REVIEW ARTICLE

Genetic Factors in the Etiology of Type 2 Diabetes: Linkage Analyses, Candidate Gene Association, and Genome-Wide Association – Still a Long Way to Go!

Deepak N Parchwani, SMS Murthy, Amit A Upadhyah, Digisha D Patel

Gujarat Adani Institute of Medical Sciences, Bhuj, India

Correspondence to:
Deepak N Parchwani
(drdeepakparchwani@yahoo.com)

Received: 31.08.2012
Accepted: 01.10.2012

DOI: 10.5455/njppp.2013.3.57-68

ABSTRACT

Type 2 diabetes is a complex and pleomorphic metabolic disorder arising from a complex interaction between genes and the environment. During the last decade there has been an outpouring of studies providing clues into the genetic architecture underlying type 2 diabetes mellitus. This review provides an overview of the genetics of type 2 diabetes in the context of recent progress in the understanding of the genetic susceptibility of the disease. Approximately 40 variants have been identified so far and the identification of these susceptibility loci for diabetes has introduced novel genes, pathways and mechanisms of diabetes pathogenesis. The genetic loci so far identified account for only a small fraction (approximately 10%) of the overall heritable risk for type 2 diabetes mellitus. Uncovering the missing heritability is essential to the progress of type 2 diabetes genetic studies and to the translation of genetic information into clinical practice. However, it may be a long time before all the susceptibility genes are found. It may take even more time before their roles in different pathways have been elucidated and the mechanisms involved in their interaction with other factors in the disease etiology clarified.

KEY WORDS: Type 2 Diabetes Mellitus; Genetic Factors; Genome Wide Association Studies

INTRODUCTION

Type 2 diabetes mellitus (T2DM) is a complex and pleomorphic metabolic disorder, characterized by defects in insulin secretion and insulin action which lead to hyperglycemia.[1] Nearly 285 million individuals worldwide are affected by diabetes mellitus. Global estimates for the year 2030 predict a further growth of almost 50%.[2] In 2000, it is estimated that 2.8% of world's population had diabetes mellitus and that by 2030 this number will be 4.4% of the world's population. According to Wild et al[2], the 'top' three countries in terms of the number of T2DM individuals with diabetes are India (31.7 million in 2000; 79.4 million in 2030), China (20.8 million in 2000; 42.3 million in 2030) and the US (17.7 million in 2000; 30.3 million in 2030). Clearly, T2DM has attained epidemic proportions in the 21st century, raising serious concern. Although the current rise in T2DM prevalence is driven mainly by changes in life-style, complex genetic determinants are widely considered to contribute to an inherent susceptibility to this disease. Among data supporting a strong heritable component are the marked differences

in T2DM prevalence across populations. This prevalence ranges from the high-risk Pima and South Sea Island populations, where it now exceeds 50%, to relatively low-risk European populations, where it had been closer to 5%, and intermediate-risk Asian Indians, United States minority populations of African and Hispanic ancestry, where the prevalence now approaches 20%.[3,4] Additional support derives from the strong familial aggregation, nearly four – fold increased risk for T2DM in siblings of a diabetic proband compared with the general population (λs of 3.5 to 4), the increased risk of T2DM in the offspring of one affected parent to an odds ratio of 3.4 – 3.5 and to 6.1 if both parents are affected.[5] The pathogenesis of T2DM is heterogeneous, suggesting that the contribution from individual genetic factors is modest. Identification of the genetic components of type 2 diabetes is the most important area of diabetes research because elucidation of the diabetes genes (alleles) will influence all efforts toward a better understanding of the disease, its complications, treatment, cure, and prevention. Linkage analysis and the candidate gene approach were the primary methods to link genotype and phenotype before the development of genome wide association studies (GWAS). This review will highlight the state of polygenic T2DM genetics i.e. we provide an overview of the genetics of type 2 diabetes in the context of recent progress made via linkage analyses, candidate gene association studies, and genome-wide association to understand the genetic architecture of the disease.

PROGRESS IN GENE DISCOVERY

In contrast to monogenic disorders, where results from single mutations lead to predictable phenotypes, the complex genetic architecture of susceptible and protective alleles in polygenic type 2 diabetes is more difficult to discern. Indeed, accumulating data suggest that type 2 diabetes is likely a collection of many closely related diseases with varying but often overlapping primary mechanisms that involve both impaired insulin secretion and insulin resistance. Adding to the challenge, type 2 diabetes is generally diagnosed later in life as a consequence of significant interactions of life-long environmental influences with multiple genetic factors. Because of the limited individual impact of single genetic loci, a full understanding of the complex gene-gene and gene-environment interactions in this disease has proven quite challenging. Prior to the GWAS era, the importance of genetic factors in the etiology of T2DM had been well established through family and twin studies.[6,7] The primary methods to identify susceptibility loci for diseases or phenotypic traits were linkage analysis and candidate gene association studies.

LINKAGE STUDIES

In the first phase of diabetes gene discovery, investigators used techniques based on linkage analysis to identify potential diabetes-associated genes. This approach, best suited for discovering genes with strong effects within relatively small family-based studies, involves genotyping affected family members for a set of markers to identify regions that are coinherited more commonly in affected family members and therefore potentially point to a genomic region containing a susceptibility locus.

Discovery of calpain 10 *(CAPN10)* was the first reported success from this line of investigation.[8] Although the haplotype described originally could not be strongly replicated in other ethnic groups, selected small nucleotide polymorphism (SNPs) in the promoter region do display evidence of association on meta-analysis, albeit with modest *P* values.[9-11] While the reason for this heterogeneity remains unexplained, functional evidence for a possible role of this protein in glycemic physiology is proved.[12] Well-replicated linkage on chromosome 20q resulted in the identification of noncoding variants in the *HNF4α* gene, which have been replicated in some studies.[13-16] Similarly, linkage with type 2 diabetes or related traits in chromosome 1q (q21 – q23) has been observed across ethnic groups and in multiple populations and encompasses over 400 expressed genes, including many strong candidates.[17-21] Investigators of the International Chromosome

1q Consortium carried a detailed fine-mapping and association analysis of this region, though no convincing locus predisposing to type 2 diabetes was reported. Finally, a thorough exploration of multiple (and modest) linkage signals by deCODE investigators resulted in the identification of the gene encoding the transcription factor 7-like 2 (*TCF7L2*) as the locus that confers the strongest effect on type 2 diabetes risk yet found.[22] This strong association has been replicated in many ethnic groups (with an OR of about 1.4 per risk allele).[23,24] Multiple risk allele appear to have an additive effect, one allele confers approximately 40% relative risk of diabetes, whereas two copies confer 80% relative risk.[25] Five SNPs and 1 tetranucleotide repeat polymorphism (DG10S478) within *TCF7L2* showed strong association with T2DM in 3 independent cohorts, and the SNPs (rs12255372 and rs7903146) showed strong linkage disequilibrium (LD) with composite at-risk alleles of the microsatellite marker (DG10S478). The association between the SNPs (rs12255372 and rs7903146) and decreased insulin secretion was also reported in American subjects with impaired glucose tolerance.[25] Subsequently, the association of *TCF7L2* with T2DM was replicated not only in populations of European origin but also in other ethnic groups.[26-29] The precise mechanisms by which *TCF7L2* variants increase risk are not well understood, although various lines of evidence suggest that they involve the enteroinsular axis, impaired insulin secretion and possibly reduced β cell proliferation.[30,31] Although this approach had proven extremely successful in identifying rare genetic variants of strong effects for single-gene disorders such as maturity onset diabetes of the young (MODY)[32], but proved limited in unveiling common genetic variants that underlie polygenic diseases. Once it became clear that association methods conducted in much larger samples was preferable to linkage analysis[33], investigators generally turned their attention to biological candidate genes.

CANDIDATE GENES

Candidate genes are previously discovered genes that, based on their inferred physiologic role, are hypothesized to contribute to the disease of interest if abnormal. In the case of type 2 diabetes, defects in genes encoding proteins that play a role in pathways involved in insulin control and glucose homeostasis would all be considered reasonable candidates for contributing to the genetic basis of disease. A powerful approach to finding 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 type 2 diabetes mellitus patients with a matched control group. This approach may show 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 mellitus. The majority of these studies have failed to uncover any association, possible explanation for this include small sample sizes, differences in T2DM susceptibility across ethnic groups, variation in environmental exposures, gene-environment interaction and in part due to adoption of low statistical thresholds before declaring association. Still, some findings stood the test of time such as role for some of the gene products involved in insulin secretion or insulin action, such as IRS-1, the glucagon receptor, the sulphonylurea receptor (SUR) and the peroxisome proliferator activated receptor-γ(PPARγ). Because of the current controversy regarding validity of association of many described gene(s), this review will focus only on a few of the most promising candidate genes that have been convincingly associated with T2DM:

PPARγ (peroxisome proliferator-activated receptor-γ)[34]: This gene has been widely studied because it is important in adipocyte and lipid metabolism. In addition, it is a target for the hypoglycemic drugs known as thiazolidinediones. A proline-to-alanine change in codon 12 (P12A) of the peroxisome proliferator–activated receptor γ (*PPAR*γ) gene was the first genetic variant to be definitively implicated in the common form of type 2 diabetes and is very common in most populations.

***ABCC8* (ATP binding cassette, subfamily C, member 8)[34]:** This gene encodes the high-affinity sulfonylurea receptor (SUR1) subunit that is coupled to the Kir6.2 subunit (encoded by U*KCNJ11*U, also known as the potassium channel, inwardly rectifying subfamily J, member 11). Both genes are part of the ATP-sensitive potassium channel, which plays a key role in regulating the release of insulin and glucagon in the beta cell. Mutations in either gene can affect the potassium channel's activity and insulin secretion, ultimately leading to the development of T2DM. Interestingly, *ABCC8* and *KCNJ11* are only 4.5 kb apart, and not far from the *INS* gene. Variant forms of *KCNJ11* (Lys) and *ABCC8* (Ala) genes have been associated with T2DM, as well as other diabetes-related traits.

***CAPN10* (calpain 10)[34]:** *CAPN10* encodes an ubiquitously expressed intracellular calcium-dependent cysteine protease. A haplotype that was initially linked to T2DM included an intronic A to G mutation at position 43, which appears to be involved in *CAPN10* transcription. Two amino acid polymorphisms (Thr504Ala and Phe200Thr) have also been associated with T2DM risk. It has been suggested that the coding and noncoding polymorphisms do not independently influence T2DM risk, but instead contribute to an earlier age at diagnosis. Physiological studies suggest that variations in calpain 10 activity effects insulin secretion, and therefore, susceptibility to T2DM.

***HNF1B*:** Research[35] on this MODY (maturity onset diabetes of the young) gene has produced a conclusive association of an intronic SNP (rs757210) in hepatocyte nuclear factor 1b (*HNF1B*) (previously known as *TCF2*) with type 2 diabetes.

The Glucagon Receptor (GCGR): Glucagon is a key hormone in the regulation of glucose levels. As such, the GCGR gene which encodes its receptor is a candidate diabetes susceptibility gene. A missense mutation in the glucagon receptor gene has been associated with decreased tissue sensitivity to glucagon and type 2 diabetes.

One of the major drawbacks of the candidate gene approach is that it will not lead to the identification of entirely new genes or pathways involved in T2DM. In order to identify new genes for T2DM, genome wide scans using polymorphic markers need to be performed.

GENOME WIDE ASSOCIATION STUDIES (GWAS)

A significant breakthrough in understanding the genetic basis of complex traits including T2DM, was facilitated by the arrival of GWAS. GWAS is a powerful biology-agnostic method to detect genetic variations that predispose to a disease. In GWAS, the entire genomes of individuals with and without the disorder of interest (i.e., cases and controls) are screened for a large number of common SNPs. If one type of the variant (one allele) is more frequent in people with the disease, the SNP is said to be "associated" with the disease. The associated SNPs are then considered to mark a region of the human genome which influences the risk of disease. The ability to interrogate the entire genome was made possible by several key advances: First, completion of the Human Genome Project and the International HapMap project identifying a large number of haplotype-tagged SNPs. Second, development of highly efficient, yet affordable genotyping technologies, which enabled rapid execution of a substantial number of GWAS. Third, development of analytical tools which made possible corroborative analysis and interpretation of huge amount of data from various sources. And fourth, cumulative assembly of well-phenotyped cohorts through different international collaborations.

Following the completion of the Human Genome Project, a search for genetic variation that might explain phenotypic diversity and an individual's risk of disease ensued. Assaying SNPs became a mainstream way to study the association of genetic variation and disease. A SNP is a single nucleotide DNA sequence variant in the genome that differs between members of the same species or a pair of chromosomes in an individual. SNPs occur on an average every 300

base pairs, have a low rate of recurrent mutation, and are most often binary in nature. Several million SNPs were discovered and deposited in public databases.[36] Initially, the HapMap genotyped 3.9 million SNPs in 270 DNA samples among four different ethnic groups and defined the underlying patterns of the inheritance of genetic variation. The inheritance pattern is quantified by linkage disequilibrium (LD), which represents the likelihood that alleles of nearby SNPs will stay together and preserve their linear arrangement on a haplotype during meiosis. This likelihood is dependent on recombination rates, with recombination events more likely to separate alleles that lie further apart. In this manner, two SNPs in strong LD will be inherited together more frequently than two SNPs in weak LD. By knowing this correlation structure, investigators only have to query a smaller subset of SNPs, or "tag" SNPs, to design genotyping arrays and conduct association analyses that essentially capture the majority of remaining common genomic variation. Genetic variants that are not directly genotyped can then be imputed from the genotyped "tag" SNPs subset. Imputation presumes the allele of a SNP at a different location inferred by its degree of LD with an allele at a directly genotyped variant.[37]

The first GWAS for T2DM was conducted in a French cohort composed of 661cases of T2DM (body mass index [BMI]<30 kg/m^2, first-degree family history of T2DM) and 614 non-diabetic controls. In total, 392,935 SNPs sourced from two different genotyping platforms were analyzed for association with T2DM. Although the two associations were not reproducible in follow-up studies (*LOC387761, EXT2*), this study identified novel and reproducible association signals at *SLC30A8* (OR 1.26, p<10^{-6}) and *HHEX* (OR 1.21, p<10^{-5}), and validated the well-known association at *TCF7L2*.[38] The realization that *SLC30A8* encodes a zinc transporter (ZnT-8; that transport zinc from the cytoplasm into insulin secretory vesicles) expressed in insulin containing granules in β cells[39] and the *HHEX* encodes a transcription factor involved in early pancreatic development[40] provided initial reassurance the GWA approach was useful for identifying functionally relevant loci. Shortly after the initial GWAS, the Icelandic company deCODE Genetics and their collaborators confirmed the association between T2DM and *SLC30A8*, *HHEX*, and the newly identified *CDKAL1*.[41] At the same time, 3 collaborating groups, the Wellcome Trust Case Control Consortium/ United Kingdom Type 2 Diabetes Genetics consortium (WTCCC/UKT2DM), the Finland-United States Investigation of NIDDM (FUSION), and the Diabetes Genetics Initiative (DGI), published their findings replicating the association of *SCL30A8* and *HHEX* with T2DM and independently discovering novel associations at *CDKAL1*, *IGF2BP2*, and *CDKN2A/B*.[42-44] With the exception of *LOC387761* and *EXT2*, these novel loci and 2 previously-known variants, *PPARG* P12A and *KCNJ11* E23K, were confirmed by multiple replication studies composed of European and non-European populations. Thus, the first round of European GWAS confirmed 8 T2DM susceptibility loci across multiple ethnic groups: *TCF7L2*, *SLC30A8*, *HHEX*, *CDKAL1*, *IGF2BP2*, *CDKN2A/B*, *PPARG*, and *KCNJ11*. In addition to these 8 loci, the WTCCC/UKT2D study identified a strong association between *FTO* variants and T2DM, although the effect of *FTO* variants on conferring susceptibility to T2DM was mostly mediated through increase in body weight.[45] In a fruitful collaboration, the various type 2 diabetes GWAS conducted by the FUSION group[44], the Wellcome Trust Case Control Consortium[43], and the Diabetes Genetics Initiative[42] combined their data to form the Diabetes Genetics Replication And Metaanalysis (DIAGRAM.)[46] This effort yielded six new loci (*JAZF1*, *CDC123-CAMK1D*, *TSPAN8-LGR5*, *THADA*, *ADAMTS9*, and *NOTCH2-ADAM30*) associated with type 2 diabetes at genome-wide statistical significance. The putative functional mechanisms of currently identified genes/genetic loci which affect type 2 diabetes risk are listed in Table 1. Other genetic loci associated with type 2 diabetes are *ZBED3* (Zinc finger, BED type containing), *TP53INP1* (Tumor protein p53 nuclear protein also known as stress inducible protein), *CHCHD9* (Coiled-coil-helix-coiled coil-helix domain containing 9), *CENTD2* (Arf-GAP with Rho-GAP domain, ANK repeat and PH domain containing protein 1), *KCNQ1* (Potassium voltage gated

Table-1: Genetic Variants Associated with Type 2 Diabetes Mellitus

Gene Region, Name (Chromosome Number)	Function
TCF7L2, Transcription factor 7-like 2 (10)	Encodes a high mobility group box- containing transcription factor that plays a key role in *Wnt* signaling pathway.
PPARG, Peroxisome proliferator-activated receptor γ (3)	Transcription factor involved in adipocyte development.
KCNJ11, Potassium inwardly-rectifying channel, subfamily J, member 11 (11)	Potassium channel that is part of the sulfonylurea receptor complex.
WFS1, Wolfram syndrome 1 (4)	Endoplasmic reticulum transmembrane protein expressed in the brain, heart and β-cells.
HNF1B, Hepatocyte nuclear factor-1β (17)	Transcription factor involved in pancreatic development.
SLC30A8, Solute carrier family 30 (zinc transporter), member 8 (8)	Expressed in β-cells – it is a Zinc transporter, this being necessary for insulin storage in secretory granules as well as being part of the secretory mechanism.
HHEX, Hematopoietically expressed homeobox (10)	Encodes a member of the homeobox family of transcription factor involved in pancreatic development.
CDKAL1, CDK5 regulatory subunit associated protein 1-like 1 (6)	The protein product shares homology with CDK5 regulatory subunit–associated-protein-1, a neuronal protein that inhibits activation of CDK-5.
IGF2BP2, IGF-2 mRNA binding protein 2 (3)	Regulates IGF-2 translation by binding to the 5′ UTR of IGF-2 mRNA; pancreatic development.
CDKN2A/B, Cyclin-dependent kinase inhibitor 2A and 2B (9)	Function as cell growth regulators that control cell cycle G1 progression by inhibiting CDK; islet development.
FTO, Fat mass and obesity associated (16)	Affects fat mass thereby indirectly predisposing to type 2DM
JAZF1, Juxtaposed with another zinc finger gene 1 (7)	Encodes a nuclear protein with three zinc fingers; functions as a transcriptional repressor.
CDC123-CAMK1D, Cell division cycle 123 homolog (*S. cerevisiae*) and Calcium /calmodulin-dependent protein kinase 1D (10)	CDC123 is a putative regulator of the cell-cycle while CAMK1D is a protein kinase that may be important in response to chemokines.
TSPAN8-LGR5, Tetraspanin 8 and Leucine-rich repeat- containing G protein- coupled receptor 5 (12)	Tetraspanin 8 is a cell surface glycoprotein that complexes with integrins regulating development and growth. Lgr5 is a potential marker of intestinal stem cells and hair follicles in humans. It is a target of *Wnt* signaling.
THADA, Thyroid adenoma associated (2)	Thyroid adenoma; associates with *PPARG*
ADAMTS9, ADAM metallopeptidase with thrombospondin type 1motif, 9 (3)	A member of the ADAMTS (a disintegrin and metalloproteinase with thrombospondin motifs) protein family, expressed in muscle and pancreas.
NOTCH2, Notch homolog 2 (1)	Transmembrane receptor implicated in pancreatic organogenesis
KCNQ1, Potassium voltage-gated channel, KQT-like subfamily, member 1 (11)	Pore-forming subunit of voltage-gated K-channel (*KvLQT1*); risk allele impairs insulin secretion.
IRS1, Insulin receptor substrate (2)	Plays a key role in transmitting signals from the insulin & IGF-1 to intracellular pathway.
MTNR1B, Melatonin receptor 1B (11)	Encodes one of two high affinity forms of a receptor for melatonin; risk allele associated with insulin secretion.
PROX1, Prospero protein homeobox 1 (1)	Corepressor of hepatocyte nuclear factor 4α which plays an important role in β-cell development
GCKR, Glucokinase regulator (2)	Regulatory protein that inhibits glucokinase.
ADCY5, Adenylate cyclase 5 (3)	Formation of adenylate cyclase.
UBE2E2, Ubiquitin conjugating enzyme E2F2 (3)	
BCL11A, B-cell lymphoma/ leukemia 11A (2)	Encodes C2H2 type zinc-finger protein.
GCK, Glucokinase (7)	Three tissue-specific forms phosphorylate glucose to produce glucose-6-phosphate in the liver and the β-cell.
DGKB/TMEM195, Diacylglycerol kinase beta and Transmembrane protein 195 (7)	DGKB encodes an isotype of DAG kinase which increases DAG and therefore increases insulin secretion. TMEM195 is a membrane phosphoprotein.
*C2CD4A/B*C2, calcium-dependent domain containing 4B (15)	Encodes nuclear localized factor 2 which is expressed in endothelial cells and the endocrine and exocrine pancreas.
KLF14, Kroppel like factor 14, also known as Basic transcription element binding protein (BTED5) (7)	Regulates the transcription of various genes, including TGF βR11.

channel subfamily KQT member 1), *HMGA2* (High mobility group AT hook 2), *HNF1A* (Hepatocyte nuclear factor 1 home box A), *PRC1* (Protein regulator of cytokinesis 1), *ZFAND6* (Zinc finger,

AN-1 type domain), *DUSP9* (Dual specificity phosphatase 9).[34,47,48,49]

Over the past 2 decades, many Asian countries have experienced a dramatic increase in the incidence of T2DM. Cumulative evidence suggests that Asians may be more susceptible than populations of European ancestry to insulin resistance and diabetes, which was thought to be due to interethnic genetic inheritance.[50] Several of the T2DM loci identified by European GWAS, have been confirmed in Asian populations. However, there are significant interethnic differences in the risk allele frequency at several loci. For example, risk allele frequencies of *TCF7L2* SNPs showing the strongest effect on T2DM in European populations are very few in the Japanese (~5%) compared to populations of European descent (~40%). As a result, *TCF7L2* variants have a little effect on susceptibility to T2DM in the Japanese.[47] In addition, the associations between T2DM and some loci are not consistent in other Asian populations. Therefore, to explain T2DM heritability in populations of Asian descent, it may be necessary to identify ethnic group-specific T2DM susceptibility loci, those were not captured in the European study.

Various epidemiological studies reported very high prevalence rates of diabetes among Indians,[51-53] probably due to genetic predisposition. Indians are diagnosed with diabetes a decade earlier and at a lower BMI than Europeans, which may be partly explained by their excess truncal obesity[54,55] and might be the reason for their tendency to insulin resistance, referred to as "Asian Indian phenotype".[56,57] A report from the Indian Genome Variation Consortium also suggested that most of the populations in the Indian subcontinent are distinct from HapMap populations.[58] Hence, genes known to be associated with T2DM in other populations need to be assessed for their role in the Indian population.

Among all the loci for polygenic T2DM, *TCF7L2* so far has shown the strongest association with the largest effect size in Indians.[59-61] This gene encodes a high mobility group box- containing transcription factor that plays a key role in *Wnt* signaling pathway and the gene product has been implicated in blood glucose homeostasis[62], and the variant rs7903146 is reported to be associated with measures of glucose metabolism. Studies evaluating the association of the Pro12Ala variant of the *PPARγ* gene with type 2 diabetes in Indians have reported contradictory results. This variant has been shown to confer protection against type 2 diabetes by many studies[60,61], similar to their western counterparts. Whereas, in some other studies it has been found to predict susceptibility for type 2 diabetes[63,64] despite the fact that frequency of the Ala allele at the PPAR*γ*-Pro12Ala locus is the same as in Caucasians.

Variant E23K (rs5219) in *KCNJ11* , variants of the *CDKAL1*, variant rs10811661, rs4402960 in *IGF2BP2* which lies near the *CDKN2A* gene, as well as polymorphisms in *SLC30A8* and *HHEX* (identified by GWA studies) have been significantly associated with type 2 diabetes in Indians with an increased effect size compared with Europeans.[61,65] In Indians, the D1057D genotype of insulin receptor substrate 2 (IRS-2) gene is susceptible to diabetes by interacting with obesity[66], so as the genetic polymorphisms (Thr394Thr, Gly482Ser and +A2962G) of peroxisome proliferator activated receptor-co-activator-1 alpha (*PGC-1*) gene.[67] A strong association of apolipoprotein E (APOE) (Hha1), angiotensin-1 converting enzyme (ACE) I/D, APOA1-CIII-AIV gene cluster with lipid levels in T2DM has been reported in North western Indian Punjabi populations.[68,69] A meta-analysis of one of the polymorphisms in *IRS-1* gene (Gly972 Arg) in four populations comprising a small south Indian population showed a significant association of this polymorphism with type 2 diabetes.[70] Studies in North Indians show genetic association of interleukin-1 beta (-511C/T) and interleukin-1 receptor antagonist (86bp repeat) polymorphism with T2DM.[71] A study on 10 candidate genes: the glucagon receptor, insulin receptor substrate 1, insulin receptor, human beta 3 adrenergic receptor, fatty

acid binding protein 2, mitochondrial tRNA [Leu (UUR)], sulphonylurea receptor, human uncoupling protein and the glycogen-associated regulatory subunit of protein phosphatase-1 genes suggested that none of them were associated with type 2 diabetes in south Indians.[72]

Among monogenic forms of T2DM, the most common is the maturity onset diabetes of the young (MODY), a genetically heterogeneous disease caused by mutations in the genes encoding hepatocyte nuclear factor-4a (MODY 1), glucokinase (MODY 2), hepatocyte nuclear factor-1a (MODY 3), insulin promoter factor-1 (MODY 4), hepatocyte nuclear factor-1b (MODY 5) and neuro D (MODY 6). A study showed high prevalence of MODY(4.8%) in south Indian population.[64] Another study reported that MODY 3 mutations in south Indians may be different from that observed in Western populations. [73]

Considering the fact that incidence and prevalence of T2DM continue to rise in Indian populations an extensive consortium based approach is required to identify the susceptibility locus and genes responsible for common form of familial diabetes in India. By defining the genetic susceptibility loci, such studies will eventually facilitate a direct, systematic exploration of the interactions of environmental factors, obesity, insulin resistance, and genetic predisposition in the pathogenesis of T2DM and prediabetic traits and also will open new pathways of exploration and therapy. Substantive facilitation of international collaboration in research and support of cross-disciplinary research will facilitate this process.

CONCLUSION

In a short span of five years, GWAS investigating the links between genetics and complex traits have transformed our knowledge. The identification of susceptibility loci for both monogenic and typical diabetes has introduced novel genes, pathways and mechanisms of diabetes pathogenesis and may open new vistas for elucidating the underlying pathophysiology of this complex disease. Understanding the complex interactions among genetic profiles, individual lifestyles, and environmental factors lies at the core of effective diabetes treatment. Although our knowledge of genetic basis of type 2 diabetes is increasing exponentially, the known variants represent approximately only 10 % of the heritability, suggesting the existence of a large portion of “missing heritability”. Major work is still required to identify the causal variants, test their role in disease prediction and ascertain their therapeutic implications. More importantly, considering the incidence and prevalence of T2DM in India, time has come for Indian research policy makers to switch gear and give new direction towards diabetic research.

REFERENCES

1. Ferrannini E. Insulin resistance versus insulin deficiency in non-insulin-dependent diabetes mellitus:problems and prospects. Endocr Rev 1998;19(4):477–490.
2. Wild G, Roglic A, Green R, Sicree H, King. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. Diabetes Care 2004;27(5);1047–1053.
3. Elbein SC. Evaluation of polymorphisms known to contribute to risk for diabetes in African and African-American populations. Curr Opin Clin Nutr Metab Care 2007;10(4):415-419.
4. Haffner SM. Epidemiology of type 2 diabetes: risk factors. Diabetes Care 1998;21(3):C3-C6.
5. Rich SS. Mapping genes in diabetes. Genetic epidemiological perspective. Diabetes 1990;39(1):1315-1319.
6. Poulsen P, Kyvik KO, Vaag A, Beck-Nielsen H. Heritability of type II (non-insulin-dependent) diabetes mellitus and abnormal glucose tolerance--a population based twin study. Diabetologia 1999; 42:139-145.
7. Groop L, Forsblom C, Lehtovirta M,Tuomi T, Karanko S, Nissén M et al. Metabolic consequences of a family history of NIDDM (the Botnia study): evidence for sex-specific parental effects. Diabetes 1996;45:1585-1593.
8. Horikawa Y, Oda N, Cox NJ, Li X, Orho-Melander M, Hara M et al. Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. Nat Genet 2000;26:163–175.

9. Weedon MN, Schwarz PEH, Horikawa Y, Iwasaki N, Illig T, Holle R et al. Meta-analysis and a large association study confirm a role for calpain-10 variation in type 2 diabetes susceptibility. Am J Hum Genet 2003;73:1208–1212.
10. Song Y, Niu T, Manson JE, Kwiatkowski DJ, Liu S. Are variants in the CAPN10 gene related to risk of type 2 diabetes? A quantitative assessment of population and family-based association studies. Am J Hum Genet 2004;74:208–222.
11. Tsuchiya T, Schwarz PEH, Bosque-Plata Ld, Geoffrey HM, Dina C, Froguel P et al. Association of the calpain-10 gene with type 2 diabetes in Europeans: results of pooled and meta-analyses. Mol Genet Metab 2006;89:174–184.
12. Turner MD, Cassell PG, Hitman GA. Calpain-10: from genome search to function. Diabetes Metab Res Rev 2005;21:505–514.
13. Ji L, Malecki M, Warram JH, Yang Y, Rich SS, Krolewski AS. New susceptibility locus for NIDDM is localized to human chromosome 20q.Diabetes 1997;46:876–881.
14. Zouali H, Hani EH, Philippi A, Vionnet N, Beckmann JS, Demenais Fet al. A susceptibility locus for early-onset non-insulin dependent (type 2) diabetes mellitus maps to chromosome 20q, proximal to the phosphoenolpyruvate carboxykinase gene. Hum Mol Genet 1997;6:1401–1408.
15. Ghosh S, Watanabe RM, Valle TT, Hauser ER, Magnuson VL, Langefeld CD et al. The Finland-United States investigation of non-insulindependent diabetes mellitus genetics (FUSION) study. I. An autosomal genome scan for genes that predispose to type 2 diabetes. Am J Hum Genet 2000;67:1174–1185.
16. Luo TH, Zhao Y, Li G, Yuan WT, Zhao JJ, Chen JL et al. A genome-wide search for type II diabetes susceptibility genes in ChineseHans. Diabetologia 2001;44:501–506.
17. Elbein SC, Hoffman MD, Teng K, Leppert MF, Hasstedt SJ. A genomewide search for type 2 diabetes susceptibility genes in Utah Caucasians. Diabetes 1999;48:1175–1182.
18. Vionnet N, Hani El H, Dupont S, Gallina S, Francke S, Dotte S et al. Genomewide search for type 2 diabetes-susceptibility genes in French whites: evidence for a novel susceptibility locus for early-onset diabetes on chromosome 3q27-qter and independent replication of a type 2-diabetes locus on chromosome 1q21–q24. Am J Hum Genet 2000;67:1470–1480.
19. Wiltshire S, Hattersley AT, Hitman GA, Walker M, Levy JC, Sampson M et al. A genome wide scan for loci predisposing to type 2 diabetes in a U.K. population (the Diabetes UK Warren 2 Repository): analysis of 573 pedigrees provides independent replication of a susceptibility locus on chromosome 1q. Am J Hum Genet 2001;69:553–569.
20. Meigs JB, Panhuysen CI, Myers RH, Wilson PW, Cupples LA. A genome-wide scan for loci linked to plasma levels of glucose and HbA1c in a community-based sample of Caucasian pedigrees: the Framingham Offspring Study. Diabetes 2002;51:833–840.
21. Hsueh WC, St. Jean PL, Mitchell BD, Pollin TI, Knowler WC, Ehm MG et al. Genome-wide and fine-mapping linkage studies of type 2 diabetes and glucose traits in the Old Order Amish: evidence for a new diabetes locus on chromosome 14q11 and confirmation of a locus on chromosome 1q21–q24. Diabetes 2003;52: 550–557.
22. Grant SFA, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, Sainz J et al. Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. Nat Genet 2006;38:320–323.
23. Florez JC. The new type 2 diabetes gene TCF7L2. Curr Opin Clin Nutr Metab Care 2007;10:391–396.
24. Cauchi S, El Achhab Y, Choquet H, Dina C, Krempler F, Weitgasser R et al. TCF7L2 is reproducibly associated with type 2 diabetes in various ethnic groups: a global meta-analysis. J Mol Med 2007;85:777–782.
25. Florez JC, Jablonski KA, Bayley N, Pollin TI, de Bakker PIW, Shuldiner AR et al, for the Diabetes Prevention Program Research G. TCF7L2 polymorphisms and progression to diabetes in the Diabetes Prevention Program. N Engl J Med 2006;355:241–250.
26. Groves CJ, Zeggini E, Minton J, Frayling TM, Weedon MN, Rayner NW et al. Association analysis of 6,736 U.K. subjects provides replication and confirms TCF7L2 as a type 2 diabetes susceptibility gene with a substantial effect on individual risk. Diabetes 2006;55:2640-2644.
27. Zhang C, Qi L, Hunter DJ, Meigs JB, Manson JE, van Dam RM et al. 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 2006;55: 2645-2648.
28. Scott LJ, Bonnycastle LL, Willer CJ, Sprau AG, Jackson AU, Narisu N et al. Association of transcription factor 7-like 2 (TCF7L2) variants

with type 2 diabetes in a Finnish sample. Diabetes 2006;55:2649-2653.

29. Damcott CM, Pollin TI, Reinhart LJ, Ott SH, Shen H, Silver KD et al. 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 200655:2654-2659.
30. Lyssenko V, Lupi R, Marchetti P, Del Guerra S, Orho-Melander M, Almgren P et al. Mechanisms by which common variants in the TCF7L2 gene increase risk of type 2 diabetes. J Clin Invest 2007;117:2155–2163.
31. Liu Z, Habener JF. Glucagon-like peptide- 1 activation of TCF7L2-dependent Wnt signaling enhances pancreatic beta cell proliferation. J Biol Chem 2008;283: 8723–8735.
32. Vaxillaire M, Froguel P. Genetic basis of maturity-onset diabetes of the young. Endocrinol Metab Clin North Am 2006;35:371–384.
33. Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science 1996;273:1516–1517.
34. Grant RW, Moore AF, Florez JC. Genetic Architecture of Type 2 Diabetes:Recent Progress and Clinical Implications. Diabetes care 2009;32(6):1107-1114.
35. Winckler W, Weedon MN, Graham RR, McCarroll SA, Purcell S, Almgren P, et al. Evaluation of common variants in the six known maturity-onset diabetes of the young (MODY) genes for association with type 2 diabetes. Diabetes 2007;56:685–693.
36. Sachidanandam R, et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 2001;409:928–933.
37. The International HapMap Consortium. A haplotype map of the human genome. Nature 2005;437:1299–1320.
38. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 2007;445: 828–830.
39. Chimienti F, Devergnas S, Favier A, SeveM. Identification and cloning of a β-cell-specific Zinc transporter, ZnT-8, localized into insulin secretory granules. Diabetes 2004;53: 2330–2337.
40. Bort R, Martinez-Barbera JP, Beddington RS, Zaret KS. Hex homeobox gene-dependent tissue positioning is required for organogenesis of the ventral pancreas. Development 2004;131:797–806.
41. Steinthorsdottir V, Thorleifsson G, Reynisdottir I, Benediktsson R, Jonsdottir T, Walters GB et al. A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. Nat Genet 2007;39:770-775.
42. Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H et al. Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University and Novartis Institutes for BioMedical Research. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 2007;316;1331-1336.
43. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H et al- Wellcome Trust Case Control Consortium (WTCCC). Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science 2007;316:1336-1341.
44. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 2007; 316:1341-1345.
45. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science 2007;316: 889-894.
46. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T et al Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat Genet 2008; 40:638–645.
47. Imamura M and Maeda S. Genetics of type 2 diabetes: the GWAS era and future perspectives. Endocrine Journal 2011; 58 (9):723-739.
48. Smushkin G, Vella A. Genetics of Type 2 Diabetes. Curr Opin Clin Nutr Metab Care 2010;13(4):471–477.
49. Billings LK and Florez. The genetics of type 2 diabetes: what we have learned from GWAS. Ann N Y Acad Sci 2010;1212:59–77.
50. Chan JC, Malik V, Jia W, Kadowaki T, Yajnik CS, Yoon KH et al. Diabetes in Asia: epidemiology, risk factors, and pathophysiology. JAMA 2009;301:2129-2140.
51. Ramachandran A. Epidemiology of Diabetes in India – Three Decades of Research. JAPI 2005;53:34-38.
52. Ramachandran A, Snehalata C, Kapur A, Vijay V, Mohan V, Das AK, Rao PV, Yajnik CS, Kuamar KMP, Nair JD for the Diabetes Epidemiology Study

Group in India DESI. High prevalence of diabetes and impaired glucose tolerance in India: National Urban Diabetes Survey. Diabetologia 2001;44:1094-1101.

53. Sadikot SM, Nigam A, Das S, Bajaj S, Zargar AH, Prasannakumar KM, Sosale A, Munichoodappa C, Seshiah V, Singh SK, Jamal A, Sai K, Sadasivrao Y, Murthy SS, Hazra DK, Jain S, Mukherjee S, Bandyopadhay S, Sinha NK, Mishra R, Dora M, Jena B, Patra P, Goenka K on behalf of Diabetes India. Comparing the ADA 1997 and the WHO 1999 criteria: Prevalence of Diabetes in India Study. Diabetes Research and Clinical Practice 2004;66:309-315.
54. Ramachandran A, Snehalatha C, Viswanathan V, Viswanathan M, Haffner SM. Risk of noninsulin dependent diabetes mellitus conferred by obesity and central adiposity in different ethnic groups: a comparative analysis between Asian Indians, Mexican Americans and Whites. Diabetes Res Clin Pract 1997;36:121–125
55. Shelgikar KM, Hockaday TD, Yajnik CS. Central rather than generalized obesity is related to hyperglycaemia in Asian Indian subjects. Diabet Med 1991;8:712–717.
56. Raji A, Seely EW, Arky RA, Simonson DC. Body fat distribution and insulin resistance in healthy Asian Indians and Caucasians. J Clin Endocrinol Metab 2001;86:5366-5371.
57. Mohan V, Deepa M, Deepa R, Shanthirani CS, Farooq S, Ganesan A. Secular trends in the prevalence of diabetes and glucose tolerance in urban South India-the Chennai Urban Rural Epidemiology Study (CURES-17). Diabetologia 2006;49:1175-1178.
58. Indian Genome Variation Consortium. Genetic landscape of the people of India: a canvas for disease gene exploration. J Genet 2008;87:3–20.
59. Bodhini D, Radha V, Dhar M, Narayani N, Mohan V. The rs12255372(G/T) and rs7903146(C/T) polymorphisms of the TCF7L2 gene are associated with type 2 diabetes mellitus in Asian Indians. Metabolism 2007;56:1174–1178.
60. Sanghera DK, Ortega L, Han S, Singh J, Ralhan SK, Wander GS, Mehra NK, Mulvihill JJ, Ferrell RE, Nath SK, Kamboh MI. 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 2008;9:59-64.
61. Chauhan G, Spurgeon CJ, Tabassum R, Bhaskar S, Kulkarni SR, Mahajan A . 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 2010;59:2068-2074.
62. Yi F, Brubaker PL, Jin T. TCF-4 mediates cell type-specific regulation of proglucagon gene expression by beta-catenin and glycogen synthase kinase-3 beta. J Bio Chem 2005;280:1457–1464.
63. Hegele RA, Cao H, Harris SB, Zinman B, Hanley AJ, Anderson CM. Peroxisome proliferator-activated receptor-gamma2 P12A and type 2 diabetes in Canadian Oji-Cree. J Clin Endocrinol Metab 2000;85:2014–2019
64. Radha V, Mohan V. Genetic predisposition to type 2 diabetes among Asian Indians. Indian J Med Res 2007;125:259-274.
65. Ng MC, Park KS, Oh B, Tam CH, Cho YM, Shin HD, Lam VK, Ma RC, So WY,Cho YS, Kim HL, Lee HK, Chan JC, Cho NH. Implication of genetic variants near TCF7L2, SLC30A8, HHEX, CDKAL1, CDKN2A/B, IGF2BP2, and FTO in type 2 diabetes and obesity in 6,719 Asians. Diabetes 2008;57:2226–2233
66. Bodhini D, Radha V, Deepa R, Ghosh S, Majumder PP, Rao MRS. The G1057D polymorphism of IRS-2 gene and its relationship with obesity in conferring susceptibility to type 2 diabetes in Asian Indians. Int J Obesity 2007;3:97-102.
67. Vimaleswaran KS, Radha V, Ghosh S, Majumder PP, Deepa R, Babu HNS. Peroxisome Proliferator-activated receptor-g coactivator- 1a (PGC-1a) gene polymorphisms and their relationship to type 2 diabetes in Asian Indians. Diab Med 2005;22:1516-21.
68. Singh P, Singh M, Gaur S, Kaur T. The ApoAI-CIII-AIV gene cluster and its relation to lipid levels in type 2 diabetes mellitus and coronary heart disease: determination of a novel susceptible haplotype. Diab Vasc Dis Res 2007;4:124-129.
69. Singh PP, Naz I, Gilmour A, Singh M, Mastana S. Association of APOE (Hha1) and ACE (I/D) gene polymorphisms with type 2 diabetes mellitus in North West India. Diabetes Res Clin Pract 2006;74:95-102.
70. Hitman GA, Hawrami K, McCarthy MI, Viswanathan M, Snehalatha C, Ramachandran A. Insulin receptor substrate-1 gene mutations in NIDDM; implications for the study of polygenic disease. Diabetologia 1995;38:481-486.
71. Achyut BR, Srivastava A, Bhattacharya S, Mittal B. Genetic association of interleukin-1beta (-511C/T) and interleukin-1 receptor antagonist (86 bp repeat) polymorphisms withType 2 diabetes mellitus in North Indians. Clin Chim Acta 2007;377:163-169.

72. Lepretre F, Vionnet N, Budhan S, Dina C, Powell KL, Genin E. Genetic studies of polymorphisms in ten non-insulin-dependent diabetes mellitus candidate genes in Tamil Indians from Pondicherry. Diabetes Metab 1998;24:244-50.
73. Radha V, Jakob EK, Pederson O, Mohan V, Hansen T. Identification of novel mutations in the HNF- 1a gene in south Indian patients with maturity onset diabetes of the young. Diabetes 2003;52(Suppl 1):A254.

Cite this article as: Parchwani DN, Murthy SMS, Upadhyah AA, Patel DD. Genetic factors in the etiology of type 2 diabetes: linkage analyses, candidate gene association, and genome-wide association – still a long way to go! Natl J Physiol Pharm Pharmacol 2013; 3:57-68.
Source of Support: Nil
Conflict of interest: None declared