INTRODUCTION

The completion of the first draft of human genome sequence is considered as the main hallmark in field of biological research and science. The advances in technology have led to the paradigm shift to an era of personalized medicine that uses the genetic profile of an individual to make decisions in context of diagnosis, treatment and prevention of disease (Wilson and Nicholls, 2015). Diabetes is a cluster of metabolic diseases characterized by hyperglycaemia that results from alterations in either insulin secretion, insulin action, or both. Diabetes can be mainly classified into Type 1 diabetes (T1D) which occurs due to the autoimmune destruction of pancreatic $\beta$-cells and Type 2 diabetes (T2D) which is related to insulin resistance coupled with inadequate insulin secretion by pancreas (American Diabetes Association, 2015). T2D is a global public health catastrophe threatening all economies especially those of the developing countries (Hu, 2011; Abdullah et al., 2014). T2D is associated with various symptoms like polyphagia (extreme hunger), polydipsia (increased thirst), polyuria (increased urination), glycosuria (glucose in urine); fatigue and weight reduction. The threat imposed by the alarming rise in T2D prevalence makes it a medical calamity of modern world. This emerging epidemic is also expected to prompt a sudden increase in the diabetes related complications, such as retinopathy, nephropathy, neuropathy, coronary artery disease and stroke (Mohan et al., 2013).

The global prevalence of T2D has escalated dramatically in the past several years. International Diabetes Federation (IDF), 2013 has predicted that by the year 2035, 592 million people will be suffering from T2D across the world. Western Pacific region has emerged as ‘diabetes epicentre’ of the world with 138 million diabetics followed by 72 million diabetics in South-East Asia. India ranks 2nd among top countries with 109 million people predicted to have diabetes mellitus by 2025 (International Diabetes Federation, 2013). India is presently experiencing an epidemic of T2D due to adaptation of western dietary habits, sedentary lifestyle and is often referred as the “diabetes capital of the world” (Mohan et al., 2007). A significant upsurge in the prevalence of T2D has been observed in both urban and rural areas of India (Anjana et al., 2011).
T2D is a multifactorial disease with a strong genetic component which is already evidenced by higher prevalence in particular ethnic groups and the difference in concordance rates between monozygotic and non-identical (dizygotic) twins (Weijnen et al., 2002; Blackett and Sanghera, 2013). It involves a complex interplay of environmental and genetic factors contributing to its pathology (Stumvoll et al., 2008; Tuomi et al., 2014). T2D is most often co-related with various risk factors like obesity, family history of diabetes, physical inactivity and ethnicity (Singh, 2011). India is a multi-ethnic country comprising of various caste, tribe and religious groups, which make them distinctive in comparison to rest of the world. The stringent endogamy customs and long term firm socio-religious boundaries along with the evolutionary forces have further enhanced the existing diversity levels in India (Reich et al., 2009). Some ethnic groups have a higher predisposition to develop T2D compared with others, even when exposed to similar environmental conditions (Abate and Chandalia, 2007; Sharma et al., 2013; Sun et al., 2014). To understand the role of genetic heterogeneity in T2D predisposition, mechanistic evaluation of pathogenesis of diabetes has to be studied in various ethnicities.

Identification of the genes that contribute to the T2D pathogenesis has been a challenging task due to the heterogeneous nature of the disease (Doria et al., 2008; Ahlqvist et al., 2011). The genetic architecture can also be influenced by gene-gene interactions (epistasis) that can act jointly to increase the risk of disease (Prasad and Groop, 2015). To unravel genetic risk variants for such common complex disorders, different strategies like linkage studies, candidate gene association studies and genome-wide association studies (GWAS) are involved, which have identified approximately 120 susceptibility loci conferring risk to T2D (Prasad and Groop, 2015).

1.1 **Calpain-10 (CAPN10)**

Calpains belong to the family of cytoplasmic cysteine proteases that are activated by Ca$^{2+}$ and stimulate proteolysis of enzymes involved in glucose metabolism and various other cell proteins. *CAPN10* is the first gene discovered by positional cloning with a putative role in T2D aetiology (Horikawa et al., 2000). It is highly expressed in tissues that regulate glucose homeostasis such as pancreatic $\beta$-islet cells, liver, skeletal muscle
and adipocytes (Baier et al., 2000; Carlsson et al., 2005, Pihlajamaki et al., 2006). CAPN10 is mapped on 2q37.3 chromosomal locus and spans 15 exons, with twelve single nucleotide polymorphism (SNPs) located in different intronic regions. CAPN10 was found in association with measure of insulin action in Pima Indians with normal glucose tolerance showing susceptibility through its effects on oxidation of glucose in skeletal muscles (Ling et al., 2009). It regulates the exocytosis of insulin in β-cells by acting as a fuel sensor in mitochondria and plasma membrane. Lower level of CAPN10 mRNA leads to insulin resistance (Ek et al., 2001). Saturated fatty acids may play a contributing role in triggering insulin resistance by interacting with genetic variants of CAPN10 (Ortho-Malender et al., 2002; Suzuki et al., 2004; Perez-Martine et al., 2011). These studies suggested that CAPN10 gene might affect insulin secretion, insulin action and hepatic glucose production. All the SNPs selected in the present study are located in the non-coding regions such as SNP 43; G>A in the third intron, SNP-19; Ins/del in the sixth intron and SNP-63; C>T in the thirteenth intron (Raj and Ramteke, 2012) and are believed to confer T2D risk by regulating the transcription of CAPN10 (Horikawa et al., 2000). Haplotype combination 112/121 and variants in CAPN10 (SNP-43, SNP-19 and SNP-63) are reported to be associated with a threefold increased risk of T2D in Mexican-Americans and an increased risk of diabetes in Northern European populations (Horikawa et al., 2000; Evans et al., 2001). Numerous studies performed on different populations have yielded inconsistent association of CAPN10 with risk of T2D (Horikawa et al., 2000; Tsuchiya et al., 2006; Adak et al., 2010; Ezzidi et al., 2010; Buraczynska et al., 2013; Sharma et al., 2013). Although great efforts have been made to understand the role of CAPN10 in T2D, its function is yet not clear.

1.2 Peroxisome Proliferator-Activated Receptor Gamma (PPARG)

PPARG is a nuclear receptor transcription factor involved in the metabolism of lipids, differentiation and proliferation of adipocytes and insulin sensitivity (Hegele et al., 2000). PPARG was the first gene reproducibly associated with T2D. It is located on chromosome 3p25 and contains 9 exons (Fajas et al., 1997). PPARG acts as a target of the insulin-sensitising thiazolidinediones, a class of drugs which are widely used for the treatment of T2D (Forman et al., 1995; Leonardini et al., 2009; Abbas et al., 2012). Evidences from the human and animal studies have confirmed the role of PPARG in
adipogenesis and T2D development (Barak et al., 1999; He et al., 2003; Gurnell et al., 2003; Imai et al., 2004; Zieleniak et al., 2008). The PPARG Pro12Ala polymorphism is the result of a CCA-to-GCA missense mutation at codon 12 (Stumvoll and Haring, 2002). PPARG Pro12Ala variant has been observed to improve insulin sensitivity. The association between the Pro12Ala variant and T2D or obesity have been enormously investigated in Caucasians and they have established the protective effect of Ala allele (Deeb et al., 1998; Clement et al., 2000; Altshuler et al., 2000; Ek et al., 2001; Wang et al., 2013). Although studies have described the association with T2D or obesity risk in the Indian population but they did not consider the impact of ethnicity on the associations (Radha et al., 2006; Sanghera et al., 2008; Sanghera et al., 2010). However, no study has documented its role in T2D among various endogamous groups of Punjab.

1.3 Peroxisome Proliferator-Activated Receptor Gamma, Co-activator 1 Alpha (PGC-1α)

PGC-1α (also known as PPARGC1A) was originally identified as a co-activator of PPARG (Puigserver, 2005). PGC-1α is a transcriptional co-activator that regulates genes involved in energy homeostasis, mitochondrial biosynthesis, lipid oxidation and T2D (Soyal et al., 2006). It is located at chromosome 4p15.1 and spans 67kb in length comprising of 13 exons and 12 introns (Pratley et al., 1998). Expression of PGC-1α has been demonstrated in tissues like heart, liver, brain, adipose tissue and skeletal muscles (Handschin and Spiegelman, 2006). Enhanced PGC-1α levels lead to enhanced glucose output by inducing the expression of gluconeogenic enzymes such as phosphoenolpyruvate carboxykinase and glucose-6-phosphatase (Soyal et al., 2006). Gene expression studies using various animal models demonstrated that PGC-1α stimulated 3-fold increase in secretion of glucose from liver when provided with gluconeogenic precursors (Puigserver and Spiegelman, 2003). PGC-1α also plays a key role in the regulation of hepatic glucose output through the control of gluconeogenesis (Yoon et al., 2001). The reduction in expression of PGC-1α is observed in T2D patients which makes it an important candidate gene in aetiology of T2D. In the present study, six polymorphisms of PGC-1α gene viz., Thr612Met (rs3736265), Thr528Thr (rs3755863), Gly482Ser (rs8192678), Asp475Asp (rs17574213), Thr394Thr (rs2970847) and IVS2+52C>A (rs2946385) as previously described were selected (Ek et al., 2001).
Chapter 1: Introduction

Among these Gly482Ser polymorphism is one of the most frequently studied polymorphism. Some studies have replicated its association with T2D in different populations (Ek et al., 2001; Barasso et al., 2006; Rai et al., 2007). Other PGC-1α genetic polymorphisms have also been reported to be associated with T2D or diabetes-related phenotypes like obesity and hypertension (Oberkofler et al., 2004; Vimalleswaran et al., 2005).

1.4 Rationale of Study

India is a highly diverse nation, where population is divided strictly on the basis of geography and language. Each geographic and linguistic group is further subdivided into various endogamous/religious groups and tribes. India exhibits greater genetic heterogeneity within regional units as compared to Europe, which is suggestive of evolutionary process (Reich et al., 2009). The splendid state of Punjab lies in Northern part of India (Singh, 1998). With the impact of globalization and industrialization, the tendency towards physical inactivity, hence overweight and obesity is mounting in the population of Punjab as in other societies of the world. The prevailing lifestyle conditions in the background of high disease susceptibility genetic factors are the leading cause of T2D. In North-West Indian population, the increase in diabetes and related complications has been observed due to transition from agriculture to a sedentary life style/physical inactivity associated with consumption of high calorie and fat rich diet (Bhatia, 2013; Anjana et al., 2014).

Rationale for Selecting Endogamous Groups

The main endogamous groups prevalent in Punjab are Banias (Merchant & moneylender, mainly trader community belonging to Vaishya group); Brahmins (traditionally priestly upper most class in Hindu classification but now involved in various occupations) and Jat Sikhs, which constitute the largest part of Sikh community (mainly agriculturist) (Sekhon, 2000). An endogamous group is selected on the basis of their high degree of homogeneity in terms of socio-cultural environment and its restricted mating pattern which provides conserved gene pool. This reduces the major limitations of association studies like population stratification, insufficient power and case-control matching. Lack of comparability between cases and controls can increase the risk of biases because there can be heterogeneity in exposure to environmental
challenges and population stratification (NCI-NHGRI et al., 2007). While selecting these endogamous groups (Bania, Brahmin and Jat Sikh), efforts were made to reduce population heterogeneity as to reduce the chances of false positive results to some extent.

**Rationale for Selecting Genes**

The three candidate genes (\textit{CAPN10}, \textit{PPARG} and \textit{PGC-1\alpha}) with known functional relation with T2D i.e., \textit{CAPN10} is involved in insulin secretion; \textit{PPARG} and \textit{PGC-1\alpha} are implicated in insulin resistance were selected and these are replicated in European and other Asian populations with inconsistency. List of SNPs selected for validation in target population is given in Table 1.1.

**Table 1.1: Target Single Nucleotide Polymorphisms**

<table>
<thead>
<tr>
<th>Gene &amp; Chromosome</th>
<th>SNPs</th>
<th>Location</th>
<th>Polymorphism</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calpain-10 \textit{(CAPN10)} Chromosome-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs3792267 Intron-3 SNP-43 (G&gt;A)</td>
<td></td>
<td></td>
<td></td>
<td>Horikawa et al. 2000</td>
</tr>
<tr>
<td>rs3842570 Intron-6 SNP-19 (Ins/del)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs5030952 Intron-13 SNP-63 (C&gt;T)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peroxisome Proliferator-Activated Receptor Gamma \textit{(PPARG)} Chromosome-3</td>
<td>rs1801282 Exon-4 Pro12Ala (C&gt;G)</td>
<td></td>
<td></td>
<td>Li et al. 2003</td>
</tr>
<tr>
<td>Peroxisome Proliferator-Activated Receptor Gamma-activator 1 alpha \textit{(PGC-1\alpha)} Chromosome-4</td>
<td>rs3736265 Exon-9 Thr612Met (G&gt;A)</td>
<td></td>
<td></td>
<td>Ek et al. 2001</td>
</tr>
<tr>
<td>rs3755863 Exon-8 Thr528Thr (C&gt;T)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs8192678 Exon-8 Gly482Ser (A&gt;G)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs17574213 Exon-8 Asp475Asp (G&gt;A)</td>
<td></td>
<td></td>
<td></td>
<td>Ek et al. 2001</td>
</tr>
<tr>
<td>rs2970847 Exon-8 Thr394Thr (C&gt;A)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs2946385 Intron 2 IVS2+52C&gt;A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Considering the high ethnic/genetic heterogeneity, the studies from Indian subcontinent are inadequate in terms of their presumption of Indians as one homogeneous population (Indian Genome Variation Consortium, 2008). Most of these population groups are highly endogamous and tends to maintain their social affinities. The unique genetic heterogeneity of the Indian population has frequently yielded contradictory results in various association studies. In the presence of highly sub-structured population in India, there is need to replicate studies in a more systematic manner (Sengupta et al., 2004;
Mastana, 2014). Very few reports are available on the genetic basis of T2D from North-West Indian population of Punjab. Therefore, to fill the existing lacunae, present case-control study was designed for selected (CAPN10, PPARG and PGC-1α) genetic polymorphisms for screening different endogamous groups (Bania, Brahmin and Jat Sikh) of Punjab for T2D predisposition. This study will help to delineate the epidemiological factors along with polymorphisms which are responsible for T2D predisposition in these population groups. This will further aid to construct statistical models which could be used for predicting the development of co-morbidities associated with T2D in future. Further, the role of T2D susceptibility in ethnic groups is important in directing public health interventions. Therefore, treatments may need to be tailored, by dose, type and intervention threshold, to a particular ethnic group in order to benefit from health care resources.

1.5 Objectives

- To conduct a candidate gene case-control association study in some endogamous population groups (Bania, Brahmin and Jat Sikh) of Punjab.
- To determine the association of various epidemiological and anthropometric (BMI, WC, WHR, WHtR, hypertension) and biochemical (Cholesterol, Triglycerides, HDL-C) variables with the development of T2D.
- To analyse CAPN10 (SNP-43, SNP-19 and SNP-63), PPARG (Pro12Ala) and PGC-1α (Thr612Met, Thr528Thr, Gly482Ser, Asp475Asp, Thr394Thr and IVS2+52C>A) genetic polymorphisms involved in T2D among Bania, Brahmin and Jat Sikh groups of Punjab.
- To examine the impact of the haplotype, linkage disequilibrium profile of the CAPN10 and PGC-1α genetic polymorphisms on the risk of T2D in population groups under study.
- To analyse the possible interaction between candidate genes for better understanding the role of these genes in the aetiology of T2D.