Genomics of Type 2 Diabetes – Review

Suparna Deepak, Rosalind Marita*

Department of Biochemistry & Biotechnology, Haffkine Institute for Training, Research & Testing, Parel, Mumbai, Maharashtra, India

ABSTRACT
Type 2 diabetes is a common life style disease affecting millions of people worldwide. The prevalence and the disabilities associated with diabetes in India are among the highest in the world. Further Indians are diagnosed with the disease, one or two decades earlier than Caucasian population. Urgent preventive measures are needed to reduce the social and economic burden of our society. Type-2 diabetes has a complex etiology. Both traditional risk factors such as obesity, stress, high calorie diets, lack of exercise as well as genetic factors are involved in the pathogenesis. The heritability of diabetes has propelled the search for disease susceptibility genes by linkage analysis and by the candidate gene approach. These classical approaches are now complemented by the genome wide association studies coupled with sophisticated analytical tools to evaluate the statistical significance. These advances have resulted in the identification of fifty six genetic susceptibility loci for Type 2 Diabetes. Among these, transcription factor 7-like 2, TCF7L2 gene has shown the most significant association with Type 2 diabetes. Further the association of this gene has been consistent in all populations including in Indian subjects. Genomic studies using large number of samples are extremely useful in early identification of subjects at risk besides offering valuable knowledge to design novel therapies for Diabetes.

Key words: Single nucleotide polymorphism, Asian Indians, Genomics

Abbreviations: T2D-type 2 diabetes; BMI-Body Mass Index; SNP-Single nucleotide polymorphism; GWAS-Genome wide association study.

Introduction:
Type-2 Diabetes has become a global health problem affecting over 300 million people worldwide. The prevalence in India is more than 9% with 63 million subjects currently suffering from Type-2 diabetes. It is predicted that the number would go up to 100 million by 2030 [1].

*Corresponding author:
Department of Biochemistry and Biotechnology, Haffkine Institute for Training, Research & Testing, Parel, Mumbai – 400 012. Tel: 22-24160961/62
email: r.marita@haffkineinstitute.org

Chronic uncontrolled diabetes leads to life threatening complications including heart disease, cataract, foot infections and kidney failure [2, 3]. Diabetes is a chronic metabolic disease that occurs either when the pancreas does not produce enough insulin or when the body does not effectively use the available insulin [2]. Diabetes Mellitus can be broadly classified into two types namely, Type-1 diabetes which is caused by an auto-immune destruction of the pancreatic β-cells leading to insulin deficiency and Type-2 diabetes characterized by insulin resistance and relative insulin deficiency [2, 4].
Type-2 diabetes is the most common form of the disease and comprises 90% of patients.

**Pathophysiology**

Type-2 diabetes results from an imbalance between insulin sensitivity and insulin secretion. Both longitudinal and cross-sectional studies have demonstrated that the earliest detectable abnormality in type-2 diabetes is the inability of the body tissues to respond to insulin. Impaired insulin action, defined as insulin resistance is observed in several tissues including skeletal muscle, adipose tissue and the liver [5]. Peripheral insulin resistance leads to increased insulin secretion from beta cells of the pancreas, as a compensatory mechanism. The resultant hyperinsulinemia maintains glucose level within the normal range.

In individuals at high risk of developing diabetes, beta cells are unable to increase insulin secretion and the beta cell function eventually declines precipitating high blood glucose levels or hyperglycemia [6-8]. Previous studies have identified several risk factors for type-2 diabetes. These include obesity, high cholesterol, hypertension, smoking, physical inactivity, dietar composition, family history and specific genes [9-11]

**Genetics of Type 2 Diabetes**

In complex disease like type-2 diabetes, both genetic and environmental factors determine the phenotype. The worldwide rise in prevalence of type-2 diabetes has led to an intense search for genetic factors influencing the susceptibility to the disease. Environmental influences, such as high-calorie diets and reduced physical activity are reported to accelerate disease development in those with genetic predisposition. Understanding genetic variants that increase the risk of type-2 diabetes could have significant clinical impact [12]. More detailed insight into the genetic risk factors could help us unravel the underlying molecular mechanisms involved in the pathogenesis besides opening up new pathways for therapeutic intervention.

Extensive studies on Caucasian population have uncovered numerous gene variants associated with type-2 diabetes. Asian Indian subjects are different from other populations and are reported to have high incidence of diabetes and insulin resistance [13]. Moreover, Asian Indians have a tendency for upper body obesity or abdominal obesity, referred to as “Asian Indian phenotype” [14-17]. Although Indian population in general cannot be considered obese based on body mass index, BMI our subjects are more insulin resistant compared to other ethnic groups. Moreover, many studies have shown that the incidence of type-2 diabetes in the offspring of diabetic parents of Indian origin is higher than the incidence among offspring of the parents of Caucasian origin. Previous research on Indian subjects [13-17] suggests insulin resistance as a causal link to increased susceptibility to diabetes [18].

**Heritability of Type-2 Diabetes**

The clinical assessment of type-2 diabetes has often incorporated genetic information in the form of family history. Family studies have estimated that the risk among offspring is 3.5-fold and 6-fold higher for those with a single or two diabetic parents, respectively [19]. Further, higher concordance rate of type-2 diabetes in monozygotic than dizygotic twins and the high prevalence in specific ethnic groups such as Pima Indians and Mexican Americans, all lend support to the greater contribution of genetic determinants [20]. It is also evident that type 2 diabetes-related intermediate and quantitative traits show substantial heritability [21]. The patterns of inheritance suggest that type-2 diabetes is polygenic and heterogeneous; hence multiple genes in different combinations play a significant role in different subsets of individuals suffering from diabetes.

**Linkage studies**

The traditional method of mapping disease genes is to identify long stretches of genetic markers which are inherited together in affected families by performing linkage analysis. Such markers are considered to be in linkage disequilibrium since they are inherited more
often than predicted by random association. By genotyping about 400 – 500 genetic markers, disease loci can be mapped on a genome-wide level. However, the classical approach of gene localisation by linkage analysis is not a suitable strategy for type-2 diabetes (T2D) for several reasons. First, there is a lack of a Mendelian inheritance pattern; second, the mean age of diagnosis is around 60 years for Caucasians and about 40 years in the Indian population. Further, one or both of the patient’s parents are often no longer available for study. Hence, it is difficult to obtain families with enough diabetic patients [22]. Although great efforts have been put into linkage studies of T2D, only two genes have been identified by this method. These are calpain 10 [CAPN10] gene and transcription factor 7-like 2 [TCF7L2] gene. The region on chromosome 2q37 containing CAPN10 was initially found in a linkage scan of T2D in Mexican Americans [23]. By large efforts in positional cloning, a haplotype of three single-nucleotide polymorphisms [SNPs] in an intron of CAPN10 was found to be associated with risk of T2D and this could explain most of the linkage to the disease [24]. These variants are also associated with an insulin resistance phenotype. From CAPN10, it became clear that intronic variation may contribute to risk of complex diseases. TCF7L2 is associated with an increase in risk of ~40% per minor T-allele. TCF7L2 was discovered by typing of microsatellite markers under a previously identified linkage peak [25, 26]. The association of gene variants in TCF7L2 with type-2 diabetes has since been widely replicated in populations of different ethnic origins [27, 28]. Gene, TCF7L2 encodes a transcription factor that is active in the Wnt-signaling pathway [29] and is implicated in pancreatic beta cell function and insulin secretion.

The candidate gene approach

Defects in genes encoding proteins that play a role in pathways involved in insulin and glucose homeostasis are excellent candidates for type-2 diabetes mellitus. A powerful approach to find such defects is the identification of a significant association between diabetes mellitus and a functional polymorphism in a candidate gene. Generally, this is achieved by comparing a random sample of unrelated diabetic patients with a matched control group. This approach may reveal a polymorphic allele that is increased in frequency in the patient group and such a significant association might point towards a disease susceptibility locus.

To date, over 250 candidate genes have been studied for their role in type-2 diabetes [30]. Majority of these studies have failed to uncover any association. In most instances, however, the initial association was not replicated in subsequent analyses. The candidate gene studies produced more unequivocal evidence for common variants involved in type-2 diabetes than did the linkage approach. The most robust candidate gene variants were the E23K variant in the KCNJ11 gene [31, 32], the P12A variant in the peroxisome proliferator-activated receptor-γ, PPARγ gene [33, 34], and common variants in the HNF1B and the Wolfram syndrome 1, WFS1 genes [35, 36]. KCNJ11 encodes a component of a potassium channel with a key role in β-cell physiology that is a target for the sulphonylurea class of oral antidiabetic drugs. PPARγ encodes a transcription factor involved in fat cell (adipocyte) differentiation and lipid metabolism. Peroxisome proliferator-activated receptor γ emerged as a type-2 diabetes candidate gene from the knowledge that thiazolidinedione (TZD) class of antidiabetic drugs, are high-affinity ligands for the PPARγ receptor. Subsequent association studies found a modest, protective effect of the P12A variant of the PPARγ gene variant on the risk of type-2 diabetes [37, 38] which has been confirmed in some GWA studies [39]. The P12A variant seems to influence susceptibility to type-2 diabetes by changing insulin sensitivity [40, 41]. Plasma cell glycoprotein-1/encoding ectonucleotide pyrophosphate phosphodiesterase-1 [PC-1/ENPP1] encodes a protein with inhibitory effect on insulin receptor function and subsequent insulin signaling [42,
Overexpression of ENPP1 in mice leads to insulin resistance and hyperglycemia. ENPP1 is widely expressed in liver, adipose tissue and skeletal muscle. A common K121Q polymorphism may result in gain-of-function leading to enhanced inhibition of the insulin receptor [44]. A recent meta-analysis demonstrated a marginal 8% increase in risk of type-2 diabetes [OR 1.08, 95% confidence interval [CI] 1.01-1.15; P=0.01] per copy of the rare Q-allele using data from more than 35,000 individuals of European ancestry [45].

Genome-wide scan
Genome-wide scans examine, in populations of related individuals like sib-pairs, extended pedigrees, or nuclear families, whether the disease or trait of interest co-segregate across generations with certain chromosomal region throughout the whole genome [46]. Genome-wide association studies [GWAS] and subsequent meta-analyses have identified 56 susceptibility loci for T2D that collectively explain ~10% of the disease risk [38]. These studies used large numbers of patients and cost several million dollars. Due to the vast amount of genetic variants analyzed in a GWA study, a high number of statistical tests are performed, thus leading to a substantial risk of false positives. The need for controlling this problem has resulted in the general use of a more stringent genome-wide significance level before an association is considered statistically significant. Current consensus has defined a genome-wide significance level of $P < 5 \times 10^{-8}$ to account for $10^6$ independent genome-wide hypotheses tested in a dense GWA [47], even though a significance level of $P < 10^{-7}$ has also been suggested [48, 49].

In 2007, results from some of the initial GWA studies investigating type-2 diabetes genes were published, namely the French DECODE, DECODA, DGI, WTCCC and FUSION studies [50-52]. These five independent GWA studies were all conducted using a two-stage strategy consisting of a GWA screen in an initial cohort of unrelated cases and controls followed by replication of the most significant findings in additional patient series. Each of these early GWA studies identified numerous potential susceptibility variants and only nine gene loci emerged as being consistently associated with risk of type-2 diabetes across multiple studies. These nine gene loci were TCF7L2, solute carrier family 30 member 8 [SLC30A8], hematopoietically expressed homeobox [HHOX], CDK5 regulatory subunit-associated protein 1-like 1 [CDKL1], cyclin-dependent kinase inhibitor 2A/2B [CDKN2A/B], insulin-like growth factor 2 mRNA-binding protein 2 [IGF2BP2], fat mass- and obesity-associated gene [FTO], KCNJ11 and PPARγ. Among these, three genes [TCF7L2, KCNJ11 and PPARγ] have been reported using other approaches.

A large amount of data available on the genetics of T2D from association studies of candidate gene variation include variants of ATP binding cassette subfamily C member 8 [ABCC8] [31, 53, 54], hepatocyte nuclear factor-1A [HNF1A] [55], hepatocyte nuclear factor-4A [HNF4A] [56, 57], glucokinase [GCK] [31, 58], insulin receptor substrate-1 [IRS-1] [59], protein tyrosine phosphatise-1B [PTP1B] [60, 61], the nuclear lamina gene, LMNA [62, 63], melatonin receptor-1B [MTNR1B] [64]. Mc Carty and co-workers have reported some of the susceptibility variants for T2D, namely PPARγ, KCNJ11, TCF7L2, CDKAL1, CDKN2A/2B, HNF1A/1B, IRS-1 and FTO. These genes have been known to be closely associated in most of the populations [34].

Asian Indian Population
Mohan and co-workers have carried out genetic studies on diabetic patients from South India. He has reported significant association of TCF7L2 gene with diabetes while no significant association was reported for the PPARγ gene [18, 66]. Radha and Mohan have reported the involvement of IAPP and TCFL2 genes in insulin secretion. They further suggested the involvement of PPARγ, PC-1, IRS-1, IRS-2, Calpain 10 and Apolipoprotein D genes to be associated with insulin resistance. Uncoupling protein-3/2 gene and PGC-1α genes have been shown to be associated with obesity [67]. Abate

J.PHARM.SCI.TECH.MGMT. Vol.1 Issue 1 2015 75
et al studied plasma cell membrane glycoprotein (PC-1), K121Q and G972A polymorphism in IRS-1 as a significant contributor to susceptibility of Asian Indians to insulin resistance [68]. Sangera et al have reported the association of four Type-2 Diabetes risk polymorphisms in north Indians which include PPARγ2, IGF2BP2, TCF7L2 and FTO [69]. Bhat et al have reported the association of PGC-1 alpha variants and mitochondrial gene G10398A/T16189C in the north Indian Population [70, 71]. Chauhan et al have replicated some of the well established genes in the western Indian population and have reported greater effect as compared to Caucasians [72]. Studies on PPARγ and TCF7L2 genes from our laboratory reveal increased incidence of type-2 diabetes in subjects with SNPs in TCF7L2 gene (rs7903146) in a worksite cohort from Mumbai, Western India [73].

The discovery of novel genes and pathways characterises the importance of conducting genome-wide scans in complex diseases like type-2 diabetes mellitus. New and improved technologies, such as microarrays that can type thousands of SNPs in a single assay, will also be of great importance in finding genetic variation in these new genes. Combining these genetic variations with new developments in the fields of bioinformatics, genomics and proteomics will lead to a greater understanding of the pathogenesis and susceptibility to type-2 diabetes mellitus. These studies may provide new information for diagnosis, prevention and early treatment of the disease. This genetic information may also form the basis for the development of new drug therapies enabling clinical translation and pharmacogenomics.

Acknowledgements: The authors are grateful to Dr. Abhay Chowdhary, Director for his encouragement. Authors thank Mr. Nikhil Kasarpalkar for help in the preparation of the manuscript.

Conflict of interest: Authors have no conflict of interest

References
and pathophysiology. JAMA, 301, (2009), pp. 2129-2140.


[31] A.L. Gloyn, M.N. Weedon, K.R. Owen et al. Large-scale association studies of variants in genes encoding the pancreatic beta-cell KATP channel subunits Kir6.2 [KCNJ11]and SUR1 [ABCC8] confirm that the KCNJ11 E23K variant is associated with


[57] Y.L. Muller, A.M. Infante, R.L. Hanson et al. Variants in hepatocyte nuclear factor 4-alpha are modestly associated with type 2 diabetes in Pima Indians. Diabetes. 54, (2005), pp.3035-3092


