Genetics of Type 2 diabetes

Torben Hansen
Steno Diabetes Center and Hagedorn Research Institute, Niels Steensens Vej 2, DK 2820, Gentofte, Denmark

Type 2 diabetes is a genetically heterogenous disease consisting of several sub-groups with various combinations of susceptibility genes. The last decade has revealed a molecular understanding of several monogenic forms of Type 2 diabetes. Much less is known about the genetic predisposition which is required for the onset of the more common form(s) of the disease, although much is known about environmental factors contributing to Type 2 diabetes, such as obesity, sedentary lifestyle, smoking, and certain drugs. Identification of genes predisposing individuals to develop Type 2 diabetes will facilitate early diagnosis, effective treatment and intervention. The genetic susceptibility must involve the key elements in the pathogenesis of Type 2 diabetes: insulin resistance and deranged β-cell function, and most likely multiple genes are responsible, each contributing a small amount to the overall risk. A more complete molecular description of Type 2 diabetes is a major undertaking, requiring a multidisciplinary effort, including new strategies for patient sampling, phenotyping, genotyping and genetic analysis.

Changes in human behaviour and lifestyle over the last century have resulted in a dramatic increase in the incidence of diabetes worldwide. The number of adults with diabetes in the world will rise from 135 million in 1995 to 300 million in the year 2025 (ref. 1). Current estimates from different countries in Europe and the United States have shown that diabetes and its complications account for 8–16% of the total health costs for society and these will increase dramatically unless major efforts are made to prevent the ongoing epidemic.

Type 2 diabetes (T2D) is a multifactorial disease with both a genetic component and an important non-genetic component(s) which undoubtedly interacts in order to precipitate the diabetic phenotype6. A model for the natural history of the development of T2D illustrating the complex interaction between genetic predisposition and environmental factors is shown in Figure 1. There are convincing arguments to support its partial genetic determination. The lifetime risk of a first-degree relative of a patient with T2D has been estimated at about 35% with the relative risk of diabetes compared with the general population of between three- and fourfold9. Furthermore, twin studies have shown much higher concordance among monozygotic compared with dizygotic twins4–8. Also, differences in the prevalence of the disease correlating with the degree of genetic admixture indicate an important genetic component9,10. The more common forms of late onset T2D show a complex mode of inheritance and segregation studies have supported an oligo-polypgenic inheritance11. Also, prediabetic phenotypes are familial. Impaired insulin action has been suggested to be the primary defect in the prediabetic state in several studies12–14, but also pancreatic beta-cell dysfunction has been shown in the prediabetic state15, and a significant familiality of both insulin action and beta-cell function have been demonstrated15–17.

Knowledge of the genetics of T2D would increase our understanding of the complex gene–gene and gene–environment interplay causing the disease and will serve to facilitate early diagnosis, treatment and intervention. The purpose of this review is to summarize recent progress in the field of genetics of T2D.

Monogenic forms of T2D

Most successes in defining the genetic basis of T2D have been obtained for monogenic forms of the disease. Classical positional cloning and/or screening of candidate genes have revealed the genetic basis for 5–10% of T2D.

Maturity-onset diabetes of the young

The most common monogenic form of diabetes is maturity-onset diabetes of the young (MODY). MODY is a monogenic subtype of T2D, characterized by an autosomal dominant inheritance, and an age of onset at 25 years or younger18. Phenotypically, MODY is primarily associated with insulin secretion defects and patients with MODY have normal insulin sensitivity and are in most cases lean19,20. It has been estimated that 2–5% of all patients with T2D actually may have MODY21. Genetic studies have revealed that MODY is genetically heterogeneous and at present there are at least six different MODY forms, entitled MODY1–MODY6 (Table 1). The different forms are caused by mutations in the genes encoding hepatocyte nuclear factor-4α (HNF-4α), glucokinase, hepatocyte nuclear factor-1α (HNF-1α), insulin promoter factor-1 (IPF-1), hepatocyte nuclear factor-1β (HNF-1β) and NeuroD, respectively (Table 1)22–27.
MODY1–3 seem to be the most common forms of MODY. While MODY2 is the least severe form of MODY characterized by mild hyperglycemia with less than 50% of the affected subjects presenting overt diabetes, MODY3 and MODY1 represent more severe forms of diabetes, often evolving to insulin requirement (Table 1). Comparing carriers of MODY mutations in the GCK and HNF-1α genes, it has been demonstrated that the genetic cause of the beta-cell defect results in clear differences in both the fasting glucose and the response to an oral glucose load and this can help prioritising diagnostic genetic testing in MODY.\(^2^9\). Also, up to 10% of diabetic patients originally classified as having T1D could have diabetes caused by mutations in the HNF-1α gene, it has been demonstrated that the genetic cause of the beta-cell defect results in clear differences in both the fasting glucose and the response to an oral glucose load and this can help prioritising diagnostic genetic testing in MODY.\(^2^9\). Additionally, MODY5 is associated with severe kidney disease, including chronic renal failure and renal cyst, however, these abnormalities may precede the development of diabetes and thereby not be a diabetic complication but rather a diabetes-independent consequence of mutations in the MODY5 gene.\(^2^6,3^2\–3^4\).

More MODY gene(s) are to be found. In the UK and France the gene has not been identified in 15–20% of MODY families, but in other ethnic groups and in other European studies more than 50% of MODY families remains MODYX (ref. 28) (own unpublished data). Research clearly needs to continue to define the other genetic subgroups and to characterize these clinically and physiologically. In line with this, it has recently been recognized that an alternative splice-form of HNF-4α is predominant in the β-cell.\(^3^5\). This splice-form utilizes a far upstream promoter of HNF-4α, P2, with an alternative exon 1 that in man is 45 kb 5′ to the previously identified P1 promoter. The P2 promoter contains functional binding sites for HNF-1α, HNF-1β and IPF-1, and a large MODY family with a mutation in the IPF-1 binding site of the P2 promoter, which cosegregated with diabetes (LOD = 3.25) has been described.\(^3^6\). Further studies are needed in order to evaluate the importance of the P2 promoter for MODY and late-onset T2D. Also, identification of new MODY forms could give important clues to the genetics of T2D.
**Maternally inherited diabetes and deafness**

Diseases due to mitochondrial mutations are recognized by a maternal dominant inheritance pattern and variable penetrance due to different proportions of the abnormal gene in each tissue (a phenomenon called heteroplasm). A subtype of T2D featuring the maternal inheritance of diabetes and neurosensory deafness have been shown to be caused by variation within the mitochondrial 4.18 gene (position 3243). The prevalence of this mutation among patients with diabetes varies between racial groups, being highest in Japan, where it accounts for 0.5–3% of diabetics with a family history. Patients with mitochondrial diabetes have early onset T2D that frequently progresses to insulin requirement owing to progressive β-cell loss. They tend to be nonobese and might have evidence of neurological and optic features seen in other mitochondrial disorders. Interestingly, the same mutation leads not only to maternally inherited diabetes and deafness, but also to mitochondrial syndrome MELAS (myopathy, encephalopathy, lactic acidosis and stroke-like episodes). This phenotypic heterogeneity is almost certainly the result of chance differences in the tissue distribution of mutated mitochondria during development. Also, other variants within the mitochondrial genome have been described to be associated with T2D.

**Other monogenic forms**

Other rare monogenic forms of diabetes are known. Over 40 mutations in the insulin receptor gene have been described, many of whom are causing diabetes. Also, Wolfram syndrome, a rare recessive disorder, defined as a combination of familial juvenile onset diabetes mellitus, optic atrophy, diabetes insipidus and deafness have been shown to be caused by various mutations in the WFS1 gene. Furthermore, recent data indicate that variation in the WFS1 gene may influence susceptibility to common forms of type 2 diabetes.

**Candidate genes for T2D**

Genes known to be involved in insulin sensitivity, β-cell function and obesity have been obvious candidates for inherited defects leading to T2D (Figure 1). Also, genes potentially involved in intrauterine growth and components of the metabolic syndrome have been considered as candidate genes for T2D. Since more than 200 biological candidates have been studied, this review will concentrate on those providing the strongest evidence for a susceptibility role (Figure 2).

The screening of candidate genes for nucleotide variants that are associated with T2D is a core component of much diabetes genetics research. Such association analyses seek to demonstrate that a particular variant has higher than expected prevalence on ‘disease-gene-carrying’ chromosomes through, for example, comparison of genotype frequencies between groups of unrelated cases and control subjects. Because such studies rely on linkage disequilibrium for detection of a signal, positive results are only likely at, or very close to, functional variants. Association studies therefore provide a powerful approach for analysis of small regions, but remain problematical for screens of larger genomic regions.

The peroxisome proliferator-activated receptor-γ (PPARγ) is a transcription factor, involved in adipogenesis and in the regulation of adipocyte gene expression and glucose metabolism. Recently two mutations in the ligand-binding domain of PPARγ were found in three Caucasian subjects with severe insulin resistance and T2D (ref. 46). Within a unique domain of PPARγ that enhances ligand-independent activation a prevalent Pro12Ala polymorphism has been identified. The polymorphism has in several studies been shown to be involved in the pathogenesis of obesity and recently, using a family-based design to control for population stratification it was reported that the Ala-allele of the codon 12 polymorphism was associated with decreased risk of T2D (ref. 48). A significant association of the Ala-allele of the Pro12Ala polymorphism with increased whole body insulin sensitivity has been demonstrated. Therefore, it is likely that increased insulin sensitivity is one of the mechanisms by which the Ala-allele protects against T2D. A widespread missense polymorphism in the peroxisome proliferator-activated receptor-γ coactivator-1 (PGC-1) gene, a novel transcriptional coactivator of a series of nuclear receptors including PPARγ is reproducibly associated with T2D (ref. 50). Similarly, a missense variant in the insulin receptor substrate-1 (IRS-1) gene has been shown to be associated with decreased insulin sensitivity and and an impairment of insulin-stimulated PI3-kinase activity, and a rare Pro387Lys variant in the protein tyrosine phosphatase-1B (PTP-1B) gene was shown to be associated with type 2 diabetes probably via an impairment in serine phosphorylation of the PTP-1B protein.
Figure 2. A complex interaction between polymorphisms in susceptibility genes and environmental factors defines subsets of late-onset Type 2 diabetes. Various susceptibility genes with effects on prediabetic phenotypes are shown. References are provided in the text. PPARγ, peroxisome proliferator-activated receptor-γ; PGC-1, peroxisome proliferator-activated receptor-γcoactivator-1; IRS-1, insulin receptor substrate-1; PTP-1B, protein tyrosine phosphatase-1B; HNF-1α, hepatocyte nuclear factor-1α; IPF-1, insulin promoter factor-1; InsVNTR, a regulatory minisatellite upstream of the insulin gene; SUR1/Kir6.2, sulphonylurea receptor/ inward rectifier K+ channel member; B2AR and B3AR, beta2- and beta3-adrenoreceptors; GCK, glucokinase.

have focused on the known interplay between obesity and risk of T2D (refs 65, 66). The largest drawback of such studies is the sample sizes used. Very large study populations are likely to be needed in order to demonstrate interactions between gene variants contributing a minor amount to the overall risk. Furthermore, studies of genes or gene–gene interactions with effects on components of the metabolic syndrome including blood pressure and lipid profiles are likely to be important for definition of subtypes of T2D. Furthermore, the observation that reduced growth in early life is strongly linked with impaired glucose tolerance and non-insulin dependent diabetes lead to the ‘thrifty phenotype’ hypothesis. Malnutrition during intra-uterine life may cause a ‘programming’ of the fetus inducing changes in important metabolic pathways leading to increased risk of T2D later in life67,68. However, genetic variation may also influence both fetal growth and risk of T2D. Therefore, predisposition to T2D is likely to be the result of both genetic and fetal environmental factors69–71.

Genome scans

For monogenic forms of T2D, linkage analysis, recombinant mapping in families and positional cloning have proved to be powerful tools to define genomic T2D regions, to further refine the location and ultimately to identify the gene encoding the disease. These findings have encouraged researchers to use genome-wide scans for linkage to identify the genetic basis for the more common forms of T2D. Given a λs of 2–4 for T2D, imposing a limit on the magnitude of any single gene effect it might be envisioned that most such studies would be underpowered to detect any gene effects. However, the first published genome scan on T2D on 170 Mexican American-affected sib-pair families showed significant evidence of linkage to T2D near the terminus of 2q (ref. 72). Using linkage disequilibrium (LD) mapping, polymorphisms within the gene encoding calpain-10 (CAPN10) were found to explain the previous described linkage and to be associated with T2D (ref. 73). Further analysis in Mexican American and European populations indicate that the disease susceptibility is best described by a combination of risk haplotypes. Also, a contribution to susceptibility to T2D in Mexican Americans is due to an interaction of genes on chromosomes 2q and 15 (ref. 74). Calpains are ubiquitously expressed cysteine proteases that are thought to regulate a variety of normal cellular functions, and recent data indicate a role for calpains in the regulation of both insulin secretion and insulin action75.

The results from several additional genome scans for T2D are now available76–84, and several replicated chromosomal regions have been described suggesting that genes contributing risk for diabetes exist on chromosome 1q21-q23 and on chromosome 12 and 20. Also, other regions have shown ‘bumps’ in several scans for T2D and/or obesity.
Conclusions

The discovery of genes encoding monogenic forms of T2D represents important milestones in the understanding of the disease. It is clear that more monogenic forms of T2D remain to be identified, and collections of families with similar disease phenotypes and distinct inheritance pattern will be of great importance for further description of subtypes of T2D. For the common form(s) of T2D it appears that multiple genes are involved, each contributing a small amount to the overall risk. The identification of susceptibility genes will probably require a combination of different methodologies including expressing profiling of genes in various tissues to define metabolic pathways that might be altered in genetic predisposition to T2D, usage of both polygenic and genetically manipulated animal models, and a detailed knowledge of patterns of linkage disequilibrium in human populations. How do we validate the many polymorphisms in the human genome evolving from various approaches as being of importance for T2D susceptibility? An editorial in Nature Genetics provides guidance on ideal features of a genetic association study: ‘large sample sizes, small P values, reported associations that make biological sense and alleles that affect the gene product in a physiologically meaningful way’. Therefore, in order to dissect the complex etiology of T2D we need to use a wide variety of in vitro and in vivo tools for susceptibility SNP identification and validation. Understanding how genetic variation contributes to disease within populations will require a simultaneous acquisition of detailed genetic and environmental (lifestyle) data from very large population cohorts (Figure 2). Achievement of a detailed description of the genetic risk profile at the population level will allow for identification of individuals at greatest personal risk of future diabetes and presymptomatic intervention can be targeted to those likely to benefit. Furthermore we will get a sound basis for subdividing patients according to predominant pathophysiological defect(s) and predicted therapeutic response. In the near future, rational individualized treatment seems to become an achievable goal.

50. Ek, J. et al., ibid, 2001, 44, 2220–2226.
73. Horikawa, Y. et al., *ibid*, 2000, **26**, 163–175.