CHAPTER 21

Genetics of Type 2 Diabetes

Venkatesan Radha, Viswanathan Mohan

Genes and Type 2 Diabetes

INTRODUCTION

Type 2 diabetes mellitus is a heterogeneous disease resulting from defects of both insulin secretion and insulin action. The etiopathogenesis of type 2 diabetes involves interplay of both genetic and environmental factors. Although the recent increase in diabetes prevalence reflects the effects of changing environmental factors, multiple lines of evidence support the view that genetic factors are equally important:

- The observation that the disease prevalence varies substantially among ethnic groups that share a similar environment, supports the idea that genetic factors contribute to predisposition to the disease.
- Familial aggregation suggests a genetic contribution to the disease, although shared common environmental traits play a role. The odds ratio for offspring of a single affected parent is 3.5 compared to those with no parental diabetes history, and this increases to 6.1, if both parents are affected.
- The high concordance in monozygotic twins (over 80%) compared to 50% concordance in dizygotic twins provides evidence for a genetic component in etiology of type 2 diabetes.
- Data from various studies are in support of a genetic basis for measures of both insulin sensitivity and insulin secretion.

The genetics of type 2 diabetes can be considered under two broad groups: genetics of monogenic forms of diabetes, where a single gene is causal in the development of the disease, and genetics of polygenic forms of diabetes, where a number of genes are responsible for the susceptibility of the disease.

Patterns of inheritance suggest that type 2 diabetes is both polygenic and heterogeneous, that is, multiple genes are involved and different combinations of genes play a role in different subsets of individuals. Exactly how many genes and what their relative contributions are still remains ambiguous. It is generally accepted that type 2 diabetes results from a complex interplay of genetic and environmental factors influencing a number of intermediate traits of relevance to diabetic phenotypes (β-cell mass, insulin secretion, insulin action, fat distribution and obesity) (Fig. 21.1).

To unravel the genetics of type 2 diabetes, four approaches have been used over the past two decades, each with some success. The first approach was to focus on forms of type 2 diabetes transmitted with a Mendelian dominant pattern of inheritance and/or other specific clinical features, which resulted in the discovery of genes involved in Maturity-Onset Diabetes of the Young (MODY), several syndromes of severe insulin resistance, neonatal diabetes, mitochondrial diabetes, and other rare
genetic syndromes. Together, these monogenic forms of type 2 diabetes account for less than 5% of all forms of type 2 diabetes. The second was to search for genetic variants in candidate genes that might be associated with the common type 2 diabetes. In general, these studies have focused on functional candidate genes, i.e. genes whose products are known to play a role in glucose homeostasis, or positional candidate genes, i.e. genes located in chromosomal regions that had been identified in linkage studies. The third approach was to perform microarray gene expression analysis in an attempt to define genetic alterations in type 2 diabetes. The fourth and more recent approach has been to perform high throughput Genome-Wide Association Studies (GWAS), which is expected to accelerate the speed of gene identification in type 2 diabetes.

GENETICS OF MONOGENIC DIABETES

Monogenic diabetes consists of different subtypes of single gene disorders comprising a large spectrum of phenotypes, namely neonatal diabetes mellitus (NDM) or monogenic diabetes of infancy, dominantly inherited familial forms of early-onset diabetes called MODY and rarer diabetes-associated syndrome diseases. All these forms are diagnosed at a very-young age and are unrelated to autoimmunity.

Neonatal Diabetes

Neonatal diabetes mellitus is defined by diabetes diagnosed within the first 6 months of life. It is rare (1:100,000–260,000 live births) and is clinically and genetically heterogenous. NDM can be either permanent [permanent neonatal diabetes mellitus (PNDM)] requiring lifelong treatment or transient [transient neonatal diabetes mellitus (TNDM)] with insulin dependence in the first months only, and a spontaneous remission usually by 18 months of age. The severe hyperglycemia and minimal ketosis appearing in the first days of life may have dramatic complications in the neonate, such as failure to thrive, acidosis, dehydration, and neurological alterations. Neonatal diabetes is a monogenic disorder, mostly unrelated to autoimmunity, and is conferred by mutations in genes that play a key role in β-cell function or development including glucokinase, the K_{ATP} channel and insulin gene.

Activating mutations in the gene encoding the Kir6.2 and SUR1 (ABCC8) subunits of the K_{ATP} channel are the most common cause of PNDM, accounting for about half of the cases. These mutations prevent the channel closure, and thereby of insulin secretion. The specific mutations determine the phenotype, and for KIR mutations, there is a correlation with functional severity of the mutation. TNDM can also result from mutations in KCNJ11 or ABCC8. Besides diabetes, around 20% of the patients with K_{ATP} channel mutations have neurological symptoms, which may constitute severe syndrome of developmental delay, epilepsy, and neonatal diabetes (DEND); sometimes an intermediate form may result, DEND, characterized by diabetes and less severe developmental delay without epilepsy. The identification of K_{ATP} channel mutations can profoundly impact the type of diabetic therapy; many of these neonates and infants can be managed by oral sulphonylurea drug.

Studies on Neonatal Diabetes in India

The molecular basis of neonatal diabetes has been reported in India in four isolated case reports. The genes implicated were KCNJ11, ABCC8 and INS. In a recent report, we identified mutations of KCNJ11, ABCC8, INS, AGPAT2, SLC2A2 and EIF2AK3 in 10 children (Table 21.1), when children with the first two mutations could be shifted to oral agents.

Maturity-Onset Diabetes of the Young

A less severe form of diabetes with a dominant mode of inheritance was first reported in three families by Tattersall in 1974; the term “MODY” for Maturity-Onset Diabetes of the Young was first used in 1975 following
Table 21.1: Mutations causing neonatal diabetes in south Indian pedigrees

<table>
<thead>
<tr>
<th>S. No</th>
<th>Diabetes subtypes</th>
<th>Mutation identified in gene</th>
<th>Nature of mutation</th>
<th>Age at onset</th>
<th>Developmental delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PNDM</td>
<td>KCNJ11</td>
<td>Known</td>
<td>82 days</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>PNDM</td>
<td>KCNJ11</td>
<td>Known</td>
<td>5 months</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>PNDM</td>
<td>KCNJ11</td>
<td>Known</td>
<td>48 days</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>PNDM</td>
<td>ABCC8</td>
<td>Known</td>
<td>6 months</td>
<td>No</td>
</tr>
<tr>
<td>5</td>
<td>PNDM</td>
<td>ABCC8</td>
<td>Known</td>
<td>2 months</td>
<td>No</td>
</tr>
<tr>
<td>6</td>
<td>PNDM</td>
<td>INS</td>
<td>Known</td>
<td>4 months</td>
<td>No</td>
</tr>
<tr>
<td>7</td>
<td>Infantile onset</td>
<td>INS</td>
<td>Known</td>
<td>10 months</td>
<td>No</td>
</tr>
<tr>
<td>8</td>
<td>Syndromic form of diabetes</td>
<td>AGPAT2</td>
<td>Novel</td>
<td>1 month</td>
<td>Yes</td>
</tr>
<tr>
<td>9</td>
<td>Syndromic form of diabetes</td>
<td>SLCA2</td>
<td>Novel</td>
<td>1 month</td>
<td>Yes</td>
</tr>
<tr>
<td>10</td>
<td>Syndromic form of diabetes</td>
<td>EIF2AK3</td>
<td>Known</td>
<td>3 months</td>
<td>Yes</td>
</tr>
</tbody>
</table>

(PNDM: Permanent neonatal diabetes mellitus)

further clinical description.\textsuperscript{39,40} MODY is defined as a dominantly inherited young-onset non-autoimmune diabetes that occurs in adolescence or young adulthood (usually less than 25 years of age) due to a primary defect in pancreatic β-cell function.\textsuperscript{41-43} Type 1 diabetes requires insulin even from the time of diagnosis due to lack of insulin secretion. MODY, which is a subtype of type 2 diabetes, accounting for 1–3% of all type 2 diabetes cases\textsuperscript{44} and does not generally require insulin initially because residual insulin secretion is still maintained for some years.

Heterozygous mutations or partial/whole gene deletions in eleven susceptibility genes, explain the clinical heterogeneity of the MODY subtypes (Table 21.2). The MODY genes encode the enzyme glucokinase (GCK),\textsuperscript{45} transcription factors HNF1α, HNF4α, HNF1β, PDX1 and NEUROD1,\textsuperscript{46,47} preproinsulin INS,\textsuperscript{48} KLF11,\textsuperscript{49} CEL,\textsuperscript{50} PAX4\textsuperscript{51} and BLK.\textsuperscript{52} Each has a crucial role in the development and/or function of the pancreatic β-cells (Table 21.2). Mutations in GCK, HNF1α and HNF4α are the most common causes of MODY.

Maturity-onset diabetes of the young 1 subtype is caused by mutations in HNF4α gene with 31 mutations reported in 40 MODY families.\textsuperscript{49} This gene plays a role in development of the liver, kidney, and intestines. Heterozygous mutation leads to progressive decrease in insulin production.\textsuperscript{50} HNF4α also determines fetal birth weight and hyperinsulinemia in utero and at birth.\textsuperscript{51}

Currently more than 600 GCK mutations were reported in more than 1,400 families worldwide.\textsuperscript{45} Heterozygous gain-of-function GCK mutations cause hypoglycemia whereas heterozygous loss-of-function GCK mutations result in alterations of both glucose-stimulated insulin secretion and hepatic glycogen synthesis, leading to mild fasting hyperglycemia (5.5–8.0 mmol/L).\textsuperscript{44,45} Homozygous GCK mutations causing PNDM have also been reported.\textsuperscript{45}

Mutations in HNF1α gene results in MODY 3 subtype. More than 200 HNF1α gene mutations were reported in ~370 families of various ethnic origins.\textsuperscript{46} MODY 3 is the commonest form of MODY worldwide.\textsuperscript{52} Clinically, the patients are non-obese; they often present with severe hyperglycemia and may develop microvascular complications. The prevalence of MODY 3 in early-onset type 2 diabetes mellitus varies from 2.5% to 36%.\textsuperscript{47,53} The penetrance of MODY 3 is high, and it often evolves insulinopenia and microvascular complications.\textsuperscript{51}

HNF1α and HNF4α mutations are associated with early-onset, and progressive diabetes often requiring oral hypoglycemic agents or even insulin. This shared phenotype is consistent with the interdependence between HNF1α and HNF4α forming part of a regulatory network in the pancreatic β-cell.\textsuperscript{54}

Two other transcription factor genes, PDX1 and NEUROD1, have an important role in the development of the endocrine pancreas, although representing a rare cause of MODY.

PDX1 or IFP (MODY 4) is co-expressed with insulin in the developing β-cell and is required for maintaining the β-cell phenotype. A frame-shift mutation in the coding sequence of the PDX1 gene was found to co-segregate with MODY in a five-generation family presenting a consanguineous link.\textsuperscript{53} In heterozygous mutation carriers,
Table 21.2: Novel Mutations/Polymorphisms identified in HNF1A, HNF4A genes in Indian MODY pedigrees

<table>
<thead>
<tr>
<th>S. no.</th>
<th>Diabetes subtypes</th>
<th>Mutation identified in gene</th>
<th>Region</th>
<th>Variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MODY 1</td>
<td>HNF-4A</td>
<td>Promoter 2</td>
<td>-1909 G/C</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-129 T/C</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-79 C/T</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-538G&gt;C</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Arg114Cys</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Val134Val</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Exon 2</td>
<td>Arg171Gly</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Glu235Gln</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Gly245Arg</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Exon 4</td>
<td>Arg283His</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Exon 7</td>
<td>Pro486Pro</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Intron 6</td>
<td>IVS6+12C&gt;T</td>
</tr>
</tbody>
</table>

(MODY: Maturity-Onset Diabetes of the Young)

the phenotypes range from impaired glucose tolerance to overt non-insulin-dependent diabetes.

Mutations in HNF1β/TCF2 were at first associated with MODY 5 in a few families. In addition to having an important role in early pancreas development, HNF1β function is also crucial in kidney development and nephron differentiation, and HNF1β mutations were shown to be a more common cause of renal cystic diseases and multiple renal malformations. This has been called renal cysts and diabetes syndrome (RCAD). More than 65 mutations were reported in 143 families. Although HNF1β was initially described as a MODY gene, patients usually present with renal disease or RCAD rather than with MODY.

The basic helix-loop-helix (bHLH) transcription factor NEUROD1 (or BETA2) with other factors (like NEUROG3) specifies the pancreatic endocrine lineage. This is also known as MODY 6. Heterozygous loss-of-function mutations in NEUROD1 were reported in a few MODY families, and very rare homozygous mutations associate with PNDM, cerebellar hypoplasia, learning difficulties, visual and hearing impairment. This syndrome highlights the critical role of NEUROD1 in the development of both endocrine pancreas and central nervous system.

Kruppel-like factor 11 is a protein that in humans is encoded by the KLF11 gene. MODY 7 is caused by mutations in the KLF11 gene. KLF11 regulates exocrine cells growth and behaves like a tumor suppressor for pancreatic malignancy.

Carboxyl-ester lipase gene (CEL) controls both exocrine and endocrine function of the pancreas. MODY 8 is caused by mutation in CEL gene on chromosome 9. It is caused by frame shift deletions in the variable number of tandem repeats (VNTR) of the carboxyl-ester lipase gene.

Paired box gene 4 encodes for PAX4 protein, and mutations in this gene results in MODY 9. The paired box gene 4 is involved in pancreatic islet development.

The proinsulin, precursor of insulin, is encoded by the INS gene. MODY type 10 is caused by mutation in INS (PROINSULIN) gene. MODY type 11 is caused by mutation in B lymphocyte kinase (BLK) gene. This gene preferentially expresses in B-lymphoid cells, and probably, it functions in a signal transduction pathway specific to this lineage.

Studies on MODY in India

Earlier studies using the clinical criteria used at that time reported on the high prevalence of MODY in South Indians (4.8%). Recently, we identified nine novel variants comprising seven mutations and two polymorphisms in the HNF1A gene. Functional studies revealed reduced transcriptional activity of the HNF1A promoter for two of the promoter variants. One of the novel mutations, Arg263His, was identified in a family of 30 individuals. The mutation co-segregated with diabetes in this family and was not seen in non-diabetic members in the family, showing that it was involved in causing MODY. There further three novel MODY 1 mutations have also been identified in HNF4A gene. Thus, about 9% of the clinically diagnosed MODY subjects are MODY 3, and 3% are MODY 1 in south India (Table 21.2).
A molecular genetic diagnosis in patients suspected of MODY aids in confirming the diagnosis, classifies the subtype, predicts the likely clinical course, defines risk for relatives and may change the patient's treatment.

**IMPLICATIONS OF MONOGENIC DIABETES GENES IN ADULT**

### Polygenic Type 2 Diabetes

Common polymorphisms in genes implicated in monogenic diabetes can increase the risk of common adult type 2 diabetes; common variation in HNF1A gene are associated with impaired insulin secretion and I27L and A98V polymorphisms in the MODY 3 gene (TCF1) with increased risk of type 2 diabetes in overweight individuals. An association of Val 88 allele with MODY phenotype was reported in South Indian patients. This allele was also associated with an earlier age at onset of type 2 diabetes mellitus. Single-nucleotide polymorphisms (SNPs) rs4810424 and rs3212198 in the HNF1A gene showed a modest association with type 2 diabetes, in line with previous studies. The -30G/A polymorphism in the GCK gene β-cell promoter has been associated with increased fasting plasma glucose and reduced β-cell function and to affect birth weight.

In a recent meta-analysis, this SNP was shown to have a modest, but significant effect on type 2 diabetes risk.

A common polymorphism in KCNJ11 (encoding Kir6.2), E23K, was shown to be associated with an increased risk of developing type 2 diabetes in European populations. This finding has been replicated in large-scale association studies and meta-analyses. The long-standing exploration of various subtypes of monogenic diabetes and their genetic dissection have improved our understanding of the β-cell physiology and regulation of insulin secretion in humans.

### Genetics of Polygenic Type 2 Diabetes

Unlike monogenic forms of diabetes as described above, the more common type 2 diabetes is a polygenic disorder. The individual susceptible and protective genes are more difficult to identify, and interact with environmental factors. The genetic susceptibility to type 2 diabetes is associated with polymorphisms that create amino acid variants in exons, influence the expression of genes in the regulatory pathways, or serve as sign posts in linkage disequilibrium. Alleles of these polymorphisms are present in both healthy individuals, and type 2 diabetes patients, although with different frequencies. The variants associated with an increase in the risk of disease are susceptibility variants but not unequivocal causal factors.

Type 2 diabetes has been at the forefront of human diseases and traits studied by new genetic analyses. The primary methods used to establish a link between genotype and phenotype were linkage analysis and candidate gene approaches.

### LINKAGE ANALYSIS

Linkage analysis is a method of mapping disease genes in affected families by genotyping about 400-500 genetic markers. Finding that affected family members share a certain marker that is identical by descent, i.e. identical because it was inherited from the same parent, is evidence that a disease causing variant is in linkage analysis with the genotyped marker. This method has been useful in identifying familial genetic variants with large effects such as those giving rise to MODY. However, it has been less successful in identifying genes that cause complex diseases such as type 2 diabetes. Despite efforts, only two genes have been identified by linkage: calpain 10 (CAPN10) and transcription factor 7-like 2 (TCF7L2). In the case of TCF7L2, a type 2 diabetes locus was mapped to chromosome 10q in both an Icelandic and a Mexican-American population. This region was later fine-mapped in the Icelandic population by use of 228 microsatellite markers covering a 10.5 Mbp region, pinpointing the locus to intron 3 of the TCF7L2 gene. The association between type 2 diabetes and a number of SNPs in the TCF7L2 gene have since been confirmed in different ethnic groups.

All of these demonstrate evidence of association with diabetic risk and consistent effect sizes. The risk allele confers a relative risk of approximately 1.4 compared to homozygous carriers of the non-risk allele, making this the strongest association with type 2 diabetes by far.

In a South Indian population, we genotyped two important SNPs in the TCF7L2 genes (rs12255372 and rs7903146). The T allele of both rs12255372 and rs7903146 polymorphisms was associated with type 2 diabetes and with non-obese type 2. This adds to the evidence that TCF7L2 is an important risk factor for type 2 diabetes in Asian Indians as well. Similar association have been demonstrated in Khatri Khals from North India. These were confirmed in meta-analyses and large GWAS.
The mechanisms by which TCF7L2 affects diabetes susceptibility are still not completely understood. T-allele of SNP rs7903146 was associated with risk of type 2 diabetes, impaired insulin secretion, incretin effects, and an enhanced hepatic glucose production.\textsuperscript{58} The other gene mapped by linkage analysis is a locus on chromosome 2.\textsuperscript{89} This locus was fine mapped and the causative gene shown to be CAPN10, the gene for calpain 10, a cysteine protease with largely unknown functions in glucose metabolism.\textsuperscript{59} Despite a number of negative replication studies, several meta-analyses have shown consistent association of CAPN10 with type 2 diabetes.\textsuperscript{51,52} Nevertheless, none of the large GWAS has identified CAPN10 as being associated with type 2 diabetes.\textsuperscript{65}

**CANDIDATE GENE APPROACH**

Identification of disease genes can also be made on the basis of association testing in populations rather than in families. Based on this hypothesis, the candidate gene approach focuses on the search for an association between type 2 diabetes and sequence variants in or near biologically defined candidate genes, chosen based on their known physiological function. The starting point for the candidate gene approach is that either altered expression and/or function of a particular gene product may affect a biological function or a disease. The importance of these or other nearby variants is tested by comparing the frequency in type 2 diabetes patients and normal glucose-tolerant subjects.

Extending the analysis of genes implicated in monogenic forms of diabetes has proved successful also for type 2 diabetes, as exemplified by HNF4A, HNF1A and KCNJ11 genes. Common variants of HNF4A (MODY1) have been associated with type 2 diabetes.\textsuperscript{55,70}

One of the main candidate genes implicated in adipogenesis, insulin resistance and type 2 diabetes is the peroxisome proliferator activated receptor-\(\gamma\) (PPAR-\(\gamma\)) gene. This is a transcription factor that is involved in adipogenesis via regulation of adipocyte gene expression and in glucose metabolism. Within a domain of PPAR-\(\gamma\)2 gene that enhances ligand independent activation, a common Pro12Ala polymorphism has been identified.\textsuperscript{93} Deeb et al. (1998) reported\textsuperscript{94} that the Ala allele of this polymorphism was associated with increased insulin sensitivity and decreased risk of type 2 diabetes. Since this initial work, the preponderance of evidence has supported PPAR's association with type 2 diabetes, with an odds ratio (OR) of -1.2.\textsuperscript{55} The risk of type 2 diabetes by this SNP has been studied prospectively in the Finnish Diabetes Prevention Study and the larger Botnia Prevention Study. In the Finnish study, 500 subjects with impaired glucose tolerance, the relative risk of developing diabetes was doubled in alanine carriers, contradicting the prior evidence that the alanine allele was protective. In the larger Botnia study, comprising more than 2,000 subjects, proline homozygotes were 1.7 times more likely to develop diabetes than alanine carriers.\textsuperscript{96} In contrast, we found that the Pro12Ala polymorphism of the PPAR-\(\gamma\) gene which is protective against diabetes of Caucasians do not offer protection in two cohorts of South Asians studied at Chennai, India and Dallas in US.\textsuperscript{97}

The SNP E23K of KCNJ11 has now been associated with type 2 diabetes. Although initial smaller studies failed to replicate the association of the E23K polymorphism with type 2 diabetes, large-scale studies and meta-analyses have consistently associated the lysine variant with type 2 diabetes, with an OR of 1.15.\textsuperscript{84}

The plasma cell glycoprotein-1 (PC-1) gene impairs insulin signaling at the insulin receptor level. The K121Q polymorphism of the ENPP1/PC-1 gene is associated with insulin resistance/atherogenic phenotypes, including earlier onset of type 2 diabetes, and myocardial infarction.\textsuperscript{84} The Q121 variant binds and inhibits insulin receptor more strongly than the K121 variant and is associated with insulin resistance and related metabolic abnormalities in the majority of populations. Prudente et al.\textsuperscript{100} suggested that the Q121 allele is a gene variant with pleiotropic deleterious effects on insulin resistance, obesity and type 2 diabetes. Our study supports that ENPP1 121Q predicts genetic susceptibility to type 2 diabetes in both South Asians and Caucasians.\textsuperscript{101}

Peroxisome proliferator-activated receptor-\(\gamma\) coactivator-1\(\alpha\) (PGC-1\(\alpha\)) is a cofactor involved in adaptive thermogenesis, fatty acid oxidation and gluconeogenesis. Dysfunction of this protein is likely to contribute to the development of obesity and the metabolic syndrome. Expression of PGC-1\(\alpha\) is downregulated in muscles of type 2 diabetic subjects. In addition, a common polymorphism of the PGC-1\(\alpha\) gene (Gly482Ser), expressing reduced PGC-1\(\alpha\) activity, has been linked to an increased risk of type 2 diabetes. These observations suggest that either reduced levels or compromised activity of PGC-1\(\alpha\) can be associated with the development of insulin resistance and type 2 diabetes.\textsuperscript{102} In a study on seven PGC1A variants only Gly482Ser polymorphism was associated with a 1.34 relative risk of type 2 diabetes.\textsuperscript{103} Studies on Thr394Thr,
Gly482Ser and +A2962G, of the peroxisome proliferator-activated receptor-co-activator-1 alpha (PGC-1A) gene with type 2 diabetes in Asian Indians showed that Thr394Thr (G-A) polymorphism is associated with type 2 diabetes\textsuperscript{106} and with body fat.\textsuperscript{108} Another study showed that the Thr394Thr and Gly482Ser variant genotypes provide protection against type 2 diabetes mellitus in North Indian populations.\textsuperscript{106}

Adiponectin, encoded by the ADIPOQ gene, is one of the adipocyte-expressed proteins that enhance insulin sensitivity. It regulates the homeostasis of glucose, lipid and energy metabolism.\textsuperscript{107,108} Genome-wide scans have mapped a susceptibility locus for type 2 diabetes and obesity/metabolic syndrome to chromosome 3q27, where the ADIPOQ gene is located.\textsuperscript{109-112} SNPs of ADIPOQ gene have been genotyped and several SNPs associated with hypoadiponectinemia, obesity and type 2 diabetes were identified.\textsuperscript{113-117} Two SNPs in the adiponectin gene, a silent T to G substitution in exon 2 (+45T/G) and a G to T substitution in intron 2 (+276G/T), were associated with type 2 diabetes and adiponectin level in Japanese population and with insulin resistance in some Caucasian populations.\textsuperscript{114,118,119} SNP 45 was associated with obesity in a German population.\textsuperscript{120} In the proximal promoter region of the APM1 gene: SNP-11426A/G and -11391A/-11377G haplotype predicted the associations with fasting plasma glucose, type 2 diabetes and adiponectin levels.\textsuperscript{111,121} Adiponectin has been associated with low diabetes risk. The metabolic effects of adiponectin are mediated by adiponectin receptors 1 (ADIPOR1) and 2 (ADIPOR2). Among six polymorphisms in ADIPOR1 and 16 polymorphisms in ADIPOR2 a significant association was seen between ADIPOR1 haplotypes and diabetes risk.\textsuperscript{122} Adiponectin is an adipose tissue specific protein that is decreased in subjects with obesity and type 2 diabetes. We showed for the first time that the +10211T>G polymorphism in the first intron of the adiponectin gene is associated with type 2 diabetes, obesity and hypo-adiponectinemia in Asian Indian population;\textsuperscript{123} thereby, suggesting adiponectin to be important for obesity and type 2 diabetes (Table 21.3).

Both the linkage analysis and candidate gene approaches failed limited success and were not able to satisfactorily explain the genetics of complex diseases. Therefore, efforts were made to study genome completely for multiple gene variants which led to GWAS.

**GENOME-WIDE ASSOCIATION STUDIES**

The ability to interrogate the entire genome was made possible by two the Human Genome Project and the International HapMap project. GWAS allowed the discovery of multiple gene variants with individually small effects. Once a specific polymorphism is associated with a disease, it is usually annotated by naming the gene in closest proximity to it. However, this does not necessarily mean that the variant in question is the molecular defect responsible for the phenotype, nor does it implicate the nearest gene; it simply flags a genomic region that harbors the causal variant, which may itself be acting at a certain distance, for instance, by modulating expression of a far-away gene. Therefore, while association signals are often identified by gene names, deep sequencing effort, fine-mapping and functional approaches are required to demonstrate a causal relationship between gene locus and the phenotype. Progress in high-throughput and affordable genotyping technology; analytical tools to assist in the data mining, cleaning and interpretation of large databases; and the assembly of international collaborations combining well-phenotyped cohorts made using these advances possible. Type 2 diabetes has been a beneficiary in this as substantial progress in our knowledge has been elucidated by GWA studies.

The first GWAS for type 2 diabetes was conducted in a French discovery cohort composed of 661 cases of type 2 diabetes and 614 non-diabetic controls: 392, 935 SNPs were analyzed for association with type 2 diabetes. This study identified novel and reproducible association signals at SLC30A8 and HHEX and validated the well-known association at TCF7L2.\textsuperscript{124} Investigators from the Icelandic company deCODE and their collaborators confirmed the association of loci SLC30A8 and HHEX with type 2 diabetes and identified an additional signal in CDKAL1.\textsuperscript{125-129} Three other collaborating groups, the Wellcome Trust Case Control Consortium (WTCCC), the Finland-United States Investigation of NIDDM Genetics (FUSION) group, and the Diabetes Genetics Initiative (DGI) published their findings replicating SLC30A8 and HHEX, and independently discovering novel associations at CDKAL1, IGF2BP2, and CDKN2A/B.\textsuperscript{130} These discoveries led to a plethora of studies which replicated the top signals in various ethnic populations.\textsuperscript{131-135} A number of loci were replicated and more number of them were unable to replicate.
A genome-wide association study for type 2 diabetes in a UK population revealed a novel locus associated with body mass index (BMI)—the Fat Mass and Obesity Associated (FTO) gene on chromosome 16.\textsuperscript{133} The representative SNP rs9903146 was associated with elevated BMI after replication in over 38,000 study participants of European ancestry. In addition, adiposity appeared to mediate the association between FTO variant and the risk of type 2 diabetes.\textsuperscript{134,135} Several other studies have also observed associations between FTO variants and obesity-related traits.\textsuperscript{136,137} Earlier studies on rs9939609 T>A and rs1783144 C>T variants of the intron 1 of the FTO gene showed an association with type 2 diabetes, which was independent of BMI in Asian Indians.\textsuperscript{138} We recently showed that rs8050136 C>A variant is associated with both generalized and central obesity among South Indians. The rs8050136 C/A polymorphism was associated with generalized obesity, and that there was no independent association of rs8050136 C>A with type 2 diabetes mellitus as its association with type 2 diabetes mellitus appears to be linked through obesity.\textsuperscript{139}

The identification of novel genes by GWAS is the discovery phase of work while the replication of these loci formed the validation phase of the work. Both these phases are essential for the discovery of novel genes/gene loci. Our own replication study in CURES resulted in the identification of genes/gene variants such as rs7756992, rs7754840, and rs6931514 of the CDKAL1, rs7020996 of the CDKNA2A/B gene, rs7923837 of the HHX gene, and rs12056034 of the BAZ1B genes were associated with type 2 diabetes in our population.\textsuperscript{139}

The coming together of large groups of investigators in collaboration resulted in meta-analysis of the GWAS identified genes,\textsuperscript{87,126} increasing the power of the study

<table>
<thead>
<tr>
<th>Table 21.3: Summary of genetic studies in type 2 diabetes in different populations using candidate gene approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>1  PPAR g gene (Pro12Ala)</td>
</tr>
<tr>
<td>2  PPAR g gene (Pro12Ala)</td>
</tr>
<tr>
<td>3  PGC-1a gene (Gly482Ser)</td>
</tr>
<tr>
<td>4  PGC-1a gene (Thr394Thr)</td>
</tr>
<tr>
<td>5  PGC-1a gene (Gly482Ser)</td>
</tr>
<tr>
<td>6  PGC-1a gene (Thr394Thr, Gly482Ser)</td>
</tr>
<tr>
<td>7  PC-1 gene (K121Q)</td>
</tr>
<tr>
<td>8  PC-1 gene (K121Q)</td>
</tr>
<tr>
<td>9  TCF7L2 gene (rs7903146)</td>
</tr>
<tr>
<td>10  TCF7L2 gene (rs7903146)</td>
</tr>
<tr>
<td>11  TCF7L2 gene (rs12255372, rs7903146)</td>
</tr>
<tr>
<td>12  TCF7L2 gene (rs7903146)</td>
</tr>
<tr>
<td>13  Adiponectin gene (+45T/G, +276G/T)</td>
</tr>
<tr>
<td>14  Adiponectin gene</td>
</tr>
<tr>
<td>15  Adiponectin gene</td>
</tr>
<tr>
<td>16  FTO gene</td>
</tr>
<tr>
<td>17  FTO gene</td>
</tr>
<tr>
<td>18  FTO gene</td>
</tr>
<tr>
<td>19  FTO gene</td>
</tr>
<tr>
<td>Table 21.4: Summary of genetic studies in type 2 diabetes in Asian Indian populations by replication and genome-wide association studies (GWAS) approach</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td><strong>1</strong> CDKAL1 gene</td>
</tr>
<tr>
<td>CDKN2A/B gene</td>
</tr>
<tr>
<td>HHEX gene</td>
</tr>
<tr>
<td>BAZ1B gene</td>
</tr>
<tr>
<td><strong>2</strong> GRB14 gene</td>
</tr>
<tr>
<td>ST6GAL1 gene</td>
</tr>
<tr>
<td>VPS26A gene</td>
</tr>
<tr>
<td>AP3S2 gene</td>
</tr>
<tr>
<td>HMG20A gene</td>
</tr>
<tr>
<td>HNF4A gene</td>
</tr>
<tr>
<td><strong>3</strong> TCF7L2 gene</td>
</tr>
<tr>
<td>TMEM106 gene</td>
</tr>
<tr>
<td>TMEM106 gene</td>
</tr>
<tr>
<td>MAP2K1 gene</td>
</tr>
<tr>
<td>TGFB3 gene</td>
</tr>
<tr>
<td>FLJ33579 gene</td>
</tr>
</tbody>
</table>

To discover common variants with frequency more than 5%. In a collaborative effort to study the population of South Asian ancestry, we performed a GWAS followed by replication of top SNPs and identified novel common genetic variants at six loci (GRB14, ST6GAL1, VPS26A, HMG20A, AP3S2 and HNF4A) newly associated with type 2 diabetes ($P = 4.1 \times 10^{-8}$ to $P = 1.9 \times 10^{-11}$). SNPs at GRB14 were also associated with insulin sensitivity ($P = 5.0 \times 10^{-4}$), and SNPs at ST6GAL1 and HNF4A were also associated with pancreatic β-cell function ($P = 0.02$ and $P = 0.001$, respectively). These provide insight into mechanisms underlying type 2 diabetes and show the potential for new discovery from genetic association studies in South Asians.141

The first Indian GWAS study identified a novel gene locus rs999451 (OR = 1.56, $P = 6.3 \times 10^{-13}$) within TMEM106 gene locus which encodes a probable vesicular transporter in nerve terminals.142 TMEM106 variants also showed association with decreased fasting plasma insulin and homeostatic model assessment of insulin resistance, indicating plausible effect through impaired insulin secretion. This study suggests that common susceptibility variants for type 2 diabetes are largely the same across populations, with population specific locus, and it also provides further insights into genetic architecture and etiology of type 2 diabetes in Indians (Table 21.4).

The GWAS have identified and replicated nearly 75 susceptibility loci associated with type 2 diabetes and related metabolic traits so far, mostly in Europeans, and some in African and South Asian populations. This is just a starting point for future genetic and functional studies. Firstly, deep sequencing and fine mapping needs to be done around these loci to pinpoint the gene region, and then functional studies need to be performed to understand the mechanisms of action by which these associated loci influence disease, and finally to predict potential implications of these findings in clinical settings. In the meantime new loci are being identified and validated, and these findings reveal new molecular pathways underlying diabetes etiology, gene-environment interactions, epigenetic modifications and gene functions.

**FURTHER READING**


**REFERENCES**


30. Fajans SS, Bell GI. MODY: history, genetics, pathophysiology, and clinical decision making. Diabetes Care. 2011;34:1876-84.


40. Ellard S, Colclough K. Mutations in the genes encoding the transcription factors hepatocyte nuclear factor 1 alpha (HNF1A) and 4 alpha (HNF4A) in maturity-onset diabetes of the young. Hum Mutat. 2006;27:954-69.


76. Gloyon AL, Weedon MN, Owen KR, et al. Large scale association studies of variants in genes encoding the pancreatic beta cell K-ATP channel subunits Kir6.2 (KCNJ11) and SUR1 (ABCC8) confirm that the KCNJ11 E23K variant is associated with increased risk of Type 2. Diabetes. 2003;52:569-72.


83. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007;447:611-78.


85. Bodhini D, Radha V, Dhar M, et al. 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;56:1174-8.


