for most of the statistical analyses) The statistical power of this study (>80) was assessed using PS software (www.vanderbilt.biostats.edu). Student’s t test to compare mean values of clinical parameters among cases and controls. Chi square ($\chi^2$) test, was used to compare allelic as well as genotypic frequencies of patients and control subjects. Once a significant overall difference between patients and control subjects was detected ($p \leq 0.05$), the odds ratio (OR) and population attributable risk (PAR) with 95% CIs were employed to assess the strength of association between polymorphism and T2D risk. Association under three different types of genetic models was also estimated, namely dominant model (M+H vs. W), co-dominant model (H vs. W+ M) and recessive model (M vs. W+H) for TCF7L2 and KCNJ11 gene polymorphisms. The unconditional logistic regression was used for mitochondrial SNP to find association for correction for age, sex and BMI as well as to explore the interaction of different genotypic combinations in the studied genes.

The present study, a total of 1813 samples comprising of 859 T2D cases and 954 non-diabetic healthy controls, belonging to seven population groups (Bania, BC, Brahmin, Jat Sikh, Khatri, Rajput and SC) of Punjab, were screened for polymorphisms in candidate genes. The polymorphisms and genes selected for the study were TCF7L2 rs7903146 (C/T); KCNJ11 E23K (G/A) (rs5219); MT-ND3 10398 (G/A) (rs2853826) associated with insulin signalling and secretion pathways. Interaction of all three polymorphisms in pooled population group of Punjab and in the background of rs2853826 polymorphism of the MT-ND3 gene was also explored for cumulative effect of all these polymorphism towards susceptibility of T2D in population of Punjab.

The results of the present study are categorized into following sections (a) Comparison of clinical parameters (b) association of risk factors like BMI and physical activity among cases and controls in endogamous groups and pooled population of Punjab (c) molecular analysis of polymorphism TCF7L2 (rs7903146), (d) molecular analysis of polymorphism KCNJ11 (rs5219) in endogamous groups and pooled population of Punjab, (e) molecular analysis of polymorphism MT-ND3 (rs2853826) in endogamous
groups and pooled population of Punjab (f) Interaction analysis of polymorphisms of TCF7L2 (rs7903146), KCNJ11 (rs5219), MT-ND3 (rs2853826) in pooled population of Punjab

**Summary**

(a) Comparison of clinical parameters among cases and control

The overall obesity was observed higher in endogamous groups as well as in pooled population of Punjab. The higher mean value was observed in cases than controls as depicted by BMI in [BCs (28.17 vs 24.52, p≤0.001), Brahmins (27.03 vs 24.61, p≤0.001), Jat Sikhs (26.72 vs 24.49, p≤0.001)] and no association of BMI with T2D was found in Banias, Rajputs, Khatris and SCs but when analysis was carried out on pooled samples a significantly higher mean values of BMI were observed (26.76 vs 25.2, p≤0.001).

Higher mean values of fasting plasma glucose levels was observed in cases of all endogamous groups than controls [BCs (153.43 vs 87.57, p ≤0.001), Brahmins (181.52 vs 99.5, p ≤0.001), Jat Sikhs (161.93 vs 98.04, p ≤0.001), Khatris (166.74 vs 83.35, p ≤0.001), Rajputs (142.29 vs 85.12, p ≤0.001), SCs (216.53 vs 114.8, p ≤0.001)]. Also significant mean value differences were observed for random blood glucose levels in all endogamous groups. When analysis for FBS and RBS was carried out in pooled population, statistically significant FBS mg/dL mean values were observed higher in cases (163.43 ±61.87) than in controls (91.47 ±19.56). Similarly, statistically significant (p≤0.001) random plasma glucose levels were observed in cases (229.13±95.93) and controls (108.65±32.84).

The significant number of cases studied had positive family history in pooled population as well as in endogamous groups.

(b) Association of BMI and physical activity and dietary pattern among cases and controls

T2D is a lifestyle disorder with various risk factors contributing towards development of disease. Risk factors such as high BMI (≥23 Kg/m² in Indian population) and
physical activity have been associated with development of T2D (Snehalatha et al., 2003). In endogamous groups some interesting trends were observed.

The association of BMI was observed in BC [p=1.3x10^{-3}, OR-2.74 (1.46-5.15)], Jat Sikhs [p=2.05x10^{-7}, OR-2.77 (1.87-4.10)] and pooled population [p=2.18x10^{-6}, OR-1.67 (1.35-2.08)] of Punjab, whereas, significance of association in Brahmins could not be attained after regression analysis (p=0.22). The risk of increasing BMI in these groups was due to adoption of sedentary life style and poor dietary habits.

Physical activity was found to be associated in Banias [OR-9.71 (4.39-21.28)], Jats Sikhs, [OR- 3.50 (2.49 -5.18)], Khatris [OR-2.49 (1.41-4.41), Rajputs [p=0.007, OR-3.36 (1.38-8.26)] and pooled population [p=6.85x10^{-10}, OR-1.86 (1.53-2.27)], whereas, dietary pattern was associated Jat Sikhs [$\chi^2=5.37$, p=0.02, OR-1.47x10^{-10} (1.06-2.04)] in Sikhs and Brahmins [$\chi^2=6.74$, p=0.009, OR-0.44 (0.24-0.87)], whereas no other group was found to be associated with T2D.

Association of different combination of life style factors increase the risk of T2D development in different group

(c) Molecular analysis of polymorphisms in TCF7L2 (rs7903146) in endogamous groups and pooled population of Punjab.

TCF7L2 finds its role in T2D by activating many genes downstream of the wnt signalling pathway involved in various metabolic pathways . The formation of complex between TCF7L2 and $\beta$-catenin leads to nuclear translocation and transcription of number of genes including intestinal proglucagon (Prunier et al., 2004; Yi et al., 2005) and thus controls incretin axis, hepatic glucose production and adipocyte function (Lyssenko et al., 2007; Cauchi et al., 2006) which leads to impaired insulin secretion

In molecular data analysis, TCF7L2 (rs7903146 C>T) was found associated with T2D in Brahmins [p=2.1x10^{-2}, OR- 2.65 (1.15-6.10)] under recessive model (TT vs CT+CC).

In pooled population of Punjab, genetic model analysis revealed TT genotype under recessive model [p=4.9x10^{-4}, OR-1.73(1.27-2.35)] and CT+TT genotype combination under dominant model [p=1.4x10^{-3}, OR-1.35(1.12-1.62)] significantly increased risk of
developing T2D. No association was observed between rs7903146 with T2D in Bania, BC, Khatri and Rajput population group. However, the frequency of minor allele T was observed to be consistently higher in case as compared to their respective control groups.

(d) Molecular analysis of polymorphisms in KCNJ11 (rs5219) in endogamous groups and pooled population of Punjab

Statistically significant association of rs5219 (KCNJ11) polymorphism under dominant \([p=4.3\times10^{-7}, OR=3.93 (2.31-6.67)]\) and co-dominant \([p= 8.1\times10^{-7}, OR=3.62 (2.19-6.16)]\) genetic models was observed in SC population of Punjab. GA+AA genotype combination and AA genotype under dominant and co-dominant provided increased risk in T2D development in SCs of Punjab. Association under all the three genetic models, dominant \([\chi^2 =22.7, p=2.1\times10^{-6}, OR=1.62 (1.32-1.98)]\), recessive \([\chi^2 =5.68, p=1.8\times10^{-2}, OR=1.63 (1.08-2.44)]\) and co-dominant \([\chi^2 =25.1, p=5.7\times10^{-4}, OR=1.40 (1.16-1.69)]\) was observed in population of Punjab. A most significant association was found under dominant and co-dominant than recessive genetic model.

(e) Molecular analysis of polymorphisms in MT-ND3 (rs2853826) in endogamous groups and pooled population of Punjab.

MT-ND3 (rs2853826) showed significant association in BCs \((p=0.01)\), Jats Sikhs \((p=0.03, OR= 1.44 (1.02-2.04))\), Khatris \((p=0.004, OR= 2.48 (1.34-4.57))\), Rajputs \((p=0.009, OR= 4.10 (1.42-11.8))\) and SCs \((p<0.0001, OR= 5.0 (2.66-9.36))\) and different distribution of genotype and allele frequencies between cases and controls was observed. No significant difference in risk allele frequency distribution was observed between cases and controls in Bania and Brahmin endogamous groups of Punjab.

(e) Interaction analysis of TCF7L2 (rs7903146), KCNJ11 (rs5219) and MT-ND3 (rs2853826)

Genotype interaction analysis was performed between polymorphism rs5219, rs7903146, rs2853826 of KCNJ11, TCF7L2 and MT-ND3. Analysis revealed that the
Presence of risk genotype CT of rs5219, risk genotype of TT rs7903146 and risk genotype of AA rs2853826 provided highly significant risk towards T2D \([p=0.028, 1.8(1.07-3.04)]\), whereas, genotype CC of rs5219, CT of rs7903146, GG of rs2853826 of KCNJ11, TCF7L2 and MT-ND3 genes respectively provided significant protection to controls \([p=0.003, 0.52(0.34-0.80)]\).

Our results confirm the association of all the polymorphism rs7903146 (TCF7L2), rs5219 (KCNJ11), rs2853826 (MT-ND3) play an important role in the development of T2D in population of Punjab but when analyzed differential pattern of association was observed.

The interaction results also unveiled the role of multigenic interaction in providing increased risk or protection against T2D. The risk alleles of all studied polymorphisms increased risk by 1.8 fold in developing T2D in pooled population of Punjab.

The cumulative effect of both environment and genetic factors along with role of ethnicity cannot be ruled out to understand the complex etiology of multifactorial disease, T2D.

In conclusion, the result of our studies contribute a bit in finding the role of most replicated and strongly associated variant rs5219 and rs7903146 along with MT-ND3 rs2853826 in T2D progression in population of Punjab but differential pattern of association is observed in different endogamous groups reflecting the role of ethnicity in progression of disease. Further studies with more combinations and larger sample size are required to understand the complex etiology of T2D in endogamous groups of Punjab.