Type 2 Diabetes: Genetic Advance and Aetiology

Mustafa Abdo Saif Dehwah, Zhang Shuang, Wang Zhen Hua, Wang Min and Qing-Yang Huang
College of Life Sciences, Hua Zhong Normal University,
Wuhan, Hubei 430079, Peoples Republic of China

Abstract: The aim of this review was to discuss the current understanding of Type 2 Diabetes Mellitus (T2DM) genetic advance and aetiology. The T2DM is a genetically heterogeneous disease, with several relatively rare monogenic forms and a number of more common forms resulting from a complex interaction of genetic and environmental factors. Earlier studies using a candidate gene approach, family linkage studies and gene expression profiling uncovered a number of T2DM genes, but the genetic basis of common T2DM remained unknown. Recent years have seen a tremendous surge in our understanding of the genetics of T2DM by Genome-Wide Association Study (GWAS). Approximately 20 genes consistently associated with T2DM mainly implicate pancreatic β-cell function in the pathogenesis of T2DM.

Key words: Type 2 Diabetes Mellitus (T2DM), susceptibility gene, aetiology

INTRODUCTION

Type 2 Diabetes Mellitus (T2DM) is a complex heterogeneous group of metabolic condition characterized by elevated levels of serum glucose, caused mainly by impairment in both insulin action and insulin secretion. T2DM is a complex trait where common genetic variants having modest individual effects act together and interact with environmental factors to modulate the risk of the disease. Traditional genetic research, including twin, adoption, and family studies, consistently supports that T2DM has a genetic component. Two broad approaches have been used to define the genetic predisposition of T2DM. First, molecular events in T2DM pathogenesis have been examined directly by testing the role of sequence variants of specific candidate genes. The candidate gene approach focuses on the search for an association between T2DM and sequence variants in or near biologically defined candidate genes which have been chosen based on their known physiological function. To overcome the shortcomings of the candidate gene studies, investigators have applied a genome-wide linkage scan strategy in which regularly spaced markers are traced in families and sibling pairs for segregation with T2DM. No prior knowledge of gene or gene effects is necessary, but the genetic locus must have sufficient impact on the disease susceptibility to be detectable. During the past decades, extensive efforts have been made to detect the underlying genetic structure for T2DM. However, until very recently, the genes involved have been poorly understood. Using a new and powerful technology in the form of a genome-wide chip that genotypes up to hundreds of thousands of SNPs, Genome-Wide Association Studies (GWAS) have recently led to the discovery of a group of novel genes that were reproducibly associated with T2DM risk. These studies have increased our understanding of the genetic aetiology of T2DM and provided invaluable insights into the way genetic studies should be conducted. The present review discusses the current understanding of T2DM genetic advance and aetiology.

EVIDENCE FOR THE GENETIC COMPONENT OF T2DM

The genetic component of T2DM comes from ethnic group differences in prevalence rates. These differences in prevalence range from 1% in Mapuche Indian tribes or Chinese population living in rural areas in mainland China, to extremely high levels found in Nauru and Pima Indians in Arizona (King and Rewers, 1993). Furthermore, the prevalence is higher in full-blooded Nauruan and Pima Indians than in those with admixture (Knowler et al., 1988; Serjeantson et al., 1983). Another source of evidence for genetic contribution in T2DM is familial aggregation. Lifetime risk of T2DM development is 40% in offspring of one T2DM parent and increases up to 70% if both parents have T2DM (Groop and Tuomi, 1997). In addition, evidence shows a greater likelihood of T2DM in offspring of affected mothers than affected fathers, indicating an excess maternal transmission of the disease (Thomas et al., 1994; De Silva et al., 2002; Arfa et al.,

Corresponding Author: Qing-Yang Huang, College of Life Sciences, Hua Zhong Normal University, Wuhan,
Hubei 430079, Peoples Republic of China

3407
Family, history of T2DM has a significant, independent and graded association with the prevalence of T2DM. In twins studies, very high concordance (35 to 100%) for T2DM has been reported among monozygotic (MZ) twins (Newman et al., 1987; Barnett et al., 1981).

Mutations in some genes cause rare forms of T2DM, giving additional support for the genetic roles in the aetiology of the disease. For example, genes KCNJ11 and ABCC8 carry rare mutations that cause a Mendelian form of neonatal diabetes (Babenko et al., 2006; Gloyd et al., 2004). Specific mutation in the mitochondrial genome was found to cause maternally inherited diabetes and deafness (Van Ouwereland et al., 1994). Maturity-onset Diabetes of the Young (MODY), which is characterized by high penetrance, early age at onset of hyperglycaemia and defective function of β-cells in the pancreas, accounts for 1-5% of all T2DM cases. Mutations in HNF4A, GCK, HNF1A, IPF1, HNF1B (TCF2) and NEUROD1 cause different subtypes of MODY1-6 (Fajans et al., 2001).

SEARCH FOR T2DM SUSCEPTIBILITY GENES

Susceptibility genes identified by linkage studies: The first T2DM gene identified by linkage is CAPN10 that belonged to the family of calcium activated non-lysosomal thiol proteinases. Horikawa et al. (2000) used relatively sparse Linkage Disequilibrium (LD) map of SNPs in the region of linkage on chromosome 2q to identify 3 common intronic variants of a previously unknown gene. Both a single intronic variant (UCSNP-43: G to A) and a specific haplotype combination defined by three polymorphisms (UCSNP-43, -19 and -63) were associated with T2DM in Mexican Americans, with lesser evidence of an association in a Northern European population from the Botnia region of Finland. Surprisingly, individuals with a combination of two different haplotypes were at the highest risk of T2DM (Horikawa et al., 2000; Wang et al., 2002). Baier et al. (2000) showed altered gene transcription and reduced muscle mRNA levels in muscle biopsies from Pima Indians with T2DM. CAPN10 SNP-44 (rs2975760 g4841T/C), located in intron 3 and 11 bp from SNP-43 (rs3792267 A/G), was independently associated with T2DM in several populations and was in LD with the missense mutation Thr504Ala and two 5'UTR variants (SNP-134 and SNP-135). An initial meta-analysis also supported an association with T2DM (Weedon et al., 2003). Furthermore, CAPN10 variants were valuable along with traditional risk factors in predicting the onset of T2DM in a prospective study of Botnia Finnish individuals (Lyssenko et al., 2005).

The TCF7L2 gene encodes for an enteronodocrine transcription factor that has a role in the Wnt signaling pathway, which is one of the key developmental and growth regulatory mechanisms of the cell. Reynisdottir et al. (2003) found suggestive linkage of T2DM to chromosome 10q in the Icelandic population. Fine-mapping with 228 microsatellite markers in Icelandic individuals with T2DM and controls throughout a 10.5 Mb interval on 10q identified one microsatellite marker, DG10S478 in intron 3 of the TCF7L2 gene, as being strongly associated with T2DM (p = 2.1×10^-4) (Grant et al., 2006). This association was replicated in both a Danish cohort (p = 4.8×10^-4) and a US cohort (p = 3.3×10^-4). Replications have appeared from analysis in subjects of Amish (Damcott et al., 2006), Finnish (Scott et al., 2006), UK samples (Grovès et al., 2006), French (Caučič et al., 2006), US (Zhang et al., 2006; Saxena et al., 2006), German KORA 500 K study population (Herder et al., 2008), population-based study of Caucasians (Van Hoek et al., 2008) and Chinese population (Ng et al., 2007; Chang et al., 2007). A meta-analysis of 28 original published association studies included 29,195 control subjects and 17,202 cases confirmed the association between TCF7L2 rs7903146 polymorphism and susceptibility to T2DM (OR = 1.46 95% CI 1.42-1.51; p = 5.4×10^-13). Compared with any other gene variants previously confirmed by meta-analysis, TCF7L2 can be distinguished by its tremendous reproducibility of association with T2DM and its OR twice as high (Caučič et al., 2007). Very recently, a large meta-analysis summarizes the strong evidence for an association between TCF7L2 gene and T2DM both overall and in Caucasians, North Europeans, East Asians, Indians, and Africans and suggested a potential multiplicative genetic model for all the four polymorphisms (rs7903146, rs7901695, rs12255372, rs11196205) of TCF7L2 gene among different ethnic populations except for Africans, where additive genetic mode is suggested for rs7903146 polymorphism, as well as suggests the TCF7L2 gene involved in near 1/5 of all T2DM (Tong et al., 2009). The increase in risk of T2DM associated with the TCF7L2 variant alleles so far identified is substantially greater than that associated with variants in the other confirmed T2DM-susceptibility genes (PPAR-, KCNJ11). The strongest signal for T2DM in GWAS was found for TCF7L2 rs7903146 with p-values down to <10^-8 and increase in the odds for the disease of 37% (Scott et al., 2007; Saxena et al., 2007). In the prediction study, TCF7L2 rs7903146 was significantly associated with the risk of future T2DM (OR = 1.30; p = 9.5×10^-13) and predicted progression from normal glucose tolerance to T2DM (OR = 1.27; p = 2.7×10^-7) in the MPP cohort (Lyssenko et al., 2008).
The locus at 18p was originally found to be linked to T2DM in families from Finland and Southern Sweden (Parker et al., 2001) and was confirmed in a second series derived from the Southwest of the Netherlands (Van Tilburg et al., 2003). The strongest evidence was obtained for marker D18S56 (LOD score 2.3, nominal p = 0.0006). In all three studies, the effect of the 18p region was strongest in the obese subpopulation. One of the human homologs of mouse Lpin1 gene, LPIN2, is located in the 18p11 region. LPIN2 is ubiquitously expressed in different tissues, including skeletal muscle (Zhou and Young, 2005). The LPIN2 gene is responsible for the lipodystrophy in the fdl line, which, among other traits, is characterized by severe insulin resistance. Fdl mice, in which Lpin1 is deleted, have diminished adipose tissue mass and multiple pathologies, including insulin resistance, fatty liver and progressive neuropathy of peripheral nerves (Shimomura et al., 1999; Péterfy et al., 2001). The protein Lipin is required upstream of PPARγ for normal adipocyte differentiation (Phan et al., 2004). The rs3745012 SNP of the LPIN2 gene is associated with T2DM and fat distribution (Aulchenko et al., 2007).

The human PTPN1 gene maps on chromosome 20q13.13 which showed evidence of linkage with early onset T2DM in a subset of 55 French families (Zouali et al., 1997). Multiple noncoding SNPs in the PTPN1 gene were implicated in T2DM in Caucasian and Mexican-American populations (Bento et al., 2004; Palmer et al., 2004), with the most consistent evidence for association occurring with SNPs spanning the 3' end of intron 1 of PTPN1 through intron 8 (p = 0.043-0.002). All of the associated SNPs were present in a single 100 kb haplotype block that encompassed the PTPN1 gene. Haplotype frequencies were significantly different between T2DM case and control subjects (p = 0.0035-0.0056), with a single common haplotype (36%) contributing strongly to the evidence for association, with OR = 1.3. Furthermore, the same haplotypes were associated with glucose homeostasis measures in Hispanic subjects (Bento et al., 2004). However, Flores et al. (2005) failed to replicate the findings in a large Caucasian study.

Adiponectin (APMI) is a strong candidate for T2DM given the clear role of plasma adiponectin levels in insulin sensitivity and the fall in adiponectin levels with obesity and T2DM. Adiponectin was mapped to the region on chromosome 3q27 which found to have positive linkage to metabolic traits and T2DM (Vionnet et al., 2000; Hegele et al., 1999; Kisselbach et al., 2000). Multiple studies found evidence that promoter variants in the APM1 gene were associated with T2DM in French and Swedish Caucasian and Japanese populations. However, APM1 variants did not appear to account for the 3q27 linkage in French families (Gibson and Freguel, 2004). APM1 variation at positions 45 (G allele) in exon 2 and 276 in intron 2 (T allele) were associated with a 4.5 fold increased risk of converting from impaired glucose tolerance to T2DM in the STOP-NIDDM trial (Zacharova et al., 2005). A meta-analysis of nine association studies included 2379 subjects confirmed significant association between the SNP45TG+GG and SNP276GG polymorphisms of APM1 with T2DM in Chinese populations (p = 0.05) (Li et al., 2008). The search for T2DM susceptibility genes on most other chromosomes (1q21.3-23, 2q37.3, 3p24.1, 3q28, 10q26.13, 12q24.31, and 18p11.2) is ongoing.

Susceptibility genes discovered by candidate gene association: Variants in genes encoding proteins that play a role in pathways involved in insulin control and glucose homeostasis are excellent candidates for T2DM. The candidate gene approach focuses on the search for an association between T2DM and sequence variants in or near biologically defined candidate genes which have been chosen based on their known physiological function. The importance of these variants is tested by comparing the frequency in T2DM subjects and control individuals. Numerous variants within several genes that confer an increased susceptibility to T2DM have been identified by candidate gene studies but only a small number have been identified as strong candidates for T2DM such as PPARγ and KCNJ11 genes, which they were confirmed by GWAS (Zeggini et al., 2007; Sladek et al., 2007).

The PPARγ is a member of nuclear hormone receptor superfamily of transcription factors and the target of the widely used class of insulin sensitizers, the thiazolidinediones. Multiple studies have examined a Pro12Ala polymorphism in the PPARγ isoform. The rare Ala allele is seen in about 15% of Europeans and was shown to be associated with increased transcriptional activity, increased insulin sensitivity and protection against T2DM in an initial study (Deeb et al., 1998). Subsequently, smaller studies provided inconsistent results. Larger studies have shown a consistently protective effect of the Ala12 allele when compared with the common Pro12 allele. The common Pro12 allele increases T2DM risk by an OR of 12.5 in a meta-analysis in cross sectional studies (Altshuler et al., 2000). Several, but not all studies suggest that the rare Ala12 allele protects against insulin resistance and obesity, but the association with obesity has been inconsistent. A meta-analysis of 41 published and 2 unpublished studies included 42,910 subjects (Asia, Europe and North America).
America) demonstrated that the reduced risk of T2DM in Ala12 carriers is not homogeneous (Ludovico et al., 2007). Recent replication studies reported significant associations between rs1706328, rs11709077, rs1801282 in PPARγ and T2DM in the German KORA S00 K study population (Herder et al., 2008). PPARγ Pro12 allele was associated with T2DM in HuBei Han Chinese population (Dehvy et al., 2008). In the prediction study, PPARγ SNP rs1801282 was significantly associated with the risk of future T2DM (OR = 1.20; \( p = 4.0 \times 10^{-5} \)) and predicted progression from normal glucose tolerance to T2DM (OR = 1.15; \( p = 0.03 \)) in the MPP cohort (Lyssenko et al., 2008).

KCNJ11 encodes a subunit of an inwardly rectifying ATP-sensitive potassium channel 1 (KATP). I (KATP) channels are crucial for the regulation of glucose-induced insulin secretion in pancreatic β-cell. The β-cell potassium channel comprises two subunits, the potassium channel encoded by the gene KCNJ11 and the regulatory subunit (SUR1 encoded by ABCC8) that binds sulfonylureas and ATP. Variants in both subunits have been associated with T2DM. Two variants in the ABCC8 gene were initially associated in smaller studies (Ioune et al., 1996; Hart et al., 1999; Meirhughe et al., 2001) and larger subsequent studies appeared to confirm the association (Barroso et al., 2003; Hani et al., 1997). However, other studies have not confirmed the association (Gloy et al., 2003; Florez et al., 2004). The association between these variants and altered insulin secretion was also observed (Elbein et al., 2001; Hansen et al., 1998). The adjacent KCNJ11 gene was initially noted to have multiple nonsynonymous coding variants (Ioune et al., 1997), but none were associated with T2DM (Altshuler et al., 2000; Sakura et al., 1996). However, a E23K polymorphism was shown to lower the sensitivity of the potassium channel to ATP and thus reduce insulin secretion in vitro (Schwanstecher et al., 2002) and to be associated with T2DM (Barroso et al., 2003; Love-Gregory et al., 2003; Yokoi et al., 2006). The result of Meta-analysis confirmed the association between the E23K variant and susceptibility to T2DM in UK cohort (K allele OR = 1.23; \( p = 1.5 \times 10^{-5} \); KK genotype OR = 1.65; \( p = 2 \times 10^{-6} \)) (Gloy et al., 2003) and white case-control subjects (\( n = 2.824 \), OR = 1.49, \( p = 2.2 \times 10^{-5} \)) (Nielsen et al., 2003). Subsequently, Florez et al. (2004) thoroughly examined KCNJ11 and ABCC8 and tested sufficient variants to completely tag all variation in this region. Again, the E23K variant was associated with T2DM in 3,413 subjects and the association was confirmed in a meta-analysis that included over 5000 T2DM subjects and 4747 controls (\( p < 10^{-7} \), OR = 1.15). Recent replication studies reported significant associations between common variants in the KCNJ11 genes and T2DM in a Japanese population (Omori et al., 2008).

Evidence for a role of the HNF4A gene in T2DM predisposition is also mounting (Barroso et al., 2003; Love-Gregory et al., 2003; Silander et al., 2004; Zhu et al., 2003). HNF4A regulates genes involved in glucose and fatty acid metabolism, as well as insulin secretion, and is therefore critical for maintaining lipid and glucose homeostasis. The K121Q variant of the ENPP1 gene has been best studied and shown to be associated with obesity but not T2DM in Caucasian and African American subjects ascertained from the New York Cancer Project (Matsukura et al., 2006). Other investigators found an association of K121Q with earlier onset T2DM and coronary disease among T2DM patients (Bacci et al., 2005). The association of K121Q with T2DM was replicated in three relatively small populations, two of South Asian ancestry and one Caucasian (Abate et al., 2001).

Minton et al. (2002) sequenced DNA from 29 patients and uncovered 12 SNPs of WFS1. The most abundant genetic variant alters the amino acid at position 611 from a histidine to an arginine (H611R). The arginine variant was present in 40% of diabetic cases and 45% of controls (\( p < 0.02 \), suggesting that the variant was protective from T2DM. Through large-scale candidate-pathway study, 1536 SNPs in 84 candidate genes were studied for association with T2DM, only rs10010131, rs6446482, rs752854, rs734312 SNPs in WFS1 were associated with T2DM in UK and an Ashkenazi Jewish population (OR = 0.92 95% CI 0.88-0.95, \( p = 10^{-7} \)) (Sandhu et al., 2007). The result was then replicated in northern Swedish populations, the minor allele (G) at SNP rs752854 was statistically associated with reduced risk of T2DM (OR = 0.85 95% CI 0.75-0.96, \( p = 0.01 \)), while borderline statistical significance was observed for the other three SNPs (rs10010131, rs6446482, rs734312) (Franks et al., 2008). Recent replication studies reported significant associations between WFS1 rs10012946 (proxy for rs10010131, \( r^2 = 1.0 \)) and T2DM in population-based study of Caucasian (Van Hock et al., 2008). A meta-analysis of 11 association studies included 12,979 cases subjects and 14,937 controls was robustly confirmed the association of WFS1 rs10010131 with risk of T2DM (OR = 0.89 95% CI 0.86-0.92, \( p = 4.9 \times 10^{-10} \)) (Franks et al., 2008). In the prediction study, WFS1 SNP rs10010131 was significantly associated with the risk of future T2DM (OR = 1.12; \( p = 0.001 \)) and predicted progression from normal glucose tolerance to T2DM (OR = 1.13; \( p = 0.004 \)) in the MPP cohort (Lyssenko et al., 2008).

The HNF1B (also known as TCF2) gene is located on chromosome 17cen-q21.3, a region linked to T2DM.
(Demenais et al., 2003). Although HNF1B is important for the development of the pancreas, mutations in HNF1B show other more characteristic phenotypes like cystic kidney disease, liver dysfunction and abnormal urogenital tract development (Nishigori et al., 1998; Lindner et al., 1999; Bingham et al., 2000). Polymorphisms in HNF1B were reported to be associated with T2DM in Caucasians (Bonnycastle et al., 2006; Winckler et al., 2007; Gudmundsson et al., 2007). However, the previously T2DM associated SNPs in HNF1B are not associated with T2DM in cohorts of 2,293 individuals from the Bonn study (Finland) and in 15,538 individuals from the Malmö Preventive Project (Sweden) (Holmquist et al., 2008).

Novel susceptibility genes from genome wide association studies: Using a candidate gene approach, family linkage studies and gene expression profiling uncovered a number of T2DM susceptibility genes (Huang et al., 2006). Since, 2007, a new window has opened on defining potential T2DM genes through genome-wide SNP association (GWA) studies of very large populations of individuals with T2DM. GWA revealed the first susceptibility loci for T2DM and validated some of the known candidates (Table 1, Fig. 1). To date, the total of the twelve update GWA scans for T2DM have been published (Table 2). Six of these represent high-density scans (i.e., at least 300,000 SNPs, offering genome-wide coverage >65%), in samples of Northern European descent (Sladek et al., 2007; Saxena et al., 2007; Scott et al., 2007; Steinhorsdottir et al., 2007; Zeggini et al., 2007; Salonen et al., 2007). Another four studies featured a wider array of ethnic groups including Native American (Hanson et al., 2007), Hispanic (Hayes et al., 2007) and populations of European descent (Rampersaud et al., 2007; Florez et al., 2007), but were less extensive with respect to both sample size and SNP density (all used the Affymetrix 100K array). Two recent studies in East Asian subjects were on a smaller scale (ranging between 82,000 and 207,000 typed SNPs in a few hundred cases only) (Yasuda et al., 2008; Unoki et al., 2008).

The first GWA covered 392,935 SNPs (passing quality control) and identified four novel loci including SLC30A8, LOC387761, IDE-KIF11-HHEX and EXT2-ALX4 (Sladek et al., 2007). Subsequent three GWA analyses 386,731, 393,453, and 315,635 SNPs respectively (Zeggini et al., 2007; Saxena et al., 2007; Scott et al., 2007). Common variants in CDKAL1, IGF2BP2, CDKN2A/B were significantly associated with T2DM risk, with the allele-associated odds ratios ranging from 1.07 to 1.48. These studies also confirmed the T2DM effects of SLC30A8 and HHEX. In another GWA, Steinhorsdottir et al. (2007) found that variant rs7756992 in the CDKAL1 gene was significantly associated with T2DM risk in individuals of European ancestry (allele-specific OR = 1.20 95% CI 1.13-1.27) and Han Chinese ancestry (OR = 1.25 95% CI 1.11-1.40), but not in those of African ancestry. Salonen et al. (2007)

<table>
<thead>
<tr>
<th>Locus</th>
<th>Chromosome</th>
<th>Year of discovery</th>
<th>Approach</th>
<th>Probable mechanism</th>
<th>variants</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPARγ</td>
<td>3</td>
<td>2000</td>
<td>Candidate</td>
<td>Insulin action</td>
<td>rs1801282</td>
<td>1.14</td>
</tr>
<tr>
<td>KCN11</td>
<td>11</td>
<td>2003</td>
<td>Candidate</td>
<td>β-cell dysfunction</td>
<td>rs6251</td>
<td>1.14</td>
</tr>
<tr>
<td>TCF7L2</td>
<td>10</td>
<td>2006</td>
<td>Linkage scan</td>
<td>Altered BMI</td>
<td>rs8050136</td>
<td>1.17</td>
</tr>
<tr>
<td>HHEX/IDE</td>
<td>10</td>
<td>2007</td>
<td>GWA</td>
<td>β-cell dysfunction</td>
<td>rs1111875</td>
<td>1.15</td>
</tr>
<tr>
<td>SLC30A8</td>
<td>8</td>
<td>2007</td>
<td>GWA</td>
<td>β-cell dysfunction</td>
<td>rs1326634</td>
<td>1.15</td>
</tr>
<tr>
<td>CDKAL1</td>
<td>6</td>
<td>2007</td>
<td>GWA</td>
<td>β-cell dysfunction</td>
<td>rs1094638</td>
<td>1.14</td>
</tr>
<tr>
<td>CDKN2A/2B</td>
<td>9</td>
<td>2007</td>
<td>GWA</td>
<td>β-cell dysfunction</td>
<td>rs1081161</td>
<td>1.20</td>
</tr>
<tr>
<td>IGF2BP2</td>
<td>2</td>
<td>2007</td>
<td>GWA</td>
<td>β-cell dysfunction</td>
<td>rs4402963</td>
<td>1.14</td>
</tr>
<tr>
<td>HNF1B</td>
<td>17</td>
<td>2007</td>
<td>Large-scale candidate-pathway</td>
<td>Unknown</td>
<td>rs430796</td>
<td>1.10</td>
</tr>
<tr>
<td>WFS1</td>
<td>4</td>
<td>2007</td>
<td>Large-scale candidate-pathway</td>
<td>Unknown</td>
<td>rs1001031</td>
<td>1.12</td>
</tr>
<tr>
<td>JAZF1</td>
<td>7</td>
<td>2008</td>
<td>GWA</td>
<td>β-cell dysfunction</td>
<td>rs864745</td>
<td>1.10</td>
</tr>
<tr>
<td>CDC123/CAMK1D</td>
<td>10</td>
<td>2008</td>
<td>GWA</td>
<td>β-cell dysfunction</td>
<td>rs12779970</td>
<td>1.11</td>
</tr>
<tr>
<td>TSPAN8/LGR5</td>
<td>12</td>
<td>2008</td>
<td>GWA</td>
<td>β-cell dysfunction</td>
<td>rs7961581</td>
<td>1.09</td>
</tr>
<tr>
<td>THADA</td>
<td>2</td>
<td>2008</td>
<td>GWA</td>
<td>Unknown</td>
<td>rs7576597</td>
<td>1.15</td>
</tr>
<tr>
<td>ADAMTS9</td>
<td>3</td>
<td>2008</td>
<td>GWA</td>
<td>Unknown</td>
<td>rs604083</td>
<td>1.08</td>
</tr>
<tr>
<td>NOTCH2</td>
<td>2</td>
<td>2008</td>
<td>GWA</td>
<td>β-cell dysfunction</td>
<td>rs1092393</td>
<td>1.13</td>
</tr>
<tr>
<td>KCNQ1</td>
<td>11</td>
<td>2008</td>
<td>GWA</td>
<td>β-cell dysfunction</td>
<td>rs2373892</td>
<td>1.29</td>
</tr>
<tr>
<td>MTNR1B</td>
<td>11</td>
<td>2008</td>
<td>GWA</td>
<td>Unknown</td>
<td>rs1083963</td>
<td>1.09</td>
</tr>
</tbody>
</table>

*ADAMTS9: ADAM metalloproteinase with thrombospondin type 1 motif 9; CAMK1D: Calcium/calmodulin-dependent protein kinase 1D; CDC123: Cell division cycle 123 homologue (Saccharomyces cerevisiae); CDKAL1: CDK5 regulatory subunit-associated protein-1-like1; CDKN2A/2B: Cyclin-dependent kinase inhibitor 2A/2B; FTO: Fat mass and obesity associated; HHEX: Haematoepoietically expressed homeobox; HNF1B: Hepatocyte nuclear factor 1 homeobox; I: IDE: Insulin degrading enzyme; IGF2BP2: Insulin-like growth factor 2 mRNA binding protein 2; JAZF1: Juxtaposed with another zinc finger gene 1; KCN11: Potassium inwardly rectifying channel, subfamily J, member 11; KCNQ1: Potassium voltage-gated channel, KQT-like subfamily, member 1; LGR5: Leucine-rich repeat-containing G-protein coupled; NOTCH2: Notch homologue (Drosophila); PPARγ: Peroxisome proliferator-activated receptor gamma; SLC30A8: Solute carrier family 30 (zinc transporter), member 8; TCF7L2: Transcription factor 7 like 2; THADA: Thyroid adenoma associated; TSPAN8: Tetraspanin 8; WFS1: Wolfram syndrome 1; MTNR1B: Melatonin receptor 1B. Estimates of effect size (given as per-allele odds ratios) reported for European descent populations based on available data (Fig. 1).
Fig. 1: Effect sizes of 19 common T2DM susceptibility loci. The x-axis gives the year that published evidence reached the levels of statistical confidence that are now accepted as necessary for genetic association studies. The y-axis gives the per-allele odds ratio (estimated for European-descent samples) for each locus listed on the y-axis. Loci are sorted by descending order of per-allele effect size from TCF7L2 (1.37) to ADAMTS9 (1.09) (Table 1). Loci shown in white and dark grey are those identified by GWA approaches and by Linkage scan respectively, whereas those identified by candidate-gene approaches and by large-scale candidate-pathway studies are shown in black and light grey, respectively. Odds ratios are as in Refs (Zeggini et al., 2007; Scott et al., 2007; Frayling et al., 2007; Franks et al., 2008; Zeggini et al., 2008; Yasuda et al., 2008; Prokopenko et al., 2009).

Table 2: GWA scans for type 2 diabetes

<table>
<thead>
<tr>
<th>Study</th>
<th>No. of cases</th>
<th>No. of controls</th>
<th>Sample source</th>
<th>Genotyping array</th>
<th>T2DM phenotype</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diabetes gene group</td>
<td>694</td>
<td>645</td>
<td>France</td>
<td>Illumina 200K + Illumina 100K</td>
<td>Family history of T2DM, AAO &lt; 45 years, BMI &lt; 2 kg m⁻²</td>
<td>Sladik et al., discovery (2007)</td>
</tr>
<tr>
<td>Diabetes Genetics Initiative (DGI)</td>
<td>1464</td>
<td>1467</td>
<td>Finland, Sweden</td>
<td>Affymetrix 500K</td>
<td>Partial enrichment for family history and lean T2DM</td>
<td>Saxena et al. (2007)</td>
</tr>
<tr>
<td>Wellcome Trust Case Control Consortium (WTCCC)</td>
<td>1924</td>
<td>2038</td>
<td>UK</td>
<td>Affymetrix 500K</td>
<td>Enrichment for family history of T2DM, AAO &lt; 65 year</td>
<td>Zeggini et al. (2007)</td>
</tr>
<tr>
<td>Finland-UK Investigation of NIDDM Genetics (FUSION)</td>
<td>1161</td>
<td>1174</td>
<td>Finland</td>
<td>Illumina 200K</td>
<td>Partial enrichment for family history</td>
<td>Scott et al. (2007)</td>
</tr>
<tr>
<td>deCODE Genetics</td>
<td>1399</td>
<td>5273</td>
<td>Iceland</td>
<td>Illumina 500K</td>
<td>No specific enrichment for family history, young AAO or BMI</td>
<td>Steinhauss v. al. et al. (2007)</td>
</tr>
<tr>
<td>Diagene</td>
<td>500</td>
<td>497</td>
<td>East Finland, Germany, UK, Ashkenazi</td>
<td>Illumina 300K</td>
<td>Enrichment for family history, AAO&lt;60 years</td>
<td>Sankai et al. (2007)</td>
</tr>
<tr>
<td>Pima</td>
<td>300</td>
<td>334</td>
<td>Pima Indians</td>
<td>Affymetrix 100K</td>
<td>AAO&lt;25 years, enrichment for family history</td>
<td>Hwang et al. (2007)</td>
</tr>
<tr>
<td>Starr County, Texas Framingham</td>
<td>281</td>
<td>289</td>
<td>Mexican Americans, Massachusetts</td>
<td>Affymetrix 100K</td>
<td>Controlled for admixture</td>
<td>Hayes et al. (2007)</td>
</tr>
<tr>
<td>Health Study</td>
<td>91</td>
<td>1087</td>
<td></td>
<td>Affymetrix 100K</td>
<td>Incident T2DM cases</td>
<td>Florez et al. (2007)</td>
</tr>
<tr>
<td>Old Order Amish</td>
<td>124</td>
<td>289</td>
<td>Amish</td>
<td>Affymetrix 100K</td>
<td>Enrichment for family history</td>
<td>Ranganathan et al. (2007)</td>
</tr>
<tr>
<td>Japanese multi-disease collaborative genome scan</td>
<td>187</td>
<td>1504</td>
<td>Japanese</td>
<td>JSNP Genome Scan 100K SNPs</td>
<td>T2DM cases</td>
<td>Yasuda et al. (2008)</td>
</tr>
<tr>
<td>BioBank Japan</td>
<td>194</td>
<td>1556</td>
<td>Japanese</td>
<td>Custom set of ~268K SNPs</td>
<td>Enrichment for T2DM cases with retinopathy</td>
<td>Unoki et al. (2008)</td>
</tr>
</tbody>
</table>

*AAO: Age at onset*
analyzed 315,917 HapMap-derived tagging SNPs in a two-stage study of 3,073 T2DM cases and 3,273 healthy controls. SNPs in the **AHU1-LOC441171** region were found to confer ~30% increased risk. Other three GWAS genotyping 100 k SNPs in the Framingham Heart Study (Florez et al., 2007), Amish (Rampersaud et al., 2007) and American Indians (Hanson et al., 2007) reported some genetic variants related to the risk of T2DM. However, these associations were not univocally observed. Yasuda et al. (2008) reported that the C allele of SNP rs2237892 in the **KCNQ1** gene is associated with an increased risk for developing T2DM in Japanese population. It is surprising that rs2237892 was associated with T2DM not only in the original Japanese set of cases and controls, but also in Chinese, Korean and European samples. At the same time Unoki et al. (2008) found similar evidence for SNPs closely related to rs2237892 (rs2237895 and rs2237897) in the **KCNQ1** gene were strongly associated with T2DM in the Singaporean population of East Asian descent and the Danish population of European descent. These studies have demonstrated how studies in non-European descent populations can reveal novel susceptibility loci.

Strong signals for T2DM were observed for **FTO**, which is also highly associated with fat mass and obesity (Dina et al., 2007). The effect of **FTO** variants on T2DM risk has been replicated and seems to be mediated entirely by their marked effect on adiposity (Frayling et al., 2007). Melatonin receptor 1B (**MTRNR1B**) gene is located on human chromosome 11q21-q22 and encodes one of two high affinity forms of a receptor for melatonin, the primary hormone secreted by the pineal gland. This gene product is an integral membrane protein that is a G-protein coupled 7-transmembrane receptor, its predominant expression in retina and brain (Reppert et al., 1995). GWAS have shown that variation in **MTRNR1B** is associated with insulin and glucose concentrations. One of the strongest signals for glucose-stimulated insulin secretion in the DGI scan emanated from a SNP (rs10830963) in **MTRNR1B** (p = 7x10^-4, rank order 595) (Saxena et al., 2007). Given that the melatonin pathway had previously been suggested to be involved in pathogenesis of T2DM, the **MTRNR1B** gene was a prime candidate gene for T2DM. The strong signal was observed at rs10830963, where each G allele was associated with an increase of (OR = 0.07 95% CI 0.06-0.08) mmol L^-1 in fasting glucose concentrations (p = 3.2x10^-9) and reduced β-cell function as measured by homeostasis model assessment in ~24,000 participants from the ten studies (HOMA-B, p = 1.1x10^-17). The same allele was associated with an increased risk of T2DM (OR = 1.09 95% CI 1.05-1.12 per G allele p = 3.3x10^-9) in large-scale meta-analysis of 13 case-control studies including data from GWAS, totaling 18,236 cases and 64,453 controls (Prokopenko et al., 2009). The rs10830963 variants in **MTRNR1B** seem to have a more marked effect on risk of T2DM, the effect size being comparable in magnitude (OR = 1.09 95% CI 1.05-1.12) to several other T2DM susceptibility genes recently identified in GWAS. Lyssenko et al. (2009) provided strong support for a role of melatonin and its receptor **MTRNR1B** in the pathogenesis of T2DM. A common variant in the **MTRNR1B** receptor was associated with an increase in fasting glucose over time and predicted future T2DM, most likely through impairment of insulin secretion from the pancreatic β-cell function. The result of GWA meta-analysis for fasting plasma glucose indicated that the **MTRNR1B** rs1387153 strongly modulates fasting plasma glucose in the European population (β = 0.06 mmol L^-1, p = 7.6x10^-10, n = 16,094) and increases the risk for T2DM (OR = 1.15 95% CI 1.08-1.22; p = 6.3x10^-4, cases n = 6,332) (Bouatia-Naji et al., 2009). Effects of melatonin are mediated by two distinct receptors, **MTRNR1A** and **MTRNR1B** (Pandi-Perumal et al., 2008), which are members of the G-protein coupled receptor family, specifically inhibitory G-proteins (Gi). Both receptors have been found to be expressed in human and rodent islets (Muhlbaier and Peschke, 2007), with **MTRNR1A** predominating, especially in glucagon-producing α-cells (Ramacheya et al., 2008). There is some evidence that melatonin may exert an effect on insulin secretion, in that acute effects exerted by αMP-elevating agents are inhibited by melatonin, whereas prolonged effects of the hormone may be stimulatory (Peschke, 2008).

The recent replication studies confirmed significant associations between SNPs within the **HHEX**, **CDKN2A/B**, **CDKAL1** and **KCNQ1** genes and T2DM in the Korean population (Lee et al., 2008), **HHEX/IDE**, **CDKN2A/B** and **IGF2BP2** loci in Dutch subjects (Grarup et al., 2007), **CDKAL1**, **IGF2BP2**, **CDKN2A/B**, **HHEX** and **SLC30A8** genes and T2DM in a Japanese population (Omori et al., 2008), **CDKAL1**, **IGF2BP2**, **HHEX** and **FTO** genes and T2DM in German KORA 500 K study population (Herder et al., 2008), **ADAMTS9**, **CDKAL1**, **CDKN2A/B**, **FTO**, **IGF2BP2**, **JAZF1** and **SLC30A8** in population-based study of Caucasian (Van Hoek et al., 2008), **CDKAL1**, **CDKN2A/B**, **IGF2BP2**, **SLC30A8** and **KCNQ1** genes independently or additively contribute to T2DM risk in the Chinese Han population (Wu et al., 2008; Hu et al., 2009; Liu et al., 2008, 2009), **TCF7L2** (rs12255372), **CDKAL1** (rs7756992, rs7758480), **HHEX** (rs7923873), **IGF2BP2** (rs4402960 and rs1407579), **CDKN2A/B** (rs10811661), and **SLC30A8** (rs13266634) were significantly associated with T2DM in Japanese population (Tabara et al., 2009). The association of
SLC30A8, HHEX, CDKAL1, CDKN2A/CDKN2B, IGF2BP2 and FTO with risk for T2DM was confirmed, in the recent large-scale study of Asian ancestry from Hong Kong and Korea, with odds ratios ranging from 1.15 to 1.35 (1.35×10^{-13}, p = 0.016) (Ng et al., 2008). It's very recently reported that the common variants in the CDKAL1, SLC30A8, HHEX, EXT2, IGF2BP2, LOC387761 and CDKN2B genes did not confer a significant risk for T2DM in Pima Indians (Rong et al., 2009). In the prediction study, common variants in TCF7L2 (rs7903146), PPARC (rs1801282), FTO (rs9999669), KCNJ11 (rs5219), NOTCH2 (rs10925931), WFS1 (rs10010131), CDKAL1 (rs7754840), IGF2BP2 (rs4402960), SLC30A8 (rs1326634), JAZF1 (rs864745), HHEX (rs1111875) were significantly associated with risk of future T2DM (OR = 1.30; p = 9.5×10^{-15}, OR = 1.20; p = 4.0×10^{-4}, OR = 1.14; p = 9.2×10^{-5}, OR = 1.13; p = 3.6×10^{-4}, OR = 1.13; p = 0.02, OR = 1.12; p = 0.001, OR = 1.11; p = 0.04, OR = 1.10; p = 0.008, OR = 1.10; p = 0.008, OR = 1.08; p = 0.03 and OR = 1.07; p = 0.03, respectively) in the MPP cohort (Lyssenko et al., 2008).

Additional loci from Meta-analysis: Efforts to find additional T2DM susceptibility loci have to contend with the modest effect sizes anticipated and the stringent significance thresholds required when many hundreds of thousands of SNPs are tested in parallel. The obvious solution is to increase sample size and the most effective strategy for this involves combining existing GWAS data through meta-analysis. The Diabetes Genetics Replication And Meta-analysis (DIAGRAM) consortium integrated data from three previously published GWAS ( Saxena et al., 2007; Scott et al., 2007; Zeggini et al., 2007), thereby doubling the sample size compared to the largest of the individual studies to ~4500 cases and 5500 controls. The consortium also used novel imputation approaches (Marchini et al., 2007) to infer genotypes at additional SNPs that were not directly typed on the commercial arrays used for the original GWAS, thereby extending the analysis to a total of ~2.2 million SNPs across the genome. In this study, 69 signals showing the strongest associations in the GWAS meta-analysis were genotyped in an initial replication set of 22426 individuals and the top eleven signals emerging from this second analysis were then evaluated in ~57 000 further subjects. After integrating data from all study subjects, six signals reached combined levels of significance, including the JAZF1 (rs864745) (p = 5.0×10^{-12}), CDC123-CAMK1D (rs12779790) (p = 1.2×10^{-6}), TSPAN8-LGR5 (rs961581) (p = 1.1×10^{-6}), THADA (rs7578597) (p = 1.1×10^{-6}), ADAMTS9 (rs4607103) (p = 1.2×10^{-6}) and NOTCH2 (rs10925931) (p = 4.1×10^{-6}) (Zeggini et al., 2008).

However, the replication study of these six SNPs in Khatri Sikh diabetes of North India, only CDC123/CAMK1D (rs12779790) revealed a significant evidence of association with T2DM (p = 0.031) (Sanghera et al., 2009).

**NEW GENES NEW AETIOLOGY**

The strongest signal for T2DM in GWAS was found for TCF7L2 with p-values down to <10^{-8} increase in the odds for the disease of 37% (Scott et al., 2007; Saxena et al., 2007), the TCF7L2 rs7901695, increases disease risk with an odds ratio of 1.37 (Table 1; Fig. 1). The expression of TCF7L2 was related to genotype and metabolite parameters in human islets, the risk T allele was associated with impaired insulin secretion, incretin effects, and enhanced rate of hepatic glucose production. Over expression of TCF7L2 in human islets reduced glucose-stimulated insulin secretion (Lyssenko et al., 2007). The HHEX region which also harbors IDE encoding insulin-degrading enzyme which has been implicated in both insulin signal and islet function. T2DM risk allele for HHEX/IDE gene associated with decreased pancreatic β-cell function, including decreased β-cell glucose sensitivity that relates insulin secretion to plasma glucose concentration (Pascoe et al., 2007). Both HHEX and IDE are critical for ventral pancreas development and are powerful biological candidates for T2DM.

A central theme for many of the recently discovered genes is that many of them seem to be involved in insulin secretion, pin-pointing the pivotal role of β-cell function in the pathogenesis of T2DM. These genes include TCF7L2, KCNJ11, HHEX, SLC30A8, CDKAL1, CDKN2A/2B, IGF2BP2 and KCNQ1 (Steinthorsdottir et al., 2007; Grarup et al., 2007; Pascoe et al., 2007; Staiger et al., 2007, 2008; Yasuda et al., 2008; Unoki et al., 2008). The loci seem particularly to be associated with an increased risk of developing T2DM through a reduced insulin-secretory capacity. A further highly interesting finding was the fat mass and obesity associated FTO gene which predisposes to T2DM by altering Body Mass Index (BMI). It is expressed in hypothalamus (Frayling et al., 2007) which is the key brain region for influencing appetite. Interestingly, the GWAS have identified a separate set of SNPs that seem to represent the strongest common genetic risk factor for heart disease (myocardial infarction) (Zeggini et al., 2007; Helgadottir et al., 2007; McPherson et al., 2007). There is no correlation between the T2DM signal and the heart disease signal, but the latter does fall closer to the CDKN2 genes, which
encodes p16INK4a. Over expression of p16INK4a results in a decreased islet proliferation and β-cell dysfunction in ageing mice (Krishnamurthy et al., 2006), the initial human physiology studies have not provided any evidence that the risk alleles alter insulin secretion, but the mouse phenotype strongly implicates β-cell dysfunction. CDKAL1 is highly expressed in human islets (Zeggini et al., 2007). CDKAL1 shares homology with the CDK5RAP1 gene, a known inhibitor of CDK5 activation. CDK5 is implicated in reduced β-cell function, through the formation of p35-CDK5 complexes, which down regulate insulin expression (Ubeda et al., 2006; Wei et al., 2005). T2DM risk allele for CDKAL1 gene associated with decreased pancreatic β-cell function, including decreased β-cell glucose sensitivity that relates insulin secretion to plasma glucose concentration (Pascoe et al., 2007). The IGF2BP2 binds to the key growth and insulin signaling molecule insulin-like growth factor 2 (IGFII) and is also expressed in the pancreatic islet (Zeggini et al., 2007). T2DM risk for CDKAL1, SLC30A8, IGF2BP2, and LOC387761 is specifically mediated through defects in insulin secretion (Palmer et al., 2008). The product of KCNQ1 gene can form heteromultimers with two other potassium channel proteins, KCNE1 and KCNE3. The SNP rs2237892 in KCNQ1 has known roles in cardiac muscle, it is not yet known how this gene could affect the risk for T2DM, but there is evidence that it is turned on in the pancreas (Yasuda et al., 2008).

Curiously, some of the recent studies suggest a possible explanation for previous much debated epidemiological observations that men with T2DM are less likely to develop prostate cancer. The same allele in HNF1B that predisposes to T2DM was protective of prostate cancer (Gudmundsson et al., 2007). Moreover, different variants in JAZF1 are associated with T2DM and with prostate cancer (Saxena et al., 2007; Zeggini et al., 2007; Scott et al., 2007; Thomas et al., 2008). In keeping with an effect on development and transcriptional processes these findings may not come as a surprise although the exact causal relationships remain to be investigated (Frayling et al., 2008).

Moreover, carriers of novel T2DM risk alleles within JAZF1, CDC123/CAMK1D, and TSPAN8 it’s suggested an impaired pancreatic β-cell function in glucose-tolerant regions in the cohort of middle-aged people (Grarup et al., 2008). CDC123 is regulated by nutrient availability in S. cerevisiae and has a role in cell cycle regulation. Taken together, evidence from GWAS implicating variants in or near CDKAL1, CDKN2A/B, CDC123 and CAMK1D suggests that cell cycle dysregulation may be a common pathogenetic mechanism in T2DM (Ridderstråle and Groop, 2009). Notch homologue 2, Drosophila (NOTCH2) is known to be involved in pancreatic development, but the mechanisms involved for the ADAMTS9 and THADA genes remain unclear.

MTN1B is thought to participate in light-dependent functions in the retina and in melatonin’s neuronal regulation of circadian rhythmicity and sleep cycles. As certain sleep disorders, such as obstructive sleep apnea, result from obesity and are associated with insulin resistance (De Sousa et al., 2008; Kashyap and Defronzo, 2007). Very recently, three new studies identified and reported extremely strong, incontrovertible evidence that the MTN1B is associated with high fasting glucose levels and increased risk of T2DM (Prokopenko et al., 2009; Bouatia-Naji et al., 2009; Lyssenko et al., 2009). MTN1B could represent a new interesting candidate gene linking sleep disorders with T2DM. These new data implicate an association between the sleep-wake rhythm, the so-called circadian rhythm, and fasting glucose levels, T2DM, which was not known previously. The greatest benefit of T2DM genetic study is likely to come from new and better therapies derived from an improved understanding of the aetiology of the disease.

**PREDICTION OF RISK OF T2DM**

One of objectives for identification of T2DM susceptibility genes is to predict T2DM risk and identify high-risk subjects. To predict the risk of T2DM for a healthy individual we need to know and be able to measure risk factors, their effect sizes and how they interact. Although, prediction of total risk is an ultimate goal, prediction of genetic risk that can be attributed to inherited genetic variants is an important component. Generally, the predictive value of genetic testing for prediction of T2DM is unclear. Several empirical studies on the predictive value of genetic polymorphisms have been conducted before the GWAS data were available (Lyssenko et al., 2005; Weedon et al., 2006). In a case-control study, combining the information of three polymorphisms improved disease prediction, albeit to a limited extent (Weedon et al., 2006). The predictive value was low compared with clinical characteristics (Vaxillaire et al., 2008). The complex disease is caused by multiple genetic variants, the predictive testing based on a single genetic marker will be of limited value. GWAS have dramatically increased the number of common genetic variants that are robustly associated with T2DM. The predictive value could be improved by combining multiple common low-risk variants, but all showed limited predictive value so far (Janssens et al., 2006, 2007;
Yang et al., 2003; Wray et al., 2007). The currently known and replicated genetic variants found in GWAS contributed modestly to the prediction of T2DM in population-based setting and marginally improved the risk prediction beyond clinical characteristics (Van Hock et al., 2008). Individuals carrying more risk alleles had a higher risk of T2DM and the common risk variants for T2DM do not provide strong predictive value at a population level (Lango et al., 2008). Similarly Lin et al. (2009) constructed an additive genetic score using the most replicated SNPs within 15 T2DM susceptibility genes and reported that the weighted 15 SNP-based genetic score provides additional information over clinical predictors of prevalent T2DM. However, the clinical benefit of this genetic information is limited. The prediction model for T2DM may not be so useful but has some value and incorporation of data from additional risk loci is most likely to increase the predictive power (Miyake et al., 2009). New gene discoveries from GWAS will certainly identify novel etiological pathways and novel intermediate biomarkers, which consequently may be stronger predictors of disease than the genetic variant that led to its identification. But it is likely that the complexity of complex diseases may ultimately limit the opportunities for accurate prediction of disease in asymptomatic individuals as unraveling their complete causal pathways may be impossible (Janssens and van Duijn, 2008). More extensive studies are needed to assess the usefulness of combining information from multiple variants.

CONCLUSION

Studies in human genetics have made tremendous strides in discovering T2DM genes especially through GWAS. Identification of the causal variants responsible for the association signals uncovered will provide valuable clues to understanding disease predisposition. GWAS is only the beginning in establishing the role of genetic variants in disease etiology. This would be followed by the fine-mapping of the susceptibility region through deep sequencing in large population and the validation of causality for genetic variants in experimental settings.

- Fine-map the new T2DM gene regions. That will involve deep sequencing and further rounds of genotyping to build up a full picture of all the possible common variation that might explain the association signals. This should include efforts to define Copy Number Variants (CNVs) such as duplications and deletions, and should also attempt to define independent associations in the same gene regions
- Additional association studies of the new variants are needed. Investigators will need to assess their role in other populations, especially populations with a high prevalence of T2DM. Further studies of the role of risk alleles in the general population are also important.
- The predictive value of genetic testing for prediction of T2DM is unclear. Additional studies are needed to identify and replicate new genetic susceptibility variants and gene-gene and gene-environment interactions to approach levels of discriminative accuracy that enable the identification of at-risk groups and to assess whether individuals with extreme numbers of risk alleles may benefit from genetic testing.
- The mechanisms whereby a given DNA change leads to an increased risk of T2DM need to be reconstructed. We need to know whether they influence T2DM predisposition through primary effects on β-cell function, through insulin action, or by some other mechanism. The ultimate objective is, of course, to understand how genetic findings can translate into advances in clinical management. In principle, genetic testing might offer insights into disease risk and predict response to the various therapeutic and preventative options available, but much work will be required to understand how to deploy such tests in clinically effective ways. A route is to apply the insights gained into the mechanisms of disease predisposition to identify new targets for drug development. In this respect, genes of small effect are likely to provide clues just as valuable as those of large effect.

Looking ahead, new methodologies and approaches may be needed to discover the remaining as yet unidentified genetic contributors to disease risk. At associated loci, fine mapping can help narrow down the list of possible causal variants and simplify future functional studies. Additional GWAS, in larger samples and multiple ethnicities, will almost certainly lead to new discoveries and incremental gains in the amount of risk accounted for by identified genetic variants. Exploration of these novel loci will very likely uncover additional alleles, both common and rare, that explain additional variance in phenotype, help pinpoint which gene(s) are responsible for the association and provide better clinical and molecular tools for assessing function and mechanism of disease.
REFERENCES


