Defining the genetic contribution of type 2 diabetes mellitus

Jonathan van Tilburg, Timon W van Haeften, Peter Pearson, Cisca Wijmenga

Abstract
Type 2 diabetes mellitus is a common multifactorial genetic syndrome, which is determined by several different genes and environmental factors. It now affects 150 million people worldwide, but its incidence is increasing rapidly because of secondary factors, such as obesity, hypertension, and lack of physical activity. Many studies have been carried out to determine the genetic factors involved in type 2 diabetes mellitus. In this review we look at the different strategies used and discuss the genome wide scans performed so far in more detail. New technologies, such as microarrays, and the discovery of SNPs will lead to a greater understanding of the pathogenesis of type 2 diabetes mellitus and to better diagnostics, treatment, and eventually prevention.

(J Med Genet 2001;38:569–578)

Keywords: diabetes mellitus type 2; genetic factors; genome screen; candidate gene

Diabetes mellitus (DM) affects over 150 million people worldwide, with a prevalence that varies markedly from population to population. Estimates predict that almost 300 million people will suffer from DM by 2025 (fig 1) with the vast majority being cases of diabetes mellitus type 2. Many risk factors have been identified which influence the prevalence (total number of cases as a percentage of the total population) or incidence (total number of new cases per year as a percentage of the total population). Factors of particular importance are a family history of diabetes mellitus, age, overweight, increased abdominal fat, hypertension, lack of physical exercise, and ethnic background. Several biochemical markers have also been identified as risk factors, including fasting hyperinsulinaemia, increased fasting proinsulin, and decreased HDL cholesterol. Both diabetes mellitus types 1 and 2 show a familial predisposition, which is a strong indication for the involvement of genes in people's susceptibility to the disease. However, the aetiology underlying types 1 and 2 is different and different genes are likely to be involved in each type of diabetes mellitus. The following discussion focuses on a genetic dissection of type 2 diabetes mellitus.

The two most common forms of diabetes mellitus, type 1 and type 2, are both characterised by raised plasma glucose levels. Normal glucose homeostasis depends on the balance between glucose production by the liver and kidneys and glucose uptake by the brain,
insulin secretion by the pancreatic mechanisms. After a meal, a small increase in plasma glucose levels are normally kept enhanced due to its conversion to glycogen and increases the uptake of glucose from the blood, the predominant anabolic hormone involved, kidneys, muscles, and adipose tissue. Insulin, the predominant anabolic hormone involved, increases the uptake of glucose from the blood, enhances its conversion to glycogen and triglycerides by the liver. Increased insulin levels inhibit glucose production by the liver, lipolysis in adipose tissue and proteolysis in muscle. They also inhibit ketogenesis by the liver. Although the brain uses glucose as its main energy source, it can also use ketone bodies when glucose levels are insufficient (for example, during fasting).

There are a number of glucose counter-regulatory hormones, such as glucagon, cortisol, epinephrine, and norepinephrine, which raise plasma glucose levels and therefore counteract hypoglycaemia. The balance between the insulin action and the effects of the counter-regulatory hormones ensures normal glucose homeostasis. Criteria for diabetes have heavily relied on plasma glucose levels after an oral glucose load (usually 75 g glucose in water). Two hour values over 11.1 mmol/l (=200 mg/dl) are still used as diagnostic for diabetes.

This value was originally chosen when prospective studies indicated that subjects with a two hour post-glucose load plasma glucose level of >11.1 mmol/l were at significant risk of developing (diabetic) retinopathy.

The diagnostic criteria for diabetes have recently been modified: a fasting glucose level of 7.0 mmol/l and higher is now sufficient for the diagnosis, since this (fasting) level has been shown to be associated with the two hour post-glucose load plasma glucose levels of >11.1 mmol/l. However, a random plasma glucose level of 11.1 mmol/l and higher is still diagnostic for diabetes mellitus.

Patients with type 1 diabetes mellitus require insulin therapy to prevent diabetic ketoacidosis. Since this form of the disease is usually established before the age of 20, it was formerly referred to as “juvenile onset type diabetes mellitus”. The major cause of type 1 diabetes mellitus is the autoimmune destruction of the pancreatic \( \beta \) cells.

Type 2 diabetes mellitus accounts for around 90% of all cases of diabetes mellitus. Since type 2 diabetes mellitus usually develops after the age of 40, the disease was also called “adult onset type diabetes mellitus”. Unlike type 1 diabetes mellitus, type 2 is not usually caused by autoimmune destruction of the pancreatic \( \beta \) cells, but is characterised by multiple defects in both insulin action and insulin secretion. Both insulin’s inhibitory effect on liver glucose production and its stimulatory effect on peripheral glucose uptake are diminished. Although many type 2 diabetes mellitus patients have a basal hyperinsulinaemia, rises in plasma glucose have a characteristically reduced stimulatory effect on insulin secretion. Type 2 diabetes mellitus patients are often treated by adapting their diet or with oral hypoglycaemic drugs, but many will eventually need exogenous insulin to overcome their hyperglycaemia.

Most patients with type 2 diabetes mellitus are obese, which led to the finding that obesity is associated with diminished insulin action both in the liver and in the periphery. The association between type 2 diabetes mellitus and obesity is probably the result of multiple mechanisms, including rises in plasma free fatty acids (FFA) and tumour necrosis factor-alpha (TNF\( \alpha \)) released from “full” adipocytes. Furthermore, lack of physical exercise is also associated with diabetes mellitus, which led to the finding that exercise enhances the action of insulin, presumably via upregulation of glucose transporters in muscle.

Apart from the short term complications such as thirst, malaise, tiredness, and ketoadosis, diabetes mellitus often leads to a number of long term complications, generally subdivided into micro- and macrovascular complications. It is these long term chronic complications that have the greatest impact on the health and quality of life of patients. The macrovascular complications include retinopathy, nephropathy, and neuropathy, with type 2 diabetes mellitus being one of the main causes of blindness, lower limb amputations, and renal failure in adults. The macrovascular...
complications mean that type 2 diabetes mellitus is a major risk factor for cardiovascular disease and stroke. These chronic complications have a high socioeconomic cost and put a heavy burden on public health services.10

Genetics of type 2 diabetes mellitus

Unlike single gene disorders, where expression of the disease is influenced by a mutant allele at one gene locus, in common diseases like type 2 diabetes mellitus the disease expression depends on many gene loci which all have small to moderate effects. Type 2 diabetes mellitus is a so-called multifactorial disease in which the genes (loci) not only interact with each other but also with environmental factors. It is probable that both insulin activity and secretion are subject to genetic variance at several loci. According to this multifactorial model, predisposition to the disease could be determined by many different combinations of genetic variants (genotypes) and environmental factors; the genetically predisposed subjects will not necessarily develop the overt syndrome unless they are also exposed to particular environmental factors. It is well known that exogenous factors such as age, physical activity, diet, and obesity, play a major role in the disease aetiology of type 2 diabetes mellitus.12

The following demographic observations have shown the effect of changes in environmental factors and the prevalence of type 2 diabetes mellitus has been estimated for various populations. The prevalence spectrum ranges from very low levels of about 1% in some populations, such as tribes of non-Austronesian ancestry in Papua New Guinea or in the Chinese population living in mainland China, to extremely high levels of 50% in Pima Indians (North America). The Pima Indians have changed from a traditional agricultural lifestyle to a sedentary one, with a diet similar to the general US population. However, the large variation in the prevalence of type 2 diabetes mellitus in different populations is probably a result of different environmental as well as genetic determinants. It is particularly interesting to see that the prevalence increases as ethnic groups migrate from lesser developed areas of the world to more urbanised or westernised regions. This is illustrated by the higher prevalence of type 2 diabetes mellitus seen among Japanese who migrated to Hawaii,13, 14 or by the high prevalence (13.1%) among Chinese living on the island of Mauritius compared with the prevalence among Chinese living on mainland China (1.6%).15 In general, there is a trend of increasing prevalence of diabetes mellitus with migration from rural to urban societies,16 but also with a change of environment, though not necessarily associated with a transition from rural to urban. Is simply a change of geographical location sufficient to trigger an increase in type 2 diabetes mellitus?

Twin studies have provided convincing evidence that genetic determinants contribute to the development of type 2 diabetes mellitus.17 Several studies have shown higher concordance rates in monozygotic (MZ) twins than in dizygotic (DZ) twins,17 for example, in a population based cohort of twins in Finland, the concordance rate in MZ twins was 34% whereas in DZ twins it was 16%.18 In a Japanese study, these figures were 83% for MZ twins and 40% for DZ twins.19 Such figures show the difference of environmental influences within populations (that is, the difference between MZ and DZ twins). The large variation in concordance rates between populations may be the result of bias or a different selection of the populations studied, but it may also indicate differences in genetic susceptibility between these populations.20–21 A concordance rate above 80% for MZ twins implies a high degree of heritability for type 2 diabetes mellitus as well as the involvement of environmental factors. In addition, there is a higher relative risk for a relative of a patient with type 2 diabetes mellitus compared with the population prevalence, the so-called λr (relative risk of a relative). For type 1 diabetes mellitus the λr is 20 whereas the λr for type 2 diabetes mellitus is 3.5. This relative risk also increases with the number of affected relatives.22–23 These figures imply that the genetic models involved in the two types of diabetes must be very different. The genetic model for type 1 diabetes mellitus appears to contain at least one major locus providing significant susceptibility but requiring many other contributing factors with equal and additive effects. In contrast, the model for type 2 diabetes mellitus seems more complex, involving more loci and additional environmental factors.24

The search for susceptibility genes in type 2 diabetes mellitus

In our search for a better understanding of the pathogenesis of type 2 diabetes mellitus, a genetic approach will help focus on the underlying causes of the disease, and may provide new information for diagnostic treatment and prevention. This genetic information may also form the basis for new drug therapies, such as individually specific or targeted pharmacotherapy (pharmacogenetics). Two common approaches for distinguishing genetic factors are: (1) the candidate gene approach and (2) the genome wide scan using anonymous polymorphic markers.

(1) The candidate gene approach

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

To date, over 250 candidate genes have been studied for their role in type 2 diabetes
mellitus. The majority of these studies have failed to uncover any association. A minor role for some of the gene products involved in insulin secretion or insulin action, such as IRα, PPARγ, the glucagon receptor, the sulphonylurea receptor (SUR), the peroxisome proliferator activated receptor-γ (PPARγ), and the MAPKBP1 has been observed, but the role for these candidate genes seems to be limited to a small percentage of type 2 diabetes mellitus patients or to specific populations.

There are two plausible explanations: either the genes concerned carry genetic variations which are peculiar to these specific populations and only give rise to type 2 diabetes mellitus in that specific population, or the genetic variances are spread through many populations and only manifest together with type 2 diabetes mellitus because of general genetic background differences between the populations concerned. Although the case-control study design is an easy to implement approach, it also has a history of false positive results. Such false positive associations often occur because of confounding owing to population stratification. This is because population subdivision (or any other form of non-random mating) permits marker allele frequencies to vary among segments of the population, as the result of genetic drift or founder effects. In response to this problem it was decided to use the transmission disequilibrium test (TDT), which looks at the genotypes of the parents of affected subjects. Although this approach takes advantage of population level associations, the TDT is not susceptible to false positive associations that result from stratification.

Unfortunately this approach is not suitable for late onset diseases like type 2 diabetes mellitus because the proband's parents may no longer be alive to give DNA samples. It is intrinsically likely that future genetic research into complex disorders, such as type 2 diabetes mellitus, will also involve genome wide analysis of many gene families to establish the contribution made by the genetic background.

(2) Genome wide scan

One of the major drawbacks of the candidate gene approach is that it will not lead to the identification of entirely new genes or pathways involved in type 2 diabetes mellitus. In order to identify new genes for type 2 diabetes mellitus, genome wide scans using polymorphic markers need to be performed. However, the classical approach of gene localisation by linkage analysis in multigenerational families is not the most suitable strategy for type 2 diabetes mellitus for several reasons. Firstly, there is the lack of a Mendelian inheritance pattern; secondly, the mean age of diagnosis is around 60 years. As a consequence, one or both of the patient’s parents are often no longer available for study. Thirdly, only affected subjects can be used for linkage studies because of the reduced and age dependent penetrance. Hence, it is hard to obtain families with enough type 2 diabetes mellitus patients. In addition, genetic heterogeneity can become a problem as mutations in any one of several genes may result in identical phenotypes, or a chromosomal region may cosegregate with the disease in some families but not in others. A non-parametric analysis method can overcome these problems, since this would require no knowledge of the mode of inheritance of the disease, the disease allele (gene) frequencies, or the penetrance.

A commonly used non-parametric genetic mapping approach is the affected sib pair (ASP) approach using randomly spaced polymorphic markers (usually every 10 cM). The ASP approach is discussed in detail in box 1.
using large numbers of ASPs to obtain sufficient power for detecting linkage for a given value of $\lambda_r$ (relative risk for a sib).\textsuperscript{42, 43, 49} This strategy is also very expensive and it used to be extremely time consuming. However, technological improvements, such as capillary sequencing equipment and faster computers, have decreased the time required enormously.

The most efficient and cost-beneficial way of performing a genome wide scan using ASP is “staged searching”. The initial genome scan (stage 1) is carried out with a sparse marker set (average spacing 20 cM). Regions of interest should exceed the threshold lod score of 1.0. It has been shown that the power exceeds 90% in a sample size of 200 ASPs once the $\lambda_r$ (relative risk for a relative) is greater than 1.7, given a lod of 1.0.\textsuperscript{42, 43} Loci with delicate effects are not missed when a lower threshold is used. However, this strategy also increases the false positive rate. Subsequently, the regions of interest are investigated (stage 2) with a denser marker set (average spacing 5 cM). The threshold for significant linkage would be a lod score of 3.3.\textsuperscript{42, 43} A three stage strategy, with increasing thresholds at each stage, is the most powerful approach to adopt in a genome scan.\textsuperscript{42} An alternative staged strategy, known as sample splitting, is to perform the initial screening on part of the sample and to follow up on interesting loci in the whole sample.\textsuperscript{43, 44}

An efficient study design is an important aspect of any genome wide scan. Different types of cohorts, consisting of nuclear families, multigenerational families, or affected sib pairs, can be used. To date, various research groups have completed or nearly completed genome scans for type 2 diabetes mellitus using ASPs\textsuperscript{51–58} or, occasionally, multigenerational families.\textsuperscript{59–62} Both types of genome scans (using ASPs or multigenerational families) yield varying levels of evidence (table 1).

In 1996, a genome wide significance was found on chromosome 2q37 in a combined data set of 330 Mexican-American ASPs from Starr County, Texas. This locus was designated NIDDM1.\textsuperscript{61} In a sample from Botnia, western Finland, a small number of selected pedigrees with the lowest quartile for mean 30 minute insulin levels after oral glucose tolerance tests showed significant evidence for linkage to type 2 diabetes mellitus on chromosome 12q, and this locus was designated NIDDM2.\textsuperscript{61} More recently, several studies have shown significant evidence for linkage to chromosome 20\textsuperscript{53, 55, 57} and a recent genome scan in Pima Indians showed strong evidence that chromosome 11q contains a susceptibility locus influencing both type 2 diabetes mellitus and obesity. Chromosomes 1q and 7q showed some evidence of additional diabetes mellitus susceptibility loci.\textsuperscript{60} In 42 multigenerational families with northern European ancestry from Utah, significant linkage was found under a model of recessive inheritance on chromosome 1q21-23,\textsuperscript{61} and in 49 ASPs of Canadian Oji-Cree Indian origin, both suggestive linkage and suggestive association was found with chromosomes 6, 8, 16, and 22.\textsuperscript{61} In Mexican Americans from the San Antonio Family Diabetes Mellitus Study, significant evidence was found that a susceptibility locus on chromosome 10q influences age at onset of diabetes mellitus and this locus also seems to be linked to type 2 diabetes mellitus itself.\textsuperscript{62}

Most recently, a genome wide scan in four American populations has indicated suggestive linkage to type 2 diabetes mellitus or impaired glucose homeostasis on chromosomes 5, 12, and X in whites, on chromosome 3 in Mexican Americans, and chromosome 10 in Afro-Americans.\textsuperscript{63} In an eastern and southern Chinese Han population, two loci in a region on chromosome 9 showed suggestive evidence for linkage to type 2 diabetes.\textsuperscript{63}

All these different findings need to be replicated in additional type 2 diabetes mellitus cohorts to strengthen the evidence that true type 2 diabetes mellitus susceptibility genes exist at these loci.\textsuperscript{64}

**After the genome wide scans, then what?**

What can be said about the results from the various genome wide scans? The results suggest that there may be genes on chromosome 1q contributing to the risk for type 2 diabetes mellitus in Pima Indians, which may also be true for chromosome 2 in Mexican Americans and for chromosomes 12 and 20 in...
Table 1  Linkage results of different genome wide scans in type 2 diabetes mellitus

<table>
<thead>
<tr>
<th>Ethnic group</th>
<th>Trait</th>
<th>Sample (cohort)</th>
<th>Locus marker</th>
<th>Lod</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pima Indian</td>
<td>Diabetes mellitus before age 25 years</td>
<td>264 nuclear families</td>
<td>D1S198</td>
<td>4.1</td>
</tr>
<tr>
<td></td>
<td>Age adjusted diabetes mellitus</td>
<td></td>
<td>D6S1009-D6S1003</td>
<td>1.39</td>
</tr>
<tr>
<td></td>
<td>Age adjusted diabetes mellitus</td>
<td></td>
<td>D9S209-D9S2026</td>
<td>1.22</td>
</tr>
<tr>
<td></td>
<td>Age adjusted diabetes mellitus</td>
<td></td>
<td>D11S446-D11S912</td>
<td>1.66</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus</td>
<td></td>
<td>D7S1799</td>
<td>1.8</td>
</tr>
<tr>
<td>Mexican Americans</td>
<td>Diabetes mellitus</td>
<td>330 ASPs</td>
<td>D2S125</td>
<td>4.03</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus</td>
<td>53 nuclear families</td>
<td>D2S432</td>
<td>3.91</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus age at onset</td>
<td>27 extended families</td>
<td>D3S2432</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus age at onset</td>
<td></td>
<td>D3S1566-GATA128C02</td>
<td>2.51</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus age at onset</td>
<td></td>
<td>D9S328-D9S925</td>
<td>2.06</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus age at onset</td>
<td></td>
<td>D10S587-D10S1223</td>
<td>3.75</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus</td>
<td></td>
<td>D3S1566-GATA128C02</td>
<td>2.67</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus</td>
<td></td>
<td>D4S1615-D4S175</td>
<td>1.99</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus</td>
<td>51 ASPs</td>
<td>D9S208-D9S925</td>
<td>2.38</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus</td>
<td>58</td>
<td>D10S587-D10S1223</td>
<td>2.88</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus</td>
<td>26 extended families</td>
<td>D1L1S1349</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus</td>
<td>77 nuclear families</td>
<td>D3S1404</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus (stratified on 30 min insulin)</td>
<td>60 D1S198</td>
<td>D12S853</td>
<td>2.81</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus</td>
<td>55 ASPs</td>
<td>GATA172D05 (X chr)</td>
<td>2.99</td>
</tr>
<tr>
<td>White (North American)</td>
<td>Diabetes mellitus</td>
<td>14 extended families</td>
<td>D2S197</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus</td>
<td>55</td>
<td>PKC1 (chr 20)</td>
<td>2.84</td>
</tr>
<tr>
<td>White (French)</td>
<td>Diabetes mellitus</td>
<td>716 ASPs</td>
<td>D11S937-D11S901</td>
<td>2.04</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus</td>
<td></td>
<td>D20S849-D20S905</td>
<td>1.75</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus</td>
<td></td>
<td>D20S909-D20S107</td>
<td>1.99</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus</td>
<td></td>
<td>D20S886-D20S197</td>
<td>2.04</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus</td>
<td>42 extended families</td>
<td>D1S2432</td>
<td>2.15</td>
</tr>
<tr>
<td>White (Utah)</td>
<td>Diabetes mellitus (recessive model)</td>
<td></td>
<td>CRP-APOE2 (chr 1)</td>
<td>4.3</td>
</tr>
<tr>
<td>Oji-Cree (Canadian)</td>
<td>Diabetes mellitus</td>
<td>49 ASPs</td>
<td>D6S1056</td>
<td>4.24</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus</td>
<td></td>
<td>D9S264</td>
<td>2.91</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus</td>
<td></td>
<td>D16S2616</td>
<td>4.20</td>
</tr>
<tr>
<td>Han (China)</td>
<td>Diabetes mellitus</td>
<td>168 ASPs</td>
<td>D22S683</td>
<td>2.48</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus</td>
<td></td>
<td>D9S171</td>
<td>3.29</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus</td>
<td></td>
<td>D9S161</td>
<td>2.22</td>
</tr>
<tr>
<td></td>
<td>Diabetes mellitus</td>
<td></td>
<td>D9S175</td>
<td>2.94</td>
</tr>
</tbody>
</table>

whites (table 1). The genomic regions described so far, which extend over 20 cM in many cases, now require fine mapping to pinpoint the region of interest and this can be done using linkage disequilibrium (LD) analysis.

LD occurs when a marker allele lies so close to the disease susceptibility allele that these alleles are inherited together over many generations. Thus, the same allele will be detected in affected subjects in multiple, but apparently unrelated families. The genetic mapping has to be followed by testing all the candidate genes from the region for their involvement in the disease and this should result in the positional cloning of a gene associated with type 2 diabetes mellitus. However, this last step will become obsolete because the Human Genome Project will now provide us with a detailed map of all the genes. It has been proven that it is possible to use this approach of genome wide scanning to position clone genes for complex diseases such as type 2 diabetes mellitus. Recently a putative diabetes mellitus susceptibility gene, calpain-10 (CAPN10), was found to be associated with type 2 diabetes mellitus in Mexican Americans, in the NIDDM1 region. This finding suggests a novel pathway that may contribute to the development of type 2 diabetes mellitus. Using single nucleotide polymorphism (SNPs) analysis, genetic variation in CAPN10, a member of the calpain-like cysteine protease family, was found and it appears to affect risk of type 2 diabetes mellitus. However, these findings need to be replicated in other populations and such studies may identify additional variation (SNP) associated with diabetes mellitus within CAPN10. If we consider that there may be approximately 30,000 genes in the human genome, that these genes may have multiple forms, and also interact with each other and environmental factors, this illustrates the magnitude of the problem in searching for type 2 diabetes mellitus susceptibility genes. It is clear that other strategies need to be considered as well as the ones described above.

It is also important to realise that type 2 diabetes mellitus often occurs together with obesity and hypertension, but that each may have its own genetic origin. One approach may therefore be to compare genome wide scans of patients having two or all three diseases with genome wide scans of patients having “only” one of these diseases, preferably in the same ethnic population. Alternative approaches could be used to find disease susceptibility genes and to elucidate the molecular basis of type 2 diabetes mellitus. By using families exhibiting a rare early onset form of the disease, it may be possible to identify genes involved in the disease aetiology. Other alternatives are to study genetic isolates or to use genetically engineered animals and inbred animals. All these alternatives can be valuable tools for understanding the molecular basis of type 2 diabetes mellitus (see box 2).

The discovery of a novel gene and pathway in type 2 diabetes mellitus characterises the importance of conducting genome wide scans in complex diseases like type 2 diabetes mellitus. However, it may be a long time before all the susceptibility genes are found. It may take even more time before their roles in different pathways have been elucidated and the mechanisms involved in their interaction with other factors in the disease aetiology clarified.

The discovery of thousands of SNPs and the construction of a reliable SNP linkage map will certainly be a major factor in the discovery of a
Box 2. Alternative approaches to finding genes

(1) THE USE OF RARE FAMILIES EXHIBITING PHENOTYPES VERY SIMILAR TO TYPE 2 DIABETES MELLITUS

It may be possible to find genes involved in the disease aetiology using rare families exhibiting an early onset form of the disease. In Alzheimer's disease (AD), for example, the use of a familial early onset form showed at least three AD genes.71-73 These genes are now being investigated for new ideas on the mechanisms underlying the pathogenesis of AD. A relatively rare form of diabetes mellitus, maturity onset diabetes mellitus of the young (MODY), is characterised by monogenic, autosomal dominant transmission and early age of onset. Although MODY accounts for only 2-5% of type 2 diabetes cases, by using large families expressing this form of diabetes, it has been possible to identify a number of different genes involved in MODY (table 2).

Another rare and early onset form of diabetes is maternally inherited diabetes and deafness (MIDD), in which mutations are found in the mitochondrial. The implication of mitochondrial mutations in diabetes mellitus is supported by the fact that patients with type 2 diabetes mellitus are more likely to have affected mothers than affected fathers.74 Although, the MODY and MIDD genes found so far provide a good insight into the development of diabetes mellitus, no direct linkage has been found between these genes and the common type 2 diabetes mellitus.

Table 2 Genes involved in MODY

<table>
<thead>
<tr>
<th>Location in gene</th>
<th>Gene Description</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>MODY1 20q22-2q13.1</td>
<td>Hepatocyte nuclear factor-4α (HNF-4α)</td>
<td>87</td>
</tr>
<tr>
<td>MODY2 7p15-p13</td>
<td>Glucokinase (GCK)</td>
<td>88-92</td>
</tr>
<tr>
<td>MODY3 12q24.2</td>
<td>Hepatocyte nuclear factor-1α (TCF1)</td>
<td>87</td>
</tr>
<tr>
<td>MODY4 13q21.1</td>
<td>Insulin promoter factor-1 (IPF1)</td>
<td>93-98</td>
</tr>
<tr>
<td>MODY5 17cen-q11.3</td>
<td>Hepatocyte nuclear factor-1β (TCF2)</td>
<td>99, 100</td>
</tr>
</tbody>
</table>

(2) THE USE OF GENETICALLY ISOLATED POPULATIONS

Another alternative for discovering genes involved in type 2 diabetes mellitus is the use of genetic isolates. The number of disease mutations in an isolated population is assumed to be reduced when the present population is derived from a relatively small number of founders and population expansion has occurred during a period of isolation and rapid population growth and not by immigration. The population has to be large enough to provide a sufficient number of affected subjects for study.75 This approach has been successful for some very rare monogenic diseases. A gene for benign recurrent intrahepatic cholestasis (BRIC) and progressive familial intrahepatic cholestasis type 1 (PFIC1) was mapped and cloned by using two genetic isolates, the Amish in the USA and the population of a fairly isolated fishing village in The Netherlands.76-78 Studying a genetic isolate may provide opportunities for special study designs to identify not only rare Mendelian disease genes, but also major loci contributing to complex diseases, as seen in a genome wide scan of Ashkenazi Jews.79 In this study it was suggested that susceptibility for type 2 diabetes mellitus may be encoded by loci on chromosomes 4q and 20q. The reduced genetic complexity of these genetic isolates means there is a greater contribution from the individual genes. Sub-populations and patient materials from these genetic isolates can be used to perform association studies or linkage analysis.80

(3) THE USE OF AN ANIMAL MODEL EXHIBITING THE PHENOTYPE

Genetically engineered animals and inbred animals can be valuable tools for understanding the molecular basis of type 2 diabetes mellitus.81 Today there are several mice and rat models available for studying both type 2 diabetes mellitus and obesity. By crossing the monogenic mouse (the db/db and the ob/ob mice) with other strains, it might be possible to find modifier genes.82 The use of polygenic models is another way to understand the molecular basis of type 2 diabetes mellitus, and the Goto-Kakizaki (GK) rat model is one of the best animal models for studying genetic susceptibility to type 2 diabetes mellitus. This rat manifests the main features of the metabolic, hormonal, and vascular disorders described in type 2 diabetes mellitus.83 It also exhibits a basal hyperinsulinaemia and impaired insulin response to glucose. One disadvantage of this model is the lack of obesity seen in these animals. Unlike the GK rats, the Otsuka Long-Evans Tokushima fatty (OLETF) rat is an animal model for type 2 diabetes mellitus, characterised by abdominal obesity, insulin resistance, hypertension, and dyslipidaemia. The OLETF rats develop the disorder with age, members of the same progeny are not all diabetic,84 and the rats also develop mild obesity.85-86 There has not so far been a good animal model available for type 2 diabetes mellitus, as the disease is much more complex and heterogeneous than can be found in inbred animal models. Complementary approaches in different animal strains may lead to the identification of candidate genes for type 2 diabetes mellitus and help to direct the search for candidate genes in humans.

Jonathan van Tilburg receives financial support from the Dutch Diabetes Research Foundation (DFN grant 97/251 to CW and TWvH). We thank Jackie Senior for improving the English.

Almind K, Bjorbaek C, Vestergaard H, Hansen T, Echwald
Committee on Diabetic Twins, Japan Diabetes Society.
King H, Zimmet P, Raper LR, Balkau B. Risk factors for
Fujimoto WY, Leonetti DL, Bergstrom RW, Kinyoun JL,
Gerich JE. The genetic basis of type 2 diabetes mellitus:
Ott J, Lucile P. Complex traits on the map. Recent Results Cancer Res 1998;154:288-91.
Genetics of type 2 diabetes mellitus


