The Breda Study: Search for genetic factors involved in type 2 diabetes mellitus in a defined Dutch population

De Breda Studie:
Zoeken naar genetische factoren betrokken bij
type 2 diabetes mellitus
in een gekarakteriseerde Nederlandse populatie
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de Rector Magnificus, Prof. Dr. W.H. Gispen ingevolge het besluit van het College voor Promoties in het openbaar te verdedigen op dinsdag 10 december 2002 des middags te 14:30 uur

door

Jonathan Hendrik Otto van Tilburg

geboren op 19 maart 1970, te Schiedam

Promotor: Prof. P. L. Pearson[1]

Co-promotors: Dr. C. Wijmenga[1]
Dr. T. W. van Haeften[2]

[1] Department of Medical Genetics
Univerisity Medical Center Utrecht
The Netherlands

[2] Department of Internal Medicine
University Medcial Center Utrecht
The Netherlands

ISBN 90-393-3227-4

The studies presented in this thesis were carried out in the Department of Medical Genetics of the University Medical Center (UMC) Utrecht.
The investigations were supported by grant DFN 97.114 of the Dutch Diabetes Foundation (DFN).

Assessment committee:	Prof. C. J. M. Lips
	Prof. J. P. H. Burbach
	Prof. J. J. M. Marx
	Prof. B. A. Oostra
	Prof. M. I. McCarthy

Paranymphs:	S. Postma
	E. Strengman

Publication of this thesis was financially supported by:
Diabetes Fonds Nederland - Faculty of Medicine, University Utrecht -
Department of Medical Genetics, University Medical Center Utrecht -
PEVATI BV.

Printed by PrintPartners Ipskamp BV.

The wise are like the universe
they treat the myriad things impartially
as straw dogs

Tao is like a bellows
empty yet never exhausted
used yet never used up
moving yet always yielding

Too many words
better to hold fast to the center

the taoist philosopher Lao Tsu

Contents

Scope of this thesis

The main aim of the Breda study described in this thesis was to identify genetic factors involved in type 2 diabetes mellitus in a defined Dutch population.

A representative Dutch population was obtained by collecting patient material from the town of Breda and surrounding areas. This material was used to identify genetic factors involved in type 2 diabetes mellitus by means of a genome-wide scan. The results obtained from the initial genome-wide scan revealed that the Breda study cohort consisted of two different groups of type 2 diabetes mellitus patients. Re-analysis of the genome-wide scan within the two different groups was performed in two different ways. We searched for factors influencing body mass index, as a measure of obesity in type 2 diabetes mellitus patients, and for factors involved in obesity "driven" type 2 diabetes mellitus. Furthermore, two possible candidate genes were investigated for association between genetic variants in these genes and type 2 diabetes mellitus.

Chapter 1

Introduction

Defining the genetic contribution of type 2 diabetes mellitus

Jonathan van Tilburg[1], Timon W. van Haeften[2], Peter Pearson[1], Cisca Wijmenga[1]

[1]Department of Medical Genetics, [2]Department of Internal Medicine, University Medical Centre Utrecht, the Netherlands

J. Med. Genet 2001; **38**:569-578 reprinted with permission from the BMJ Publishing Group

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 world-wide but its incidence is increasing rapidly due to 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.

Introduction

Diabetes mellitus (DM) affects over 150 million people world-wide, with a prevalence that varies markedly from population to population (Zimmet 1992). Estimates predict that almost 300 million people will suffer from DM by 2025 (see figure 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 hyperinsulinemia, increased fasting proinsulin, and decreased HDL-cholesterol (DeFronzo and Ferrannini 1991). 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 for 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

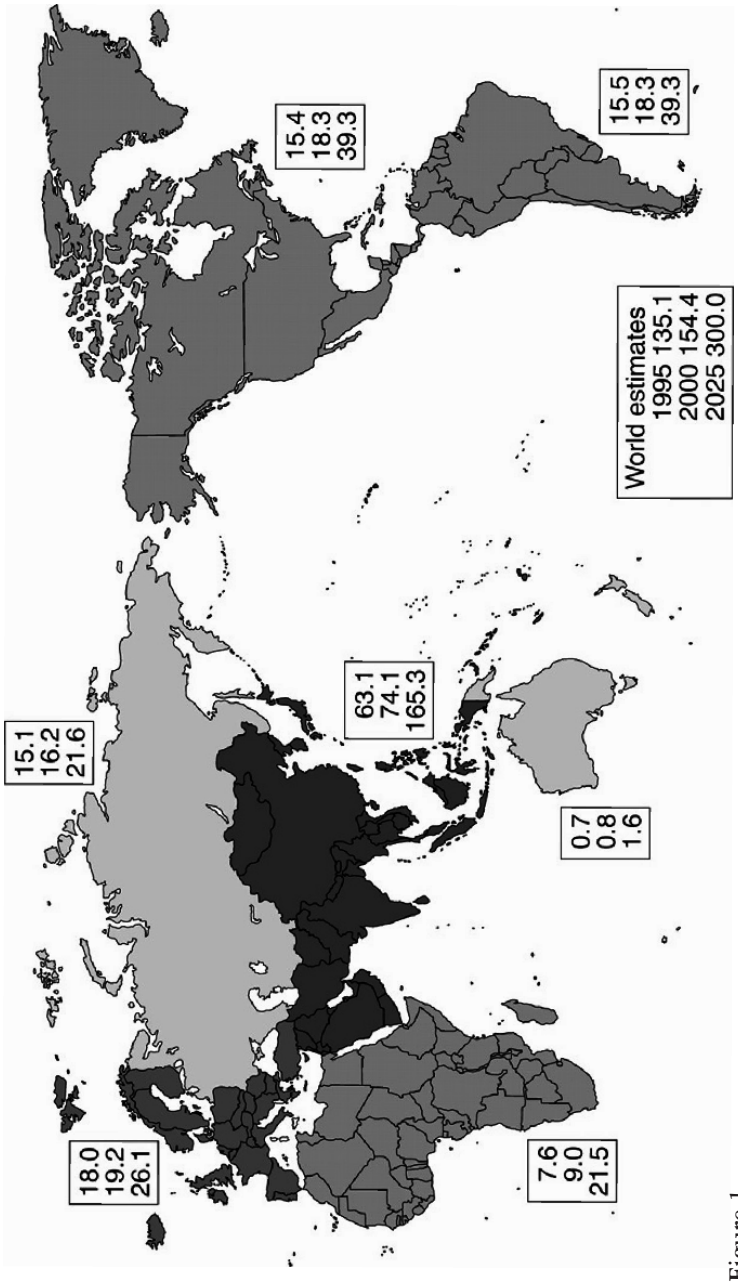


Figure 1.
Regional estimates of people with diaberes mellitus (in millions) for 1995, 2000 and 2025. Adapted from http://www.who.int/ncd/dia/databases0.html.

both characterised by elevated plasma glucose levels. Normal glucose homeostasis depends on the balance between glucose production by the liver and kidney, and glucose uptake by the brain, kidney, muscle and adipose tissue. Insulin, the predominant anabolic hormone involved, increases the uptake of glucose from the blood, enhances its conversion to glycogen and triglyceride and also increases glucose oxidation. Plasma glucose levels are normally kept within a small range (4 to 6 mmol/l) by multiple mechanisms. After a meal, a small increase in plasma glucose will lead to an increased insulin secretion by the pancreatic β-cells (figure 2).

This is associated with a decrease in glucose production by the liver, and enhanced glucose uptake in muscle and adipose tissue. These actions result from a combination of short-term rapid effects and longer-term slow effects, which involve changes in gene transcription and in the rate of translation of enzymes involved in glycogen synthesis, the glycolytic pathway and lipid metabolism (Heesom *et al.* 1997). The effect on gene expression can be either positive or negative, depending on the physiological role of the gene product.

There are a number of glucose counter-regulatory hormones, such as glucagon, cortisol, epinephrine and nor-epinephrine, which elevate 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 grams glucose in water). Two-hour values over 11.1 mmol/l (= 200 mg/dl) are still used as diagnostic for diabetes (1979 National Diabetes Data Group). This value was originally chosen when prospective studies indicated that subjects with a 2-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 2-hour post-glucose load plasma glucose levels of >11.1 mmol/l (1997 Report of Expert Committee). 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 auto-immune destruction of the pancreatic b-cell (Taylor 1997).

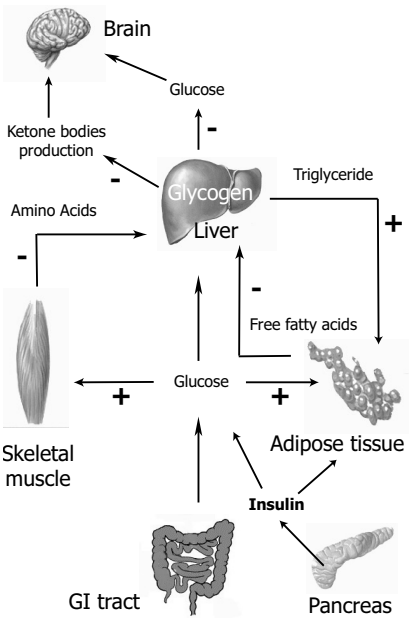


Figure 2. Insulin action after a meal.
Selected actions of insulin are indicated with + (upregulation) or – (down regulation). Insulin activates transport of glucose to muscle and adipose tissue, and also promotes synthesis of glycogen and triglycerides by the liver. Increased insulin levels inhibit glucose production by the liver, lypolysis 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 (e.g. during fasting).

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 β-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 dimin-

ished. Although many type 2 diabetes mellitus patients have a basal hyperinsulinemia, elevations 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 due to multiple mechanisms, including elevations in plasma free fatty acids (FFA) and tumour necrosis factor-alpha (TNFα) released from "full" adipocytes (Hotamisligil *et al.* 1995; Uysal *et al.* 1997). Furthermore, lack of physical exercise is also associated with diabetes mellitus, which led to the finding that exercise enhances the insulin action, presumably via upregulation of glucose-transporters in muscle (DeFronzo 1997).

Apart from the short-term complications such as thirst, malaise, tiredness, and ketoacidosis, 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 microvascular complications include retinopathy, neuropathy and nephropathy, 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 socio-economic cost and put a heavy burden on public health services (WHO 1999).

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 individuals (Valsania and Micossi 1994) 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 (Gerich 1998).

The following demographic observations have revealed 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 on mainland China, to extremely high levels of 50% in Pima Indians (Northern 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 the Japanese who migrated to Hawaii (Fujimoto *et al.* 1991; Fujimoto *et al.* 1987) or by the high prevalence (13.1%) among the Chinese living on the island of Mauritius compared with the prevalence among the Chinese living on mainland China (1.6%) (Fujimoto 2000). In general, there is a trend of increasing prevalence of diabetes mellitus with migration from rural to urban societies (King *et al.* 1984) 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 (Valsania and Micossi 1994). Several studies have shown higher concordance rates in monozygotic (MZ) twins than in dizygotic (DZ) twins (Medici *et al.* 1999) 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% (Kaprio *et al.* 1992). In a Japanese study these figures were 83% for MZ twins and 40% for DZ twins (1988 Japanese Diabetes Society). Such figures show the difference of environmental influences within populations (i.e. the difference between MZ and DZ twins). The large variation in concordance rates between populations may be due to bias or a different selection from the populations studied, but it may also indicate differences in genetic susceptibility between these populations (Hamman 1992; MacGregor *et al.* 2000).

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 λ_r, (relative risk of a relative). For type 1 diabetes mellitus the $\lambda_r = 20$ whereas the λ_r for type 2 diabetes mellitus = 3.5. This relative risk also increases with the number of affected relatives (Kahn *et al.* 1996; Rich 1990). 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 (Rich 1990).

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 reveal 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 (DeFronzo 1997). 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 IRS-1 (Almind *et al.* 1993; Almind *et al.* 1996; Porzio *et al.* 1999) the glucagon receptor (Hager *et al.* 1995; Hansen *et al.* 1996; Lok *et al.* 1994) the sulfonylurea receptor (SUR) (t Hart *et al.* 1999) the peroxisome proliferator-activated receptor-γ (PPARγ) (Altshuler *et al.* 2000; Hegele *et al.* 2000) and the MAPKBIP1 (Waeber *et al.* 2000) 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 (So *et al.* 2000; Velho and Froguel 1997).

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 differ-

ences 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 due 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 (Slatkin 1991). 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 individuals. 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 (Pritchard and Rosenberg 1999).

(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 multi-generational 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 individuals 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 sev-

eral genes may result in identical phenotypes, or a chromosomal region may co-segregate 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 (Lander and Schork 1994).

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 ASPs in genome-wide scans generally requires large numbers of ASPs to obtain sufficient power for detecting linkage for a given value of λ_s (relative risk for a sibling) (Risch 1990; Risch and Merikangas 1996). 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 λ_r (relative risk for a relative) is greater than 1.7, given a LOD of 1.0 (Risch 1990; Weeks and Lathrop 1995). 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 (Holmans and Craddock 1997; Kruglyak and Daly 1998; Weeks and Lathrop 1995). A three-stage strategy, with increasing thresholds at each stage, is the most powerful approach to adopt in a genome scan (Brown *et al.* 1994; Weeks and Lathrop 1995). 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 (Holmans 1998; Holmans and Craddock 1997).

An efficient study design is an important aspect of any genome-wide scan.

Different types of cohorts, consisting of nuclear families, multi-generational 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 (Ehm *et al.* 2000; Ghosh *et al.* 1999; Ghosh *et al.* 2000; Hanis *et al.* 1996; Hegele *et al.* 1999; Ji *et al.* 1997; Watanabe *et al.* 2000; Zouali *et al.* 1997) or occasionally, multi-generational families (Duggirala *et al.* 1999; Elbein *et al.* 1999; Hanson *et al.* 1998; Ji *et al.* 1997; Mahtani *et al.* 1996). Both types of genome scans (using ASPs or multi-generational 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* (Hanis *et al.* 1996). In a sample from Botnia, Western Finland, a small number of selected pedigrees with the lowest quartile for mean 30-min 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* (Mahtani *et al.* 1996). More recently, several studies have shown significant evidence for linkage to chromosome 20 (Ghosh *et al.* 1999; Ghosh *et al.* 2000; Ji *et al.* 1997; Zouali *et al.* 1997) and a recent genome scan in Pima Indians revealed 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 (Hanson *et al.* 1998). In 42 multi-generational families with Northern European ancestry from Utah, significant linkage was found under a model of recessive inheritance on chromosome 1q21-23 (Elbein *et al.* 1999) 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 (Hegele *et al.* 1999). 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 (Duggirala *et al.* 1999).

Box 1. Affected sib pair analysis (ASP)

Currently, the ASP approach is the most commonly used non-parametric or model-free mapping approach (Kruglyak and Lander 1995; Lander and Schork 1994; Weeks and Lathrop 1995) and it only requires pairs of affected siblings (Holmans and Craddock 1997). The ASP approach is well suited to the analysis of type 2 diabetes mellitus because only one or two generations in a family with this disease are normally available. The basis of the ASP analysis is that individuals concordant for a given genetic trait should show greater than expected concordance for marker alleles that are closely linked to the disease. The most frequently used measure of concordance of two siblings at a locus is the number of alleles they share *identical-by-descent* (IBD). If the marker is not linked to a disease susceptibility locus, then the probabilities of a sib pair sharing 0, 1, and 2 alleles IBD are 0.25, 0.50, and 0.25, respectively (Holmans 1998; Lander and Schork 1994; McCarthy *et al.* 1998; Velho and Froguel 1997; Weeks and Lathrop 1995). The mean sharing is 0.5. If the marker is linked to a disease locus, the probability of an affected sib pair sharing IBD alleles should be higher (i.e. >0.5) (figures 3 and 4).

Figure 3. Family with two affected sibs

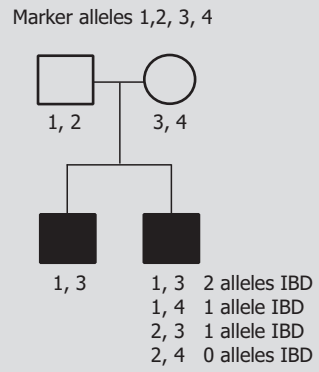
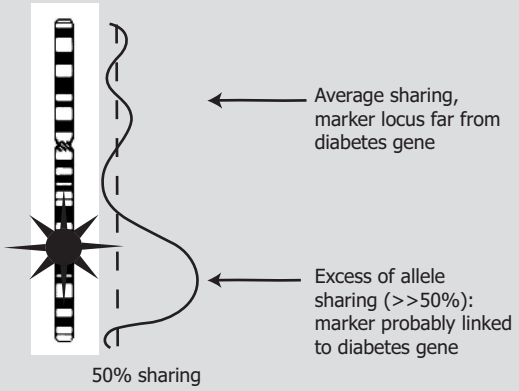


Figure 4. Overview of sharing



The most distinct approach for determining the number of alleles IBD is to count the number of pairs sharing 0, 1, and 2 alleles IBD and to compare these to the expected frequencies under the hypothesis of

no linkage (H0) using a chi-square test. Another possibility is to compare the average number of shared IBD alleles by the affected sibs (the mean test), or to count the number of pairs sharing two IBD alleles and compare this to the expected numbers. An excess of IBD allele sharing in each case is taken as evidence of linkage between the tested marker and the disease susceptibility locus (Haines 1998; Holmans 1998; Ott and Lucek 1998).
Since type 2 diabetes mellitus is a late-onset disease, the patient's parents are usually unavailable so that the IBD status of affected sib pairs cannot be determined. One way of circumventing the problem due to untyped parents is to consider the number of alleles shared'*identical-by-state* (IBS). Two individuals are said to share an IBS allele if they both have a copy of an identical allele, regardless of from whom it was inherited. One problem with this kind of analysis is the reduced power, its reliance on the allele frequency, and the higher percentage of false-positive findings. However, the alleles can be reconstructed using additional sibs approaching IBD in more than 80% of the cases (Bishop and Williamson 1990; Haines 1998) (Sandkuijl, unpublished data).

Most recently, a genome-wide scan in four American populations has revealed suggestive linkage to type 2 diabetes mellitus or impaired glucose homeostasis on chromosome s 5, 12 and X in whites, on chromosome 3 in Mexican Americans, and chromosome 10 in Afro-Americans (Ehm *et al.* 2000). In an eastern and south-eastern Chinese Han population, two loci in a region on chromosome 9 showed suggestive evidence for linkage to type 2 diabetes (Luo *et al.* 2001).

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 (Frayling *et al.* 2000).

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, this may also be true for chromosome 2 in Mexican Americans and for chromosomes 12 and 20 in Caucasians (see table 1). The genomic regions described so far, which extend over 20 cM in many cases, now require fine mapping to pinpoint the region

Table 1. Linkage results of different genome wide scans in type 2 diabetes mellitus.

Ethnic group	Trait
Pima Indian	Diabetes mellitus before age 25 years
	Age-adjusted diabetes mellitus
	Age-adjusted diabetes mellitus
	Age-adjusted diabetes mellitus
	Diabetes mellitus
Mexican Americans	Diabetes mellitus
	Diabetes mellitus
	Diabetes mellitus age at onset
	Diabetes mellitus age at onset
	Diabetes mellitus age at onset
	Diabetes mellitus
	Diabetes mellitus
	Diabetes mellitus
	Diabetes mellitus
Caucasian (Finn)	Diabetes mellitus (stratified on 30 min. insulin)
Caucasian (North American)	Diabetes mellitus
Caucasian (North American)	Diabetes mellitus
Caucasian (French)	Diabetes mellitus
Caucasian (Finn)	Diabetes mellitus
Caucasian (Utah)	Diabetes mellitus (recessive model)
Oji-Cree (Canadian)	Diabetes mellitus
Han (China)	Diabetes mellitus

Sample (cohort)	Locus marker	LOD
264 nuclear families	D1S198	4.1
(Hanson *et al.* 1998)	D6S1009-D6S1003	1.39
	D9S299-D9S2026	1.22
	D11S4464-D11S912	1.66
	D7S1799	1.8
330 ASPs(Hanis *et al.* 1996)	D2S125	4.03
- 53 nuclear families (Phase 1)	D3S2432	3.91
- 64 nuclear families (Phase 2)	D3S2432	<0.1
(Ehm *et al.* 2000)	D3S1566-GATA128C02	2.51
27 extended families	D9S288-D9S925	2.06
(Duggirala *et al.* 1999)	D10S587-D10S1223	3.75
	D3S1566-GATA128C02	2.67
	D4S1615-D4S175	1.99
	D9S288-D9S925	2.38
	D10S587-D10S1223	2.88
26 families	D12S1349	3.3
77 nuclear families (Phase 1)	D5S1404	2.8
(Ehm *et al.* 2000)	D12S853	2.81
	GATA172D05 (X chr.)	2.99
14 extended families	D20S197	3.3
(Ji *et al.* 1997).		
55 ASPs (Zouali *et al.* 1997)	ADA (chr. 20)	2.84
	PCK1 (chr. 20)	2.04
716 ASPs (Ghosh *et al.* 1999;	D11S937-D11S901	1.75
Ghosh *et al.* 2000)	D20S849-D20S905	1.99
	D20S909-D20S107	2.04
	D20S886-D20S197	2.15
42 extended families (Elbein *et al.* 1999)	CRP-APOA2 (chr. 1)	4.3
49 ASPs (Hegele *et al.* 1999)	D6S1056	4.24
	D8S264	2.91
	D16S2616	4.20
	D22S683	2.48
168 ASPs (Luo *et al.* 2001)	D9S171	3.29
	D9S161	2.22
	D9S175	2.94

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 individuals 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 scan 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 (Hanis *et al.* 1996). This finding suggests a novel pathway that may contribute to the development of type 2 diabetes mellitus (Horikawa *et al.* 2000). Using of single nucleotide polymorphisms (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 (Horikawa *et al.* 2000).

If we consider there may be approximately 30,000 genes in the human genome (McPherson *et al.* 2001; Venter *et al.* 2001) 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 (Permutt and Hattersley 2000). 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 pa-

tients having "only" one of these diseases, preferably in the same ethnic population"(Parker *et al.* 2001; Perusse *et al.* 2001).

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 new gene. New and improved technologies, such as microarrays that can type thousands of SNPs in a single assay, will also be of great importance in finding genetic variation in these new genes. Combining these genetic variations with new developments in the fields of bioinformatics, genomics and proteomics will lead to a greater understanding of the pathogenesis of type 2 diabetes mellitus, and may provide new information for diagnostics, treatment and, eventually, prevention of the disease. This genetic information may also form the basis for the development of new drug therapies such as individually specific or targeted pharmacotherapy.

BOX 2. Alternative approaches for 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 revealed at least three AD genes'(Levy-Lahad *et al.* 1998; Roses 1996; Schellenberg 1995). 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 account for only 2-5% of the 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 (see table 2). Another rare and early onset form of diabetes is the maternally inherited diabetes and deafness (MIDD), in which mutations are found in the mitochondrion. 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 (Alcolado and Alcolado 1991). 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 more common type 2 diabetes mellitus.

Table 2. Genes involved in MODY.

	Location on genome	Gene	References
MODY1	20q12-q13.1	hepatocyte nuclear factor-4α(HNF-4-alpha)	(Yamagata *et al.* 1996)
MODY2	7p15-p13	glucokinase (GCK)	(Froguel *et al.* 1992; Froguel and Velho 1993; Matschinsky 1990; Velho *et al.* 1992; Vionnet *et al.* 1992)
MODY3	12q24.2	hepatocyte nuclear factor-1α (TCF1)	(Yamagata *et al.* 1996)
MODY4	13q21.1	insulin promoter factor-1 (IPF1)	(Leonard *et al.* 1993; Miller *et al.* 1994; Ohlsson *et al.* 1993; Stoffel *et al.* 1995; Stoffers *et al.* 1997; Stoffers *et al.* 1998)
MODY5	17cent.-q21.3	hepatocyte nuclear factor-1-β (TCF2)	(Abbott *et al.* 1990; Horikawa *et al.* 1997)
MODY6	2q	Neurogenic differentiation 1 (NEUROD1)	(Malecki *et al.* 1999)

(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 individuals for study (Sheffield *et al.* 1998). 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 (Bull *et al.* 1998; De Koning *et al.* 1995; Houwen *et al.* 1994). 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 (Permutt *et al.* 2001). 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 (Peltonen *et al.* 1995).

(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 (Kim *et al.* 1998). Today there are several mice and rat models available for studying both type 2 diabetes mellitus and obesity. By crossing the monogenic mouse (the ob/ob and the db/db mice) models with other strains, it might be possible to reveal modifier genes (Ktorza *et al.* 1997). The use of polygenic models is another way towards understanding the molecular basis of type 2 diabetes mellitus, and the Goto-Kakisaki (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 (Hussain 1997). It also exhibits a basal hyperinsulinemia 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 dyslipidemia. The OLETF rats develop the disorder with age, individuals of the same progeny are not all diabetic (Ktorza *et al.* 1997) and the rats also develop mild obesity (Kanemoto *et al.* 1998; Nara *et al.* 1997; Wei *et al.* 1999). There has not so far been a good animal model available for type 2 diabetes mellitus, 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.

Acknowledgements

Jonathan van Tilburg receives financial support from the Dutch Diabetes Research Foundation (DFN grant 97.114 to CW and TWvH). We thank Jackie Senior for improving the English text.

References

(1979) Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. National Diabetes Data Group. Diabetes 28:1039-57.

(1988) Diabetes mellitus in twins: a cooperative study in Japan. Committee on Diabetic Twins, Japan Diabetes Society. Diabetes Res Clin Pract 5:271-80.

(1997) Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care 20:1183-97.

Abbott C, Piaggio G, Ammendola R, Solomon E, Povey S, Gounari F, De Simone V, et al. (1990) Mapping of the gene TCF2 coding for the transcription factor LFB3 to human chromosome 17 by polymerase chain reaction. Genomics 8:165-7.

Alcolado JC, Alcolado R (1991) Importance of maternal history of non-insulin dependent diabetic patients. Bmj 302:1178-80.

Almind K, Bjorbaek C, Vestergaard H, Hansen T, Echwald S, Pedersen O (1993) Aminoacid polymorphisms of insulin receptor substrate-1 in non-insulin- dependent diabetes mellitus. Lancet 342:828-32.

Almind K, Inoue G, Pedersen O, Kahn CR (1996) A common amino acid polymorphism in insulin receptor substrate-1 causes impaired insulin signaling. Evidence from transfection studies. J Clin Invest 97:2569-75.

Altshuler D, Hirschhorn JN, Klannemark M, Lindgren CM, Vohl MC, Nemesh J, Lane CR, et al. (2000) The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. Nat Genet 26:76-80.

Bishop DT, Williamson JA (1990) The power of identity-by-state methods for linkage analysis. Am J Hum Genet 46:254-65.

Brown DL, Gorin MB, Weeks DE (1994) Efficient strategies for genomic searching using the affected-pedigree- member method of linkage analysis. Am J Hum Genet 54:544-52.

Bull LN, van Eijk MJ, Pawlikowska L, DeYoung JA, Juijn JA, Liao M, Klomp LW, et al. (1998) A gene encoding a P-type ATPase mutated in two forms of hereditary cholestasis. Nat Genet 18:219-24.

De Koning TJ, Sandkuijl LA, De Schryver JE, Hennekam EA, Beemer FA, Houwen RH (1995) Autosomal-recessive inheritance of benign recurrent intrahepatic cholestasis. Am J Med Genet 57:479-82.

DeFronzo RA (1997) Pathogenesis of type 2 diabetes: metabolic and molecular implications for identifying diabetes genes. Diabetes Reviews 5:177-269.

DeFronzo RA, Ferrannini E (1991) Insulin resistance. A multifaceted syndrome responsible for NIDDM, obesity, hypertension, dyslipidemia, and atherosclerotic cardiovascular disease. Diabetes Care 14:173-94.

Duggirala R, Blangero J, Almasy L, Dyer TD, Williams KL, Leach RJ, O'Connell P, et al. (1999) Linkage of type 2 diabetes mellitus and of age at onset to a genetic location on chromosome 10q in Mexican Americans. Am J Hum Genet 64:1127-40.

Ehm MG, Karnoub MC, Sakul H, Gottschalk K, Holt DC, Weber JL, Vaske D, et al. (2000) Genomewide Search for Type 2 Diabetes Susceptibility Genes in Four American Populations. Am J Hum Genet 66:1871-1881.

Elbein SC, Hoffman MD, Teng K, Leppert MF, Hasstedt SJ (1999) A genome-wide search for type 2 diabetes susceptibility genes in Utah Caucasians. Diabetes 48:1175-82.

Frayling TM, McCarthy MI, Walker M, Levy JC, O'Rahilly S, Hitman GA, Rao PV, et al. (2000) No evidence for linkage at candidate type 2 diabetes susceptibility loci on chromosomes 12 and 20 in United Kingdom Caucasians. J Clin Endocrinol Metab 85:853-7.

Froguel P, Vaxillaire M, Sun F, Velho G, Zouali H, Butel MO, Lesage S, et al. (1992) Close linkage of glucokinase locus on chromosome 7p to early-onset non- insulin-dependent diabetes mellitus. Nature 356:162-4.

Froguel P, Velho G (1993) Non-sense mutation of glucokinase gene. Lancet 341:385.

Fujimoto WY (2000) The importance of insulin resistance in the pathogenesis of type 2 diabetes mellitus. Am J Med 108 Suppl 6a:9S-14S.

Fujimoto WY, Leonetti DL, Bergstrom RW, Kinyoun JL, Stolov WC, Wahl PW (1991) Glucose intolerance and diabetic complications among Japanese-American women. Diabetes Res Clin Pract 13:119-29.

Fujimoto WY, Leonetti DL, Kinyoun JL, Newell-Morris L, Shuman WP, Stolov WC, Wahl PW (1987) Prevalence of diabetes mellitus and impaired glucose tolerance among second-generation Japanese-American men. Diabetes 36:721-9.

Gerich JE (1998) The genetic basis of type 2 diabetes mellitus: impaired insulin secretion versus impaired insulin sensitivity. Endocr Rev 19:491-503.

Ghosh S, Watanabe RM, Hauser ER, Valle T, Magnuson VL, Erdos MR, Langefeld CD, et al. (1999) Type 2 diabetes: evidence for linkage on chromosome 20 in 716 Finnish affected sib pairs. Proc Natl Acad Sci U S A 96:2198-203.

Ghosh S, Watanabe RM, Valle TT, Hauser ER, Magnuson VL, Langefeld CD, Ally DS, et al. (2000) The Finland-United States Investigation of Non-Insulin-Dependent Diabetes Mellitus Genetics (FUSION) Study. I. An Autosomal Genome Scan for Genes That Predispose to Type 2 Diabetes. Am J Hum Genet 67:1174-1185.

Hager J, Hansen L, Vaisse C, Vionnet N, Philippi A, Poller W, Velho G, et al. (1995) A missense mutation in the glucagon receptor gene is associated with non-insulin-dependent diabetes mellitus. Nat Genet 9:299-304.

Haines JL (1998) Sib Pair Analysis. In: Pericak-Vance JLHMA (ed) Approaches to Gene Mapping in Complex Human Diseases. John Wiley & Sons, Inc. Publications, pp 273-303.

Hamman RF (1992) Genetic and environmental determinants of non-insulin-dependent diabetes mellitus (NIDDM). Diabetes Metab Rev 8:287-338.

Hanis CL, Boerwinkle E, Chakraborty R, Ellsworth DL, Concannon P, Stirling B, Morrison VA, et al. (1996) A genome-wide search for human non-insulin-dependent (type 2) diabetes genes reveals a major susceptibility locus on chromosome 2. Nat Genet 13:161-6.

Hansen LH, Abrahamsen N, Hager J, Jelinek L, Kindsvogel W, Froguel P, Nishimura E (1996) The Gly40Ser mutation in the human glucagon receptor gene associated with NIDDM results in a receptor with reduced sensitivity to glucagon. Diabetes 45:725-30.

Hanson RL, Ehm MG, Pettitt DJ, Prochazka M, Thompson DB, Timberlake D, Foroud T, et al. (1998) An autosomal genomic scan for loci linked to type II diabetes mellitus and body-mass index in Pima Indians. Am J Hum Genet 63:1130-8.

Heesom KJ, Harbeck M, Kahn CR, Denton RM (1997) Insulin action on metabolism. Diabetologia 40 Suppl 3:B3-9.

Hegele RA, Cao H, Harris SB, Zinman B, Hanley AJ, Anderson CM (2000) Peroxisome proliferator-activated receptor-gamma2 P12A and type 2 diabetes in Canadian Oji-Cree. J Clin Endocrinol Metab 85:2014-9.

Hegele RA, Sun F, Harris SB, Anderson C, Hanley AJ, Zinman B (1999) Genome-wide scanning for type 2 diabetes susceptibility in Canadian Oji- Cree, using 190 microsatellite markers. J Hum Genet 44:10-4.

Holmans P (1998) Affected sib-pair methods for detecting linkage to dichotomous traits: review of the methodology. Hum Biol 70:1025-40.

Holmans P, Craddock N (1997) Efficient strategies for genome scanning using maximum-likelihood affected-sib-pair analysis. Am J Hum Genet 60:657-66.

Horikawa Y, Iwasaki N, Hara M, Furuta H, Hinokio Y, Cockburn BN, Lindner T, et al. (1997) Mutation in hepatocyte nuclear factor-1 beta gene (TCF2) associated with MODY. Nat Genet 17:384-5.

Horikawa Y, Oda N, Cox NJ, Li X, Orho-Melander M, Hara M, Hinokio Y, et al. (2000) Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. Nat Genet 26:163-75.

Hotamisligil GS, Arner P, Caro JF, Atkinson RL, Spiegelman BM (1995) Increased adipose tissue expression of tumor necrosis factor-alpha in human obesity and insulin resistance. J Clin Invest 95:2409-15.

Houwen RH, Baharloo S, Blankenship K, Raeymaekers P, Juyn J, Sandkuijl LA, Freimer NB (1994) Genome screening by searching for shared segments: mapping a gene for benign recurrent intrahepatic cholestasis. Nat Genet 8:380-6.

Hussain MA (1997) Polygenic models of non-insulin-dependent diabetes mellitus. Eur J Endocrinol 137:453-4.

Ji L, Malecki M, Warram JH, Yang Y, Rich SS, Krolewski AS (1997) New susceptibility locus for NIDDM is localized to human chromosome 20q. Diabetes 46:876-81.

Kahn CR, Vicent D, Doria A (1996) Genetics of non-insulin-dependent (type-II) diabetes mellitus. Annu Rev Med 47:509-31.

Kanemoto N, Hishigaki H, Miyakita A, Oga K, Okuno S, Tsuji A, Takagi T, et al. (1998) Genetic dissection of "OLETF", a rat model for non-insulin-dependent diabetes mellitus. Mamm Genome 9:419-25.

Kaprio J, Tuomilehto J, Koskenvuo M, Romanov K, Reunanen A, Eriksson J, Stengard J, et al. (1992) Concordance for type 1 (insulin-dependent) and type 2 (non-insulin- dependent) diabetes mellitus in a population-based cohort of twins in Finland. Diabetologia 35:1060-7.

Kim JH, Nishina PM, Naggert JK (1998) Genetic models for non insulin dependent diabetes mellitus in rodents. J Basic Clin Physiol Pharmacol 9:325-45.

King H, Zimmet P, Raper LR, Balkau B (1984) Risk factors for diabetes in three Pacific populations. Am J Epidemiol 119:396-409.

Kruglyak L, Daly MJ (1998) Linkage thresholds for two-stage genome scans. Am J Hum Genet 62:994-7.

Kruglyak L, Lander ES (1995) Complete multipoint sib-pair analysis of qualitative and quantitative traits. Am J Hum Genet 57:439-54.

Ktorza A, Bernard C, Parent V, Penicaud L, Froguel P, Lathrop M, Gauguier D (1997) Are animal models of diabetes relevant to the study of the genetics of non-insulin-dependent diabetes in humans? Diabetes Metab 23 Suppl 2:38-46.

Lander ES, Schork NJ (1994) Genetic dissection of complex traits. Science 265:2037-48.

Leonard J, Peers B, Johnson T, Ferreri K, Lee S, Montminy MR (1993) Characterization of somatostatin transactivating factor-1, a novel homeobox factor that stimulates somatostatin expression in pancreatic islet cells. Mol Endocrinol 7:1275-83.

Levy-Lahad E, Tsuang D, Bird TD (1998) Recent advances in the genetics of Alzheimer's disease. J Geriatr Psychiatry Neurol 11:42-54.

Lok S, Kuijper JL, Jelinek LJ, Kramer JM, Whitmore TE, Sprecher CA, Mathewes S, et al. (1994) The human glucagon receptor encoding gene: structure, cDNA sequence and chromosomal localization. Gene 140:203-9.

Luo TH, Zhao Y, Li G, Yuan WT, Zhao JJ, Chen JL, Huang W, et al. (2001) A genome-wide search for type II diabetes susceptibility genes in Chinese Hans.

Diabetologia 44:501-6.
MacGregor AJ, Snieder H, Schork NJ, Spector TD (2000) Twins. Novel uses to study complex traits and genetic diseases. Trends Genet 16:131-4.
Mahtani MM, Widen E, Lehto M, Thomas J, McCarthy M, Brayer J, Bryant B, et al. (1996) Mapping of a gene for type 2 diabetes associated with an insulin secretion defect by a genome scan in Finnish families. Nat Genet 14:90-4.
Malecki MT, Jhala US, Antonellis A, Fields L, Doria A, Orban T, Saad M, et al. (1999) Mutations in NEUROD1 are associated with the development of type 2 diabetes mellitus. Nat Genet 23:323-8.
Matschinsky FM (1990) Glucokinase as glucose sensor and metabolic signal generator in pancreatic beta-cells and hepatocytes. Diabetes 39:647-52.
McCarthy MI, Kruglyak L, Lander ES (1998) Sib-pair collection strategies for complex diseases. Genet Epidemiol 15:317-40.
McPherson JD, Marra M, Hillier L, Waterston RH, Chinwalla A, Wallis J, Sekhon M, et al. (2001) A physical map of the human genome. The International Human Genome Mapping Consortium. Nature 409:934-41.
Medici F, Hawa M, Ianari A, Pyke DA, Leslie RD (1999) Concordance rate for type II diabetes mellitus in monozygotic twins: actuarial analysis. Diabetologia 42:146-50.
Miller CP, McGehee RE, Jr., Habener JF (1994) IDX-1: a new homeodomain transcription factor expressed in rat pancreatic islets and duodenum that transactivates the somatostatin gene. Embo J 13:1145-56.
Nara Y, Gao M, Ikeda K, Sato T, Sawamura M, Kawano K, Yamori Y (1997) Genetic analysis of non-insulin-dependent diabetes mellitus in the Otsuka Long-Evans Tokushima Fatty rat. Biochem Biophys Res Commun 241:200-4.
Ohlsson H, Karlsson K, Edlund T (1993) IPF1, a homeodomain-containing transactivator of the insulin gene. Embo J 12:4251-9.
Ott J, Lucek P (1998) Complex traits on the map. Recent Results Cancer Res 154:285-91.
Parker A, Meyer J, Lewitzky S, Rennich JS, Chan G, Thomas JD, Orho-Melander M, et al. (2001) A gene conferring susceptibility to type 2 diabetes in conjunction with obesity is located on chromosome 18p11. Diabetes 50:675-80.
Peltonen L, Pekkarinen P, Aaltonen J (1995) Messages from an isolate: lessons from the Finnish gene pool. Biol Chem Hoppe Seyler 376:697-704.
Permutt MA, Hattersley AT (2000) Searching for Type 2 Diabetes Genes in the Post-genome Era. Trends Endocrinol Metab 11:383-393.
Permutt MA, Wasson JC, Suarez BK, Lin J, Thomas J, Meyer J, Lewitzky S, et al. (2001) A genome scan for type 2 diabetes susceptibility loci in a genetically isolated population. Diabetes 50:681-5.
Perusse L, Rice T, Chagnon YC, Despres JP, Lemieux S, Roy S, Lacaille M, et al.

(2001) A genome-wide scan for abdominal fat assessed by computed tomography in the Quebec Family Study. Diabetes 50:614-21.

Porzio O, Federici M, Hribal ML, Lauro D, Accili D, Lauro R, Borboni P, et al. (1999) The Gly972—>Arg amino acid polymorphism in IRS-1 impairs insulin secretion in pancreatic beta cells. J Clin Invest 104:357-64.

Pritchard JK, Rosenberg NA (1999) Use of unlinked genetic markers to detect population stratification in association studies. Am J Hum Genet 65:220-8.

Rich SS (1990) Mapping genes in diabetes. Genetic epidemiological perspective. Diabetes 39:1315-9.

Risch N (1990) Linkage strategies for genetically complex traits. II. The power of affected relative pairs. Am J Hum Genet 46:229-41.

Risch N, Merikangas K (1996) The future of genetic studies of complex human diseases. Science 273:1516-7.

Roses AD (1996) Apolipoprotein E alleles as risk factors in Alzheimer's disease. Annu Rev Med 47:387-400.

Schellenberg GD (1995) Progress in Alzheimer's disease genetics. Curr Opin Neurol 8:262-7.

Sheffield VC, Stone EM, Carmi R (1998) Use of isolated inbred human populations for identification of disease genes. Trends Genet 14:391-6.

Slatkin M (1991) Inbreeding coefficients and coalescence times. Genet Res 58:167-75.

So WY, Ng MC, Lee SC, Sanke T, Lee HK, Chan JC (2000)

Genetics of type 2 diabetes mellitus. Hong Kong Med J 6:69-76.

Stoffel M, Stein R, Wright CV, Espinosa R, 3rd, Le Beau MM, Bell GI (1995)Localization of human homeodomain transcription factor insulin promoter factor 1 (IPF1) to chromosome band 13q12.1. Genomics 28:125-6.

Stoffers DA, Ferrer J, Clarke WL, Habener JF (1997) Early-onset type-II diabetes mellitus (MODY4) linked to IPF1. Nat Genet 17:138-9.

Stoffers DA, Stanojevic V, Habener JF (1998) Insulin promoter factor-1 gene mutation linked to early-onset type 2 diabetes mellitus directs expression of a dominant negative isoprotein. J Clin Invest 102:232-41.

t Hart LM, Stolk RP, Dekker JM, Nijpels G, Grobbee DE, Heine RJ, Maassen JA (1999) Prevalence of variants in candidate genes for type 2 diabetes mellitus in The Netherlands: the Rotterdam study and the Hoorn study. J Clin Endocrinol Metab 84:1002-6.

Taylor S (1997) Diabetes Mellitus The Metabolic and molecular-bases of inherited disease. Vol. 7th Edition, pp 843-896.

Uysal KT, Wiesbrock SM, Marino MW, Hotamisligil GS (1997) Protection from obesity-induced insulin resistance in mice lacking TNF- alpha function. Nature 389:610-4.

Valsania P, Micossi P (1994) Genetic epidemiology of non-insulin-dependent diabetes. Diabetes Metab Rev 10:385-405.

Velho G, Froguel P (1997) Genetic determinants of non-insulin-dependent diabetes mellitus: strategies and recent results. Diabetes Metab 23:7-17.

Velho G, Froguel P, Clement K, Pueyo ME, Rakotoambinina B, Zouali H, Passa P, et al. (1992) Primary pancreatic beta-cell secretory defect caused by mutations in glucokinase gene in kindreds of maturity onset diabetes of the young. Lancet 340:444-8.

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, et al. (2001) The Sequence of the Human Genome. Science 291:1304-1351.

Vionnet N, Stoffel M, Takeda J, Yasuda K, Bell GI, Zouali H, Lesage S, et al. (1992) Nonsense mutation in the glucokinase gene causes early-onset non- insulin-dependent diabetes mellitus. Nature 356:721-2.

Waeber G, Delplanque J, Bonny C, Mooser V, Steinmann M, Widmann C, Maillard A, et al. (2000) The gene MAPK8IP1, encoding islet-brain-1, is a candidate for type 2 diabetes. Nat Genet 24:291-5.

Watanabe RM, Ghosh S, Langefeld CD, Valle TT, Hauser ER, Magnuson VL, Mohlke KL, et al. (2000) The Finland-United States Investigation of Non-Insulin-Dependent Diabetes Mellitus Genetics (FUSION) Study. II. An Autosomal Genome Scan for Diabetes-Related Quantitative-Trait Loci. Am J Hum Genet 67:1186-1200.

Weeks DE, Lathrop GM (1995) Polygenic disease: methods for mapping complex disease traits. Trends Genet 11:513-9.

Wei S, Wei K, Moralejo DH, Ogino T, Koike G, Jacob HJ, Sugiura K, et al. (1999) Mapping and characterization of quantitative trait loci for non-insulin- dependent diabetes mellitus with an improved genetic map in the Otsuka Long-Evans Tokushima fatty rat. Mamm Genome 10:249-58.

WHO (1999) Definition, Diagnosis and Classification of Diabetes Mellitus and its complications. WHO, pp 42.

Yamagata K, Furuta H, Oda N, Kaisaki PJ, Menzel S, Cox NJ, Fajans SS, et al. (1996) Mutations in the hepatocyte nuclear factor-4alpha gene in maturity- onset diabetes of the young (MODY1). Nature 384:458-60.

Zimmet PZ (1992) Kelly West Lecture 1991. Challenges in diabetes epidemiology—from West to the rest. Diabetes Care 15:232-52.

Zouali H, Hani EH, Philippi A, Vionnet N, Beckmann JS, Demenais F, Froguel P (1997) A susceptibility locus for early-onset non-insulin dependent (type 2) diabetes mellitus maps to chromosome 20q, proximal to the phosphoenolpyruvate carboxykinase gene. Hum Mol Genet 6:1401-8.

Chapter 2

The Breda Study Cohort

The Dutch Breda Cohort of type 2 diabetes mellitus patients

JHO van Tilburg[1], E Strengman[1], H van Someren[1], T Rigters-Aris[1], PL Pearson[1], TW van Haeften[2], C Wijmenga[1]

[1]Department of Medical Genetics, and [2]Department of Internal Medicine, University Medical Center, Utrecht, the Netherlands.

Available online: http://humgen.med.uu.nl/research/diabetes/BredaCohort.html

Introduction

The prevalence of type 2 diabetes mellitus (T2D) varies markedly worldwide. In the Netherlands the prevalence is estimated to be at least 2.5% of the total population, and ranges from 2.6% (in the age group 55-59 years) up to 18% (in the age group 80-84 years) (The Hoorn Study (Heine *et al.* 1996; Nijpels 1998)).

Based on different epidemiological studies, the relative risk (i.e. ratio of the incidence of T2D among relatives of a proband with T2D to the incidence of T2D in the general population) for MZ twins of T2D probands (λ_{MZ} is predicted to be 10, whereas the relative risk for first-degree and second-degree relatives (λ_1 and λ_2) is predicted to be 3.5 and 1.5, respectively (Elbein 1997; Kaprio *et al.* 1992; Medici *et al.* 1999).

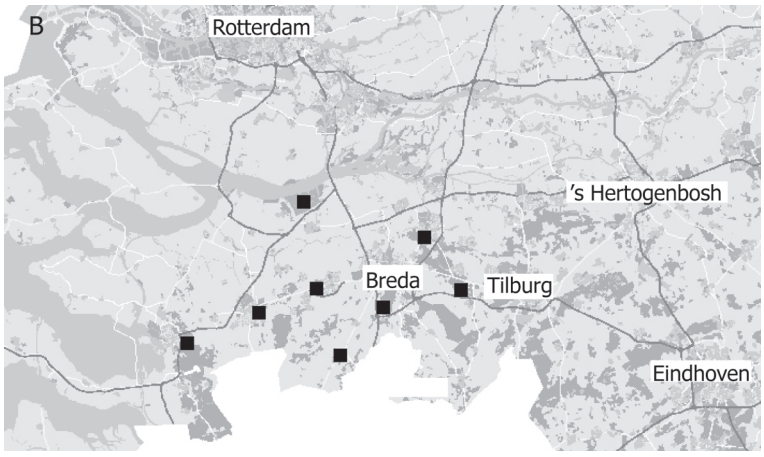
This clearly implies the involvement of a genetic component in T2D. Nevertheless, the expression of the disease does not follow Mendelian segregation. We know that the expression of T2D largely depends on environmental factors, such as a family history of diabetes, increased age, hypertension, lack of physical exercise, and obesity (DeFronzo and Ferrannini 1991). According to this multifactorial model, genetically predisposed subjects will not necessarily develop overt disease unless they are also exposed to one or more of these particular environmental factors (Valsania and Micossi 1994).

Aim of the study

The aim of the study was to perform a genome-wide scan in at least 250 Dutch Caucasian affected sibpairs with T2D to identify susceptibility loci. To date, no genome-wide scan has been performed in an outbred population of Dutch T2D patients and no sufficient Dutch cohort of affected sibpairs was available to perform such a genome-wide scan. To collect T2D sibpairs, we set up a collaboration with the Diabetes Service Breda and 80 general practitioners from the region around Breda (see figure 1). The Diabetes Service Breda is the only regional clinical and laboratory service for the western part of the province of North Brabant in the Netherlands.

Figure 1. A) Breda and the surrounding area. B) Black boxes indicates the Diabetes Service Breda and its outpatient clinics.

Since 1990 the Diabetes Service Breda has collected clinical and biochemical data on more than 13,000 patients with T2D. All patients are diagnosed according to WHO criteria (plasma glucose levels >11.1 mmol/l or a fasting plasma glucose level ≥ 7.0 mmol/l), and undergo clinical and laboratory evaluations for their diabetes at



regular 3-month intervals. A randomly selected portion of these patients served as probands to identify first-degree relatives (mainly sibs) who have T2D. The Medical Ethics Committee of the University Medical Centre in Utrecht approved our study protocol. Initially, 4,000 possible T2D probands from the Diabetes Service Breda were recruited in collaboration with their general practitioners.

All possible participants were sent an information letter describing the rationale of the study, background information, the inclusion criteria (having first-degree relatives who are also affected) and a questionnaire. These probands were also used to obtain information on the occurrence of T2D in relatives. Moreover, the probands were asked to invite their T2D relatives to participate in this study as well. All family members included in the study filled out an informed consent form and a questionnaire on clinical data, which included their diabetes-related medication, height and weight at present time, and at the age of 20 years.

Ten ml blood was collected from each proband by the Diabetes Service Breda and was sent to the research lab of the Department of Medical Genetics in Utrecht, where DNA was extracted from the blood samples. A similar procedure was applied for siblings who lived in the same area and were under control of the Diabetes Service Breda. Those siblings who were not under the control of the Diabetes Service Breda (e.g. diabetes controlled by another physician) were asked to go the Diabetes Service or to one of its out-patients-clinics in the surrounding area to donate 2 x 10 ml EDTA blood. Of this 20 ml, half was to sent to the research lab for DNA extraction at Utrecht and 10 ml of blood was analysed for lipid levels by the Diabetes Service Breda. Two 10 ml tubes for blood donation were sent to siblings living outside the control area of the Diabetes Service Breda, with an accompanying letter for the nurses at a nearby clinic asking them to send the samples to Utrecht and Breda.

Clinical data of all participants

We had a response of 60% on the returned questionnaires and signed informed consent forms: 570 (approx. 15%) probands donated blood for DNA extraction. Of these 570 probands, 227 had at least one affected sibling who also donated blood, making up the initial 227 families used for the genome-wide scan. Unfortunately no parents were available. However, if available, unaffected relatives were approached to donate blood for DNA extraction in order to reconstruct parental genotypes. After careful evaluation of the af-

fected siblings, 49 families were excluded because one of the affected siblings was younger than 35 at age of onset of T2D, or because of demonstrable non-paternity, using the program GRR (graphical representation of relationship errors (Abecasis *et al.* 2001)). The remaining 178 families, consisting of 420 patients and 142 unaffected siblings, were included in the genome-wide screen.

These 178 families comprised 312 affected sibpairs with an average sibship size of 3.1; there were 128 families with two affected siblings, 40 families with three affected siblings, 9 families with four affected siblings, and 1 family with five affected siblings. The participants' clinical parameters are shown in Table 1 and the information on the 178 families is shown in Table 2.

Table 1. Clinical information of all participating probands

	female (n=296)		male (n=246)	
	average ± SD	range	average ± SD	range
Age (y)	70 ± 9	45-93	69 ± 9	45-94
Age at onset (y)	61 ± 9	40-85	60 ± 9	40-86
Years of disease (y)#	10 ± 6	4-36	10 ± 6	4-37
Height (m)	1.63 ± 0.06	1.45-1.76	1.75 ± 0.07	1.50-1.93
Weight (kg)	75.5 ± 14.1	45-130	83.7 ± 12.2	56-124
BMI (kg/m^2)	28.5 ± 4.8	17.9-46.1	27.2 ± 3.4	19.2-38.7
HbA1c (%)	7.3 ± 1.1	4.5-12.9	7.3 ± 1.3	5.3-13.1
HDL cholesterol(mmol/l)	1.2 ± 0.3	0.6-2.7	1.1 ± 0.3	0.5-2.0
Total cholesterol (mmol/l)	5.5 ± 1.1	2.5-9.0	5.1 ± 1.0	2.0-8.4
Triglicerides (mmol/l	1.9 ± 0.9	0.5-5.1	1.8 ± 1.0	0.8-3.3
Insulin use (y/n/?)	30/264/2		22/222/2	
Elevated BP (y/n/?)	142/135/19		74/158/14	

#Duration of the disease from age at onset.
BMI: body mass index, HbA1c: haemoglobin A1c = glucose bound to haemoglobin, HDL: high density lipoprotein, BP: blood pressure, y/n/?: yes/no/unknown

Table 2. Clinical information of the families used in the genome-wide scan

	Affected		Unaffected	
Gender	female	male	female	male
Number (n)	235	185	87	55
Age (y)	69 ± 9	67 ± 9	64 ± 10	64 ± 10
Age at onset (y)	58 ± 10	57 ± 9		
Body weight (kg)	73.9 ± 12.2	83.0 ± 12.6	72.3 ± 15.4	78.6 ± 8.5
BMI (kg/m^2)	27.9 ± 4.1	26.9 ± 3.4	26.4 ± 4.2	25.8 ± 2.2
Affected sib pairs				
All possible	312			
Independent	239			

BMI: body mass index

Purpose of the Cohort

Genome-wide scan

Our goal of collecting at least 250 affected sibpairs with T2D was well met by obtaining 312 affected sibpairs from 178 families with at least 2 affected sibs. These families could be used to perform a genome-wide scan to search for susceptibility loci for T2D in the Dutch population.

Genome-wide scan of stratified cohort

The clinical data obtained from all the subjects could be used to stratify the cohort on different parameters, such as age at onset, BMI or medication use. These stratified subsets of the cohort could be used to repeat the genome-wide scan to find quantitative trait loci (QTL) associated with type 2 diabetes mellitus.

Candidate genes

This large Breda cohort of 542 patients could be used to perform association studies with polymorphisms from suitable candidate genes for type 2

diabetes mellitus. As stated in chapter 1 of this thesis, defects in genes involved in insulin signalling and insulin response are excellent candidate genes for type 2 diabetes mellitus.

Acknowledgements

We thank the Dutch families for their support and willingness to participate in this study. We are also grateful to the Diabetes Service Breda, to Marjan de Jong and Erdtsienck Ernste the research coordinators at the Diabetes Service Breda, and to the general practitioners for encouraging their patients to participate. We would also like to thank Tineke Righters-Aris for the contact with the patients and their families and to Harry van Someren for all the administrative work. This work was financially supported by the Dutch Diabetes Research Foundation (DFN grant 97.114 to Cisca Wijmenga and Timon W van Haeften).

References

Abecasis RC, Cherny SC, Cookson WO, Cardon LR (2001) *GRR: graphical representation of relationship errors.* Bioinformatics 17:742-743.

DeFronzo RA, Ferrannini E (1991) *Insulin resistance. A multifaceted syndrome responsible for NIDDM, obesity, hypertension, dyslipidemia, and atherosclerotic cardiovascular disease*. Diabetes Care 14:173-94.

Elbein SC (1997) *The genetics of human noninsulin-dependent (type 2) diabetes mellitus*. J Nutr 127:1891S-1896S.

Heine RJ, Nijpels G, Mooy JM (1996) *New data on the rate of progression of impaired glucose tolerance to NIDDM and predicting factors*. Diabet Med 13:S12-4.

Kaprio J, Tuomilehto J, Koskenvuo M, Romanov K, Reunanen A, Eriksson J, Stengard J, et al. (1992) *Concordance for type 1 (insulin-dependent) and type 2 (non-insulin- dependent) diabetes mellitus in a population-based cohort of twins in Finland*. Diabetologia 35:1060-7.

Medici F, Hawa M, Ianari A, Pyke DA, Leslie RD (1999) *Concordance rate for type II diabetes mellitus in monozygotic twins: actuarial analysis*. Diabetologia 42:146-50.

Nijpels G (1998) *Determinants for the progression from impaired glucose tolerance to non-insulin-dependent diabetes mellitus*. Eur J Clin Invest 28 Suppl 2:8-13.

Valsania P, Micossi P (1994) *Genetic epidemiology of non-insulin-dependent diabetes*. Diabetes Metab Rev 10:385-405.

Chapter 3

The genome-wide scan

A genome-wide scan in type 2 diabetes mellitus provides independent replication of a susceptibility locus on 18p11 and suggests existence of novel loci on 2q12 and 19q13

JHO van Tilburg[1], LA Sandkuijl[1,3], E Strengman[1], H van Someren[1], T Rigters-Aris[1], PL Pearson[1], TW van Haeften[2], C Wijmenga[1]

[1]Department of Medical Genetics, and [2]Department of Internal Medicine, University Medical Center, Utrecht, the Netherlands. [3]Departments of Medical Statistics and of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands.

Accepted with minor revisions by The Journal of Clinical Endocrinology & Metabolism (2002)

Abstract

A genome-wide scan was performed using non-parametric linkage analyses to find susceptibility loci for type 2 diabetes mellitus in the Dutch population. We studied 178 families from the Netherlands, who constituted 312 affected sibpairs. The first stage of the genome scan consisted of 270 DNA markers with an average inter-marker spacing of 13 cM. Since obesity and type 2 diabetes mellitus are inter-related, the data set was stratified for the sub-phenotype body mass index, corrected for age and gender. This resulted in a suggestive maximum multi-point LOD score of 2.3 (p-value 9.7 x 10^{-4}) for the most obese 20% pedigrees of the data set , between marker loci D18S471 and D18S843. In the lowest 80% obese pedigrees two interesting loci on chromosome 2 and 19 were found with LODs of 1.5 and 1.3, respectively (p-values 0.0075 and 0.0112). We provide solid and independent evidence that the chromosome 18p11 locus, reported earlier from a Finnish/Swedish population, is of definite interest for type 2 diabetes mellitus in connection with obesity. Subsequently, our results indicate that two novel loci may reside on chromosomes 2 and 19 with minor effects involved in the development of type 2 diabetes mellitus in the Dutch population.

Introduction

The aetiology of type 2 diabetes mellitus is unknown, but several studies indicate that the disease results from a combination of genetic susceptibility and external risk factors (DeFronzo and Ferrannini 1991). According to this multifactorial model, genetically predisposed subjects will not necessarily develop overt disease unless they are also exposed to particular environmental factors (Valsania and Micossi 1994). Important risk factors for the development of type 2 diabetes mellitus, apart from obesity, include a family history of diabetes, increased age, hypertension, lack of physical exercise, and ethnic background (DeFronzo and Ferrannini 1991).

The discovery of monogenic forms of diabetes, such as maturity-onset diabetes of the young (MODY), underscores the phenotypic and genotypic het-

erogeneity of this disease. Despite identification of at least six MODY loci to date, they account for only a few percent of all diabetes patients (Permutt and Hattersley 2000). However, defining the genetic basis of the far more common polygenic forms of type 2 diabetes mellitus presents methodological difficulties due to the absence of extended families and the small contribution of each polygene to the disease phenotype. Candidate gene studies have identified several loci with a modest effect on type 2 diabetes mellitus susceptibility (van Tilburg *et al.* 2001). Several genome-wide scans have been conducted in the past five years, giving rise to various genetic regions in as many various populations, illustrating the genetic heterogeneity of type 2 diabetes mellitus (Busfield *et al.* 2002; Duggirala *et al.* 1999; Ehm *et al.* 2000; Elbein *et al.* 1999; Ghosh *et al.* 1999; Hanis *et al.* 1996; Hanson *et al.* 1998; Hegele *et al.* 1999; Lindgren *et al.* 2002; Mahtani *et al.* 1996; Parker *et al.* 2001; Permutt *et al.* 2001; Watanabe *et al.* 2000; Wiltshire *et al.* 2001). However, at least one susceptibility gene was found using the genome-wide scan approach (CAPN10 (Horikawa *et al.* 2000)). It is therefore of great importance that replication studies are performed in additional populations, as well as independent genome scans in different populations, to confirm the original findings or to provide more insight into the genetic complexity of type 2 diabetes mellitus.

We describe a genome-wide linkage analysis to identify type 2 diabetes mellitus susceptibility loci in nuclear families from the province of North Brabant, the Netherlands (around the town of Breda). To minimize genetic heterogeneity, all nuclear families with at least two affected sibs were ascertained from a region of 130,000 inhabitants.

Subjects and methods

Population studied. Probands were recruited in collaboration with their general practitioners and the Diabetes Service Breda, which is the only regional clinical and laboratory service for the western part of the province of North Brabant in the Netherlands. Since 1990 the Diabetes Service Breda has collected clinical and biochemical data on more than 13,000 patients with type 2 diabetes mellitus. All patients undergo clinical and laboratory evaluations for their

diabetes at regular 3-month intervals. Initially, 4,000 possible participants were asked at random to take part in this study. Although we had a response of 60%, only those who had at least one affected sibling and were diagnosed after the age of 35 years according to WHO criteria were included. Unfortunately, no parents were available. However, if available, unaffected relatives were included in order to reconstruct parental genotypes. Initially 214 families fulfilled the criteria for inclusion in the study. After careful evaluation of the affected siblings, 36 families were excluded because one of the affected siblings was younger than 35 at age of onset of type 2 diabetes mellitus, or because of demonstrable non-paternity. The remaining 178 families, consisting of 420 patients and 142 unaffected siblings, were included in the genome-wide screen. These 178 families comprised 312 affected sibpairs with an average sib-ship size of 3.1; there were 128 families with two affected siblings, 40 families with three affected siblings, 9 families with four affected siblings, and one family with five affected siblings. The Medical Ethics Committee of the University Medical Center in Utrecht approved our study protocol and all the participants signed an informed consent. The participants' clinical parameters are shown in Table 1.

Genotyping. DNA was extracted from 10 ml of blood using standard procedures (Miller *et al.* 1988). The 551 DNA samples were divided into eight 96-well microtiter plates. Every DNA plate contained up to 80 unique DNA samples, six blind duplicate samples, three CEPH controls and one negative control. A modified version of the Weber set 6 containing 270 markers (73% of markers from the Weber 6 map) at an average spacing of 13 cM was used for the genome-wide screen. For details of the markers see: http:humgen.med.uu.nl/publications/jonathan2002_1/. Reverse primers were labelled with either 6-FAM, HEX, or TET fluorescent dyes (Isogen Bioscience, the Netherlands) at the 5'-end. PCR was carried out in a 10 ml volume containing 1 x PCR Gold-buffer*, 200 mmol/l of deoxy-NTP, 2.5 mmol/l $MgCl_2$, 25 ng/ml of each primer, 0.4 U AmpliTaq Gold* and 25 ng genomic DNA. Cycling conditions were 7 min at 94°C followed by 32 cycles of 30s at 95°C,

Table 1. Clinical information of participants

	Affected		Unaffected	
Gender	female	male	female	male
Number (n)	235	185	87	55
Age (y)	69 ± 9	67 ± 9	64 ± 10	64 ± 10
Age at onset (y)	58 ± 10	57 ± 9		
Body weight (kg)	73.9 ± 12.2	83.0 ± 12.6	72.3 ± 15.4	78.6 ± 8.5
BMI (kg/m^2)	27.9 ± 4.1	26.9 ± 3.4	26.4 ± 4.2	25.8 ± 2.2
Affected sib pairs				
All possible	312			
Independent	239			

30s at 55°C and 30s at 72°C, followed by a final extension at 72°C for 30 min. PCR products were pooled into four different running sets, before electrophoresis on an ABI 3700*, and analyzed using GeneScan version 3.1*. Allele sizes of the individual markers were determined using Genotyper version 2.1 software*. CEPH reference samples (1331-01, 1331-02 and 1347-02) were included to determine the appropriate size of the alleles. The 48 duplicate samples were included to estimate the proportion of mistyping of genotypes. All samples where double-checked by two independent investigators (JHOvT and ES), who did not know the origin of the 48 duplicate samples. The blind genotypes of the duplicate samples were compared to the original samples by a technician (Alfons Bardoel). (*Applied Biosystems, Foster City, Ca, USA).

Stratification of the dataset. In unaffected individuals (n=150) the relationship between BMI and age and gender was determined via multiple linear regressions. For all family members, BMI was adjusted for age and gender according to the resulting regression coefficient. Normal percentile values were obtained from the adjusted BMI values in unaffected individuals. All affected individuals were then classified according to these percentiles. Only sibpairs

in which both affected siblings fitted in the same percentile group were included in the stratified data set. As a consequence the 20% most obese pedigrees group (referred to as 'DS-20%') contained 44 affected sibpairs. As a counterpart to the DS-20% group, there was the group without the 20% most obese pedigrees (referred to as'DS-80%'); only sibpairs who fit in this group were included giving 146 affected sibpairs (Table 2).

Table 2. Distribution of variables used for subphenotypic classification according to BMI.

	Whole data set		Affected		Highest percentile threshold values	
	Mean ± SD	Range	Mean ± SD	Range	DS-20%	DS-80%
age (y)	67 ± 9	33 - 96	68 ± 9	41 - 96	65 ± 10	68 ± 9
age at onset (y)	58 ± 10	35 - 85	57 ± 11	35 - 85	55 ± 10	58 ± 10
Weight (kg)	77.5 ± 13.3	40 - 160	78.3 ± 13.3	40 - 128	90.3 ± 11.2	72.4 ± 11.3
BMI (kg/m^2)	27.2 ± 3.8	17.2 - 43.3	27.5 ± 3.9	17.1 - 43.3	31.9 ± 2.6	25.3 ± 2.4
No. of families	178				30	99
No. of affected sibpairs	312				44	146

Statistical analyses. To assess for linkage, we applied multi-point non-parametric linkage analysis using the MapMaker/Sibs software 2.0 package (Kruglyak and Lander 1995). Allele frequencies were calculated from the whole data set, and the weighted sibpair option in MapMaker/Sibs was used. For analysis of the entire length of the different chromosomes, we used genetic map distances estimated from the Marshfield genetic map (http://research.marshfieldclinic.org/genetics/, see also the complementary information on the web site: http:humgen.med.uu.nl/publications/jonathan2002_1/). Exclusion mapping was performed using the exclude option of MapMaker/Sibs, under an additive model and at several locus-specific values of λ_s (the ratio of the risk to sibling of an affected person relative to the risk to a member of the general population), ranging from 1.25 to 2.5. LOD scores for exclusion of a region were obtained by comparing the likelihood of the data assuming the presence of a locus with a specific effect (λ_s) to the likelihood if the region contained no relevant locus at all ($\lambda_s = 1$).

Power calculations. Computer simulations were carried out on an initial marker map with a 20 cM spacing, with subsequent markers at 2 cM intervals, 250 sibpairs, and marker information > 0.7. This simulation showed that, with the initial marker map, we could only extract about 55% of all the available information from the sibpairs. After saturation with additional markers, on average 94% of the maximal information (which would be for a completely linked and completely informative marker) could be extracted via the MapMaker/Sibs program. Under the assumption of a single disease-predisposing locus for type 2 diabetes mellitus (locus-specific λ_s = 3.5), there was ~ 100% power for detection with a LOD score of 3 (p-value < 0.0001), when the initial map of 270 markers was used. The power to detect a locus with a modest effect (locus-specific λ_s = 1.5) is ~ 70%. It should be realized, however, that if there are multiple disease-predisposing loci, each will present an opportunity for mapping. Therefore, while our sample size yields a power of 70% to detect a single locus with λ_s of 1.5, if there are two such loci, our power to detect at least one of them is 0.70+(0.70 x (1-0.70)) = 0.91. Since type 2 diabetes mellitus is expected to be genetically heterogeneous, we were aiming at a LOD score of 0.5 (p-value of 0.09 or less) for our initial scan. To detect a locus with a relative small effect (λ_s = 1.5) the power is ~ 99% for detection with a LOD score of 0.5.

Results

The autosomal genome scan was completed on 178 families consisting of 417 patients and 134 unaffected siblings, which were used to generate the final data set for statistical analysis. By including additional siblings, both affected and unaffected, we obtained an average sibship size of 3.1. The first stage of the genome scan consisted of 270 DNA markers with an average inter-marker spacing of 13 cM and a mean heterozygosity of 0.76. Forty-eight duplicate samples were included to estimate the proportion of typing errors. An average of 90% of subjects was successfully genotyped for each marker with less than 3% of mistyping of genotypes. The average informa-

tion content was 50% throughout the genome (measuring the proportion of the total inheritance information extracted at each chromosomal position given the observed genotype data (Kruglyak *et al.* 1996)), see for details http:humgen.med.uu.nl/publications/jonathan2002_1/.

Only four genomic regions initially showed multi-point LOD scores ≥ 0.5 (Figure 1). On chromosome 2 a LOD score of 0.5 was obtained between markers D2S436 and D2S410, on chromosome 11 a maximum LOD score of 0.9 was obtained between markers D11S1984 and D11S2362, chromosome 14 showed a LOD score of 0.9 between markers D14S53 and D14S606, whereas on chromosome 18 a LOD score of 0.6 was obtained near marker D18S843. These four chromosomal regions were selected for follow up studies (Table 3).

Fine maps of the regions exhibiting LOD scores ≥ 0.5 were constructed with an average inter-marker spacing of 5 cM or less. The original 178 type 2 diabetes mellitus families were then genotyped using these fine maps. The results of the follow up studies are summarized in Table 3. The addition of extra markers (10, 3, 6 and 15 markers for chromosomes 2, 11, 14 and 18 respectively) gave the following results: the multi-point LOD at chromosome 14 decreased from 0.9 to 0.3 at marker position D14S53 whereas the multi-point LOD on chromosome 11 the LOD score decreased slightly from 0.9 to 0.8. No change in the multi-point LOD for chromosome 2 was seen after addition of extra markers. However, on chromosome 18 the multi-point LOD increased slightly when more markers where analyzed in the region of interest. Analysing more markers on chromosome 18 increased the multi-point LOD from 0.6 at marker position D18S843 to a multi-point LOD of 0.7 between markers D18S1163 and D18S843 (Table 2). For additional information on the markers, see http:humgen.med.uu.nl/publications/jonathan2002_1/.

In addition to searching for evidence of linkage, we performed exclusion mapping to determine which genomic regions could be excluded as candidates for harbouring major susceptibility loci. Five different locus-specific values of λ_s were considered, 1.25, 1.5, 1.75, 2.0 and 2.5. We could exclude 87% of the genome for a λ_s of 2.0 (for details see http:humgen.med.uu.nl/

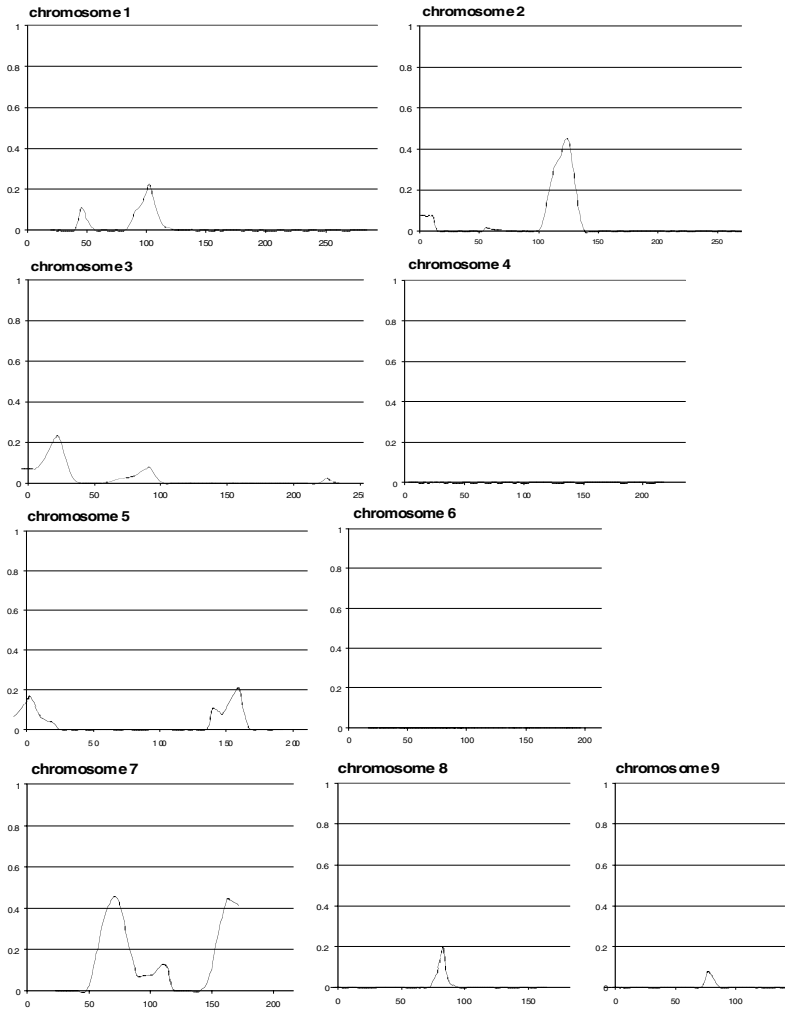
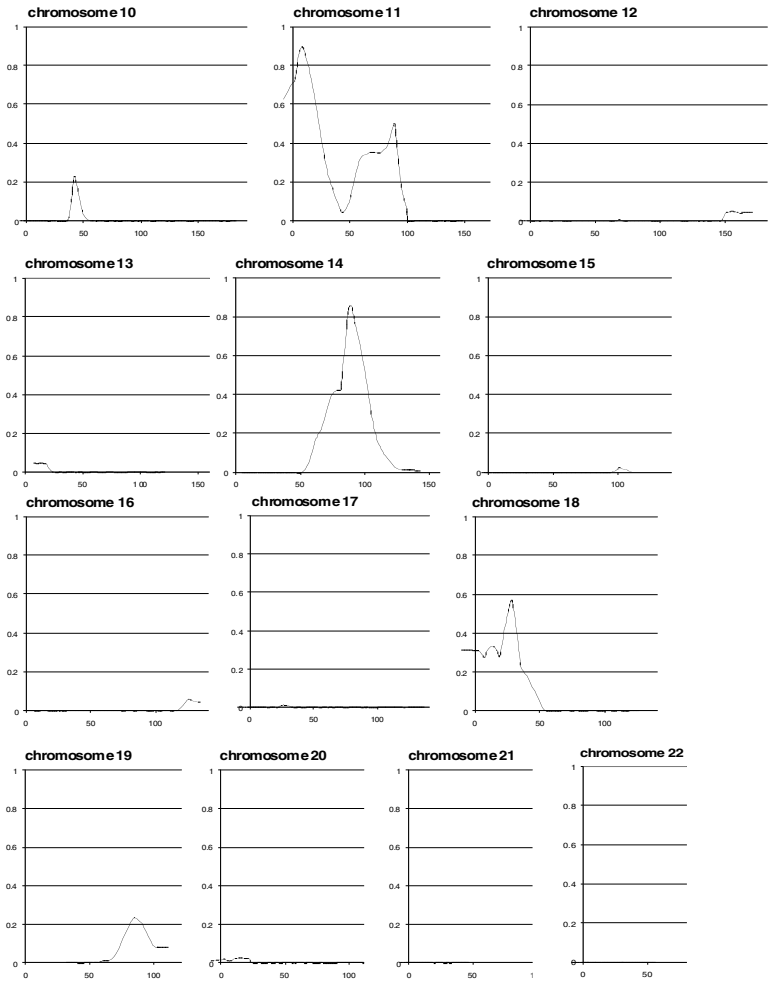


Figure 1. Graphs of the multi-point LOD scores for each autosome from the genome-wide scan in 312 affected sibpairs with 270 DNA markers, using the non-parametric linkage analysis program, MapMaker/Sibs. The X-axis represents the length of the chromosome in cM whereas the Y-axis represents the multi-point LOD score.


chromosome 10
chromosome 11
chromosome 12
chromosome 13
chromosome 14
chromosome 15
chromosome 16
chromosome 17
chromosome 18
chromosome 19
chromosome 20
chromosome 21
chromosome 22

publications/jonathan2002_1/). However, for a λ_s of 1.25 we could only exclude 7% of the genome with a sample of 312 affected sibpairs and in the absence of parental data. This is almost certainly due to the limited size of our data set in resolving such low predisposition values, suggesting that large numbers of affected sibpairs are necessary to identify loci with minor effects in type 2 diabetes mellitus.

Stratification for obesity

A previous study by Parker et al. (Parker *et al.* 2001) had reported a locus on chromosome 18 whose genetic contribution increased by stratifying the data set for BMI. When a comparable stratification was applied to our data set (Table 3), the multi-point LOD increased from 0.7 in the unstratified sample to LOD 2.5 (between markers D18S452 and D18S1163) when the DS-20% pedigrees were analyzed (Table 3). Adding another 8 markers to this region with a resolution of one marker every 1.5 cM slightly decreased the LOD to a maximum LOD of 2.3 for the DS-20% pedigrees, between markers D18S471 and D18S843 near marker D18S1163 (Table 3).

Figure 2 illustrates the evident effect of BMI stratification on the multipoint LOD score, despite the reduction in sample size. Only 30 families containing 44 affected diabetes sibpairs comprised the DS-20% type 2 diabetes mellitus pedigrees. The mean BMI of the affected individuals in this subphenotype group of DS-20% pedigrees was 31.9 kg/m^2 versus 27.5 kg/m^2 for all affected individuals in the whole data set.

Although the overall marker information content was relatively low, namely 50% (see http:humgen.med.uu.nl/publications/jonathan2002_1/), it did reach 78% at the map position showing the highest evidence for linkage on 18p, with an average information content of 91% between D18S391 and D18S1163.

Genome screen in DS-80%

The limited number of interesting loci resulting from the genome-wide scan prompted us to re-calculate the LOD scores for the affected sibpairs in which both sibs were in the DS-80% group, as we thought this group could

Table 3. Regions displaying multi-point LOD scores ≥ 0.5 in Dutch Caucasian siblings with type 2 diabetes mellitus.

Chromosome	Marker	cM	Het[1]	LOD	p-value[2]
2	D2S436	118.2	0.90	0.3	0.16
	interval	123.7		0.5	0.09
	D2S410	125.2	0.81	0.3	0.16
11	D11S1984	2.1	0.76	0.7	0.04
	interval	7.6		0.9	0.03
	D11S2362	8.9	0.93	0.9	0.04
14	D14S53	86.3	0.67	0.8	0.12
	interval	88.5		0.9	0.03
	D14S606	91.6	0.69	0.8	0.05
18	D18S843	28.1		0.6	0.07
Dense map[3]					
2	D2S436	118.2	0.90	0.0	0.58
	D2S1888	121.6	0.71	0.2	0.22
	interval	122.9		0.5	0.09
	D2S410	125.2	0.81	0.4	0.12
11	D11S2362	8.9	0.93	0.8	0.04
	interval	10.2		0.8	0.04
	D11S1999	17.9	0.76	0.9	0.04
14	D14S53	86.3	0.67	0.3	0.16
	interval	88.5		0.3	0.16
	D14S606	91.6	0.69	0.1	0.22
18	D18S1163	24.1	0.54	0.7	0.05
	interval	27.2		0.7	0.05
	D18S843	28.1		0.7	0.05
Stratification[4]					
DS-20%	D18S452	18.7	0.83	2.3	9.5×10^{-4}
	interval	24.0		2.5	5.5×10^{-4}
	D18S1163	24.1	0.54	2.4	7.1×10^{-4}
DS-20%[5]	D18S471	19.3		1.6	0.0052
	interval	24.2		2.3	9.7×10^{-4}
	D18S843	28.1		1.8	0.003

[1] Het = Heterozygosity of the marker
[2] p-value according to Holmans' possible triangle method (Holmans 1993)
[3] With the addition of more markers at a ≤ 5 cM spacing in the region of interest
[4] Analysis using data-set stratified on age- and sex- adjusted BMI
[5] With the addition of 5 extra markers in the region between D18S452 and D18S1163 on chromosome 18

represent a different sub-phenotype. The DS-80% consisted of 146 affected sibpairs from 99 families with 220 patients and 72 unaffected siblings, with an average sibship size of 2.9 (Table 2). The number of markers was increased from 270 to 319, since markers were added in order to fine map the 5 regions already identified. Recalculation of the LOD scores was performed for 319

markers; interestingly, three genomic regions showed multi-point LOD scores ≥ 1.0 (Table 4). On chromosome 2 a LOD score of 1.5 was obtained between markers D2S436 and D2S1888, a region also be identified in the whole data set. Since this region had already been saturated with markers every 5 cM or less we did not perform further fine mapping on this region. On chromosome 15 a maximum LOD score of 3.1 was obtained between markers D15S817 and the gene ACTC, whereas on chromosome 19 a LOD score of 1.5 was obtained between markers D19S400 and D19S254. All results of the recalculation can be seen on http:humgen.med.uu.nl/publications/jonathan2002_1/

Both regions on chromosome 15 and 19 were further investigated by fine mapping. Addition of 4 markers on chromosome 15 showed that the previous finding was a false positive finding because the maximum LOD score decreased from 3.1 to 0.3 near the ACTC gene. However, adding 3 markers to the map of chromosome 19 gave the following result; the maximum LOD score on chromosome 19 decreased slightly from 1.5 to 1.3 between markers D19S246 and D19S601 (Table 4).

Discussion

We performed a genome-wide linkage analysis study designed to identify type 2 diabetes mellitus susceptibility loci in type 2 diabetes mellitus nuclear families from Breda, North Brabant, the Netherlands. Our results contribute to a better understanding of type 2 diabetes mellitus, combined with results of similar studies in other populations, including African Americans (Ehm *et al.* 2000), Ashkenazi Jews (Permutt *et al.* 2001), British (Wiltshire *et al.* 2001), Chinese (Luo *et al.* 2001), European Americans (Ehm *et al.* 2000; Elbein *et al.* 1999), Finnish (Ghosh *et al.* 1999; Lindgren *et al.* 2002; Mahtani *et al.* 1996; Parker *et al.* 2001; Watanabe *et al.* 2000), French (Vionnet *et al.* 2000), Han Chinese (Luo *et al.* 2001), Mexican Americans (Duggirala *et al.* 1999; Ehm *et al.* 2000; Hanis *et al.* 1996), and native Americans from the US and Canada (Hanson *et al.* 1998; Hegele *et al.* 1999). However most of these scans have failed to generate highly significant linkage results. Replications of certain

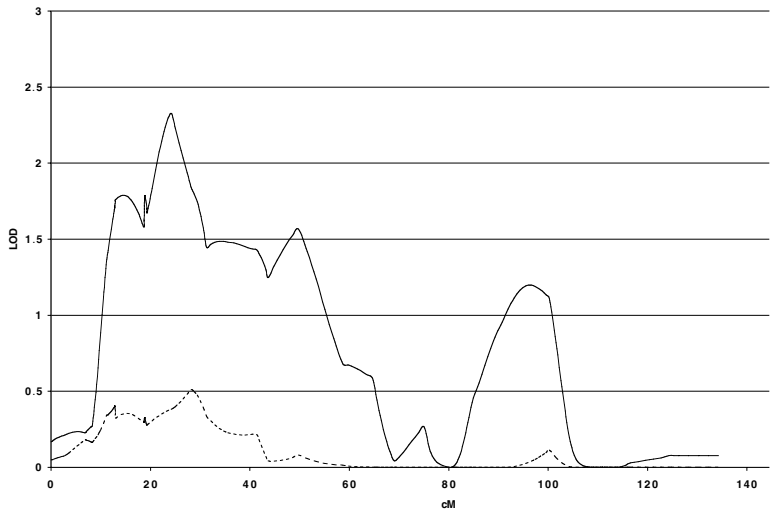


Figure 2. Fine map (33 markers) analysis of chromosome 18 in a data set stratified by mean age- and sex-adjusted BMI. The dashed line indicates the whole data set and the solid line the DS-20% set. The confidence limits define this locus as lying between D18S967and D18S1153, a region of ~20 cM.

loci of these different scans can therefore direct further investigations toward positional cloning targeting the most promising loci.

Our genome scan of a stratified sample of 146 Dutch families with type 2 diabetes mellitus using non-parametric linkage analysis and exclusion mapping revealed modest indications of linkage to regions on chromosomes regions 2q12 (multi-point LOD score of 1.5 with corresponding p-value of 0.0075) and 19q13 (multi-point LOD score of 1.3 with corresponding p-value of 0.0112). Although we identified interesting LOD scores on chromosomes 2, our locus mapped outside the previously described CAPN10 region on chromosome 2 (Cox *et al.* 1999; Hanis *et al.* 1996; Horikawa *et al.* 2000). This locus lies approximately 105 cM distal of the locus found in our population. The locus found on chromosome 19 is not previously described in literature.

The results on chromosome 18 in the DS-20% group indicate linkage

between D18S471 and D18S843 (multi-point LOD score of 2.3 with corresponding p-value of 9.7×10^{-4}). This finding on chromosome 18 clearly replicates the linkage to this region previously reported in a Finnish-Swedish population (Parker *et al.* 2001) and in a confined isolated population of Dutch Caucasians near the town of Breda (Prof. B.A. Oostra, Erasmus University Rotterdam, personal communication). This is interesting because it suggests our finding is genuine since our stratified group (DS-20%) consists of only 44 affected sibpairs.

Table 4. Regions displaying multi-point LOD scores ≥ 1.0 in 146 DS-80% sibpairs

Chromosome	Marker	cM	Het[1]	LOD	p-value[2]
2[3]	D2S436	118.2	0.90	1.0	0.0239
	interval	121.4		1.5	0.0075
	D2S1888	121.6	0.71	1.4	0.0092
15	D15S817	4.8	0.79	0.8	0.0446
	interval	20.3		3.1	1.54×10^{-4}
	ACTC	31.5	0.94	0.8	0.0423
19	D19S400	64.7	0.86	0.6	0.0683
	interval	79.5		1.5	0.0063
	D19S245	100.6	0.76	0.3	0.1600
Dense map					
15	ACTC	31.5	0.94	0.3	0.1802
19	D19S246	78.1	0.84	1.0	0.0239
	interval	80.0		1.3	0.0112
	D19S601	83.2	0.81	1.2	0.0144

[1] Het = Heterozygosity of the marker
[2] p-value according to Holmans' possible triangle method (Holmans 1993)
[3] This is already a dense map with markers at 5 cM or less.

Replication of linkage results from additional populations, whether as extension or follow up studies in the same population or as independent genome scans in different populations, may provide vital confirmation of the original findings. Guidelines have been proposed for the level of significance necessary in sibpair studies in order to show replication (Lander and Kruglyak 1995; Roberts *et al.* 1999).

As for the significance of our finding with respect to the null hypothesis - no disease susceptibility locus present in the study sample - Lander and Kruglyak (1995) defined suggestive linkage for a LOD score of 2.2 (p-value

= 0.001) or higher whereas significant linkage is defined only when the LOD score is 3.6 (p-value = 2 x 10^{-5}) or higher. The LOD score obtained on chromosome 18p11 (in DS-20%) meets the criteria for suggestive linkage. Although the loci found on chromosome region 2q12 and chromosome region 19q13 do not achieve suggestive LOD scores, they may indicate the existence of novel loci with minor effects involved in type 2 diabetes
mellitus in the Dutch population.

Apart from the findings on chromosome 2, 18 and 19, our genome scan did not provide any supporting evidence for the other linkage findings described earlier. Several reports have described linkage to chromosomes 12q (Bowden *et al.* 1997; Elbein *et al.* 1995) and 20q (Bowden *et al.* 1997; Elbein *et al.* 1999; Ghosh *et al.* 1999; Permutt *et al.* 2001). We observed no evidence for linkage to either of these two chromosome regions in our study.

This absence of strong agreement among genome-wide scans in type 2 diabetes mellitus is not unexpected, in part because of the complexity of the disease involved. Gene discovery in complex diseases has been limited by substantial etiological and genetic heterogeneity, the possibility of genes of small effect, the interaction of multiple genes with each other and environmental factors, and the need for large sample sizes (Altmuller *et al.* 2001). A typical 10 cM genome scan fails to capture a significant proportion of the inheritance information, especially in cases of small sibships and lack of parental genotype information, as is usually the case in late-onset disorders such as type 2 diabetes mellitus (Wiltshire *et al.* 2001). The inclusion of additional sibs to reconstruct parental haplotypes may compensate for part of the lost information. A two-stage screening design with denser mapping in regions of interest identified by the primary low-resolution scan has also been proposed to enhance power by recovering some of the missed information. However, this approach may still miss regions of linkage if the evidence for linkage has, by chance, been underestimated in the primary scan such that thresholds for dense mapping were not achieved (Wiltshire *et al.* 2001). On the other hand, as shown in this study and others (Parker *et al.* 2001), sub-phenotyping of the disease may lead to substantial improvement in the results from genome-

wide scans for type 2 diabetes mellitus.

The role of a susceptibility locus for type 2 diabetes mellitus on chromosome 18p11 and BMI, suggested by Parker et al. (Parker *et al.* 2001) and independently replicated in our Dutch population, is not known. It has often been suggested that obesity plays a causal role in the development in type 2 diabetes mellitus and is the most important determinant of type 2 diabetes mellitus (Trevisan *et al.* 1998). So far, the region on chromosome 18p11 has not been found in genome-wide scans for obesity. The association between type 2 diabetes mellitus and obesity is presumably due to multiple mechanisms, including elevations in plasma free fatty acids (FFA) and tumour necrosis factor-alpha (TNFα) released from "full" adipocytes (Hotamisligil *et al.* 1995; Pi-Sunyer 1993; Reaven 1988; Uysal *et al.* 1997). Additional support for this hypothesis was seen in the recent demonstration of significantly increased β-cell apoptosis in obese versus lean ZDF rats (Shimabukuro *et al.* 1998).

Whether this chromosome 18p11 locus is a primary obesity locus, or a locus which is important for the development of type 2 diabetes mellitus in already overweight individuals is a question for the future. Substantial new research will be required to resolve this issue. Analysis of the region on chromosome 18 in the sequence databases revealed no obvious candidate gene. The chromosomal 18p11 region contains 111 transcripts of which 61 are of unknown function and 50 either resemble a known function or have already been described. Much genetic analysis will have to be carried out to implicate the correct candidate gene.

In conclusion, our results indicate that a novel gene resides in the 18p11 region, which forms part of an as yet unidentified pathway involved in type 2 diabetes mellitus and obesity, and that are indications of two novel loci on chromosome 2 and 19 both with minor effects involved in the development of type 2 diabetes mellitus in the Dutch population.

Acknowledgements

We thank the Dutch families for their support and willingness to participate in this study. We are also grateful to Dr. Joop Freriks, director of the Diabetes Service Breda and the nurses for collecting the blood samples, and to the general practitioners for encouraging their patients to participate. This work was financially supported by the Dutch Diabetes Research Foundation (DFN grant 97.114 to CW and TWvH). We would also like to thank Mr. Tsing Cheng for his contribution to the chromosome 18 work, Mr. Alfons Bardoel for checking the blind controls, and Ms. Jackie Senior for improving the manuscript.

References

Altmuller J, Palmer LJ, Fischer G, Scherb H, Wjst M (2001) *Genomewide scans of complex human diseases: true linkage is hard to find*. Am J Hum Genet 69:936-50.

Bowden DW, Sale M, Howard TD, Qadri A, Spray BJ, Rothschild CB, Akots G, et al. (1997) *Linkage of genetic markers on human chromosomes 20 and 12 to NIDDM in Caucasian sib pairs with a history of diabetic nephropathy*. Diabetes 46:882-6.

Busfield F, Duffy DL, Kesting JB, Walker SM, Lovelock PK, Good D, Tate H, et al. (2002) *A genomewide search for type 2 diabetes-susceptibility genes in indigenous Australians*. Am J Hum Genet 70:349-57.

Cox NJ, Frigge M, Nicolae DL, Concannon P, Hanis CL, Bell GI, Kong A (1999) *Loci on chromosomes 2 (NIDDM1) and 15 interact to increase susceptibility to diabetes in Mexican Americans*. Nat Genet 21:213-5.

DeFronzo RA, Ferrannini E (1991) *Insulin resistance. A multifaceted syndrome responsible for NIDDM, obesity, hypertension, dyslipidemia, and atherosclerotic cardiovascular disease*. Diabetes Care 14:173-94.

Duggirala R, Blangero J, Almasy L, Dyer TD, Williams KL, Leach RJ, O'Connell P, et al. (1999) *Linkage of type 2 diabetes mellitus and of age at onset to a genetic location on chromosome 10q in Mexican Americans*. Am J Hum Genet 64:1127-40.

Ehm MG, Karnoub MC, Sakul H, Gottschalk K, Holt DC, Weber JL, Vaske D, et al. (2000) *Genomewide Search for Type 2 Diabetes Susceptibility Genes in Four American Populations*. Am J Hum Genet 66:1871-1881.

Elbein SC, Chiu KC, Hoffman MD, Mayorga RA, Bragg KL, Leppert MF (1995) *Linkage analysis of 19 candidate regions for insulin resistance in familial NIDDM*. Diabetes 44:1259-65.

Elbein SC, Hoffman MD, Teng K, Leppert MF, Hasstedt SJ (1999) *A genome-wide*

search for type 2 diabetes susceptibility genes in Utah Caucasians. Diabetes 48:1175-82.

Ghosh S, Watanabe RM, Hauser ER, Valle T, Magnuson VL, Erdos MR, Langefeld CD, et al. (1999) *Type 2 diabetes: evidence for linkage on chromosome 20 in 716 Finnish affected sib pairs*. Proc Natl Acad Sci U S A 96:2198-203.

Hanis CL, Boerwinkle E, Chakraborty R, Ellsworth DL, Concannon P, Stirling B, Morrison VA, et al. (1996) *A genome-wide search for human non-insulin-dependent (type 2) diabetes genes reveals a major susceptibility locus on chromosome 2*. Nat Genet 13:161-6.

Hanson RL, Ehm MG, Pettitt DJ, Prochazka M, Thompson DB, Timberlake D, Foroud T, et al. (1998) *An autosomal genomic scan for loci linked to type II diabetes mellitus and body-mass index in Pima Indians*. Am J Hum Genet 63:1130-8.

Hegele RA, Sun F, Harris SB, Anderson C, Hanley AJ, Zinman B (1999) *Genome-wide scanning for type 2 diabetes susceptibility in Canadian Oji- Cree, using 190 microsatellite markers*. J Hum Genet 44:10-4.

Holmans P (1993) *Asymptotic properties of affected-sib-pair linkage analysis*. Am J Hum Genet 52:362-74.

Horikawa Y, Oda N, Cox NJ, Li X, Orho-Melander M, Hara M, Hinokio Y, et al. (2000) *Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus*. Nat Genet 26:163-75.

Hotamisligil GS, Arner P, Caro JF, Atkinson RL, Spiegelman BM (1995) *Increased adipose tissue expression of tumor necrosis factor-alpha in human obesity and insulin resistance*. J Clin Invest 95:2409-15.

Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES (1996) *Parametric and nonparametric linkage analysis: a unified multipoint approach*. Am J Hum Genet 58:1347-63.

Kruglyak L, Lander ES (1995) *Complete multipoint sib-pair analysis of qualitative and quantitative traits*. Am J Hum Genet 57:439-54.

Lander E, Kruglyak L (1995) *Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results.* Nat Genet 11:241-7.

Lindgren CM, Mahtani MM, Widen E, McCarthy MI, Daly MJ, Kirby A, Reeve MP, et al. (2002) *Genomewide search for type 2 diabetes mellitus susceptibility Loci in finnish families: the botnia study*. Am J Hum Genet 70:509-16.

Luo TH, Zhao Y, Li G, Yuan WT, Zhao JJ, Chen JL, Huang W, et al. (2001) *A genome-wide search for type II diabetes susceptibility genes in Chinese Hans*. Diabetologia 44:501-6.

Mahtani MM, Widen E, Lehto M, Thomas J, McCarthy M, Brayer J, Bryant B, et al. (1996) *Mapping of a gene for type 2 diabetes associated with an insulin secretion defect by a genome scan in Finnish families*. Nat Genet 14:90-4.

Miller SA, Dykes DD, Polesky HF (1988) *A simple salting out procedure for extracting DNA from human nucleated cells*. Nucleic Acids Res 16:1215.

Parker A, Meyer J, Lewitzky S, Rennich JS, Chan G, Thomas JD, Orho-Melander M,

et al. (2001) *A gene conferring susceptibility to type 2 diabetes in conjunction with obesity is located on chromosome 18p11*. Diabetes 50:675-80.

Permutt MA, Hattersley AT (2000) *Searching for Type 2 Diabetes Genes in the Post-genome Era*. Trends Endocrinol Metab 11:383-393.

Permutt MA, Wasson JC, Suarez BK, Lin J, Thomas J, Meyer J, Lewitzky S, et al. (2001) *A genome scan for type 2 diabetes susceptibility loci in a genetically isolated population*. Diabetes 50:681-5.

Pi-Sunyer FX (1993) *Medical hazards of obesity*. Ann Intern Med 119:655-60.

Reaven GM (1988) *Banting lecture 1988. Role of insulin resistance in human disease*. Diabetes 37:1595-607.

Roberts SB, MacLean CJ, Neale MC, Eaves LJ, Kendler KS (1999)

Replication of linkage studies of complex traits: an examination of variation in location estimates. Am J Hum Genet 65:876-84.

Shimabukuro M, Zhou YT, Levi M, Unger RH (1998) *Fatty acid-induced beta cell apoptosis: a link between obesity and diabetes*. Proc Natl Acad Sci U S A 95:2498-502.

Trevisan R, Vedovato M, Tiengo A (1998) *The epidemiology of diabetes mellitus*. Nephrol Dial Transplant 13:2-5.

Uysal KT, Wiesbrock SM, Marino MW, Hotamisligil GS (1997) *Protection from obesity-induced insulin resistance in mice lacking TNF- alpha function*. Nature 389:610-4.

Valsania P, Micossi P (1994) *Genetic epidemiology of non-insulin-dependent diabetes*. Diabetes Metab Rev 10:385-405.

van Tilburg J, van Haeften TW, Pearson P, Wijmenga C (2001) *Defining the genetic contribution of type 2 diabetes mellitus*. J Med Genet 38:569-78.

Vionnet N, Hani EH, Dupont S, Gallina S, Francke S, Dotte S, De Matos F, et al. (2000) *Genomewide Search for Type 2 Diabetes-Susceptibility Genes in French Whites: Evidence for a Novel Susceptibility Locus for Early-Onset Diabetes on Chromosome 3q27-qter and Independent Replication of a Type 2-Diabetes Locus on Chromosome 1q21-q24*. Am J Hum Genet 67:1470-1480.

Watanabe RM, Ghosh S, Langefeld CD, Valle TT, Hauser ER, Magnuson VL, Mohlke KL, et al. (2000) *The Finland-United States Investigation of Non-Insulin-Dependent Diabetes Mellitus Genetics (FUSION) Study. II. An Autosomal Genome Scan for Diabetes-Related Quantitative-Trait Loci*. Am J Hum Genet 67:1186-1200.

Wiltshire S, Hattersley AT, Hitman GA, Walker M, Levy JC, Sampson M, O'Rahilly S, et al. (2001) *A genomewide scan for loci predisposing to type 2 diabetes in a U.K. population (the Diabetes UK Warren 2 Repository): analysis of 573 pedigrees provides independent replication of a susceptibility locus on chromosome 1q*. Am J Hum Genet 69:553-69.

Chapter 4

QTL mapping

Confirming evidence for a susceptibility locus influencing BMI and type 2 diabetes on chromosome 11q

Jonathan H van Tilburg[1], Lodewijk A Sandkuijl[1,3], Eric Strengman[1], Peter L Pearson[1], Timon W van Haeften[2] and Cisca Wijmenga[1]

[1]Department of Medical Genetics, and [2]Department of Internal Medicine, University Medical Center, Utrecht, the Netherlands. [3]Departments of Medical Statistics and of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands.

Submitted for publication

Abstract

We analyzed data from 178 small family samples for linkage to loci influencing BMI; the data came from the Dutch Breda Cohort ascertained for type 2 diabetes. Subsequently, we also analyzed the data from the 20% most obese type 2 diabetes pedigrees for linkage to type 2 diabetes. We used variance-components analysis implemented in GENEHUNTER 2 to determine QTL influencing BMI and we used affected sib pair analysis implemented in MAPMAKER/SIBS to search for linkage in obese-driven type 2 diabetes. Our findings support previous results that a susceptibility locus influencing BMI in type 2 diabetes may reside on chromosome 11q. Additionally, we found evidence to suggest linkage for type 2 diabetes on chromosome regions 1q, 11p, 12q and 18p11, and to confirm previous findings to the corresponding regions. However, it appears that the linkage found in the present Breda Cohort of type 2 diabetes patients is influenced by obesity. This supports the notion that a genetic predisposition to obesity is probably intertwined with a genetic predisposition to type 2 diabetes; further efforts should address the question how, on a genetic level, the two interact.

Introduction

The etiology of type 2 diabetes is ill defined: several studies indicate that the disease results from a combination of genetic susceptibility and external risk factors (DeFronzo and Ferrannini 1991). According to this multi-factorial model, genetically predisposed subjects will not necessarily develop overt disease unless they are also exposed to particular environmental factors (Valsania and Micossi 1994). Important risk factors for the development of type 2 diabetes include a family history of diabetes, increased age, hypertension, lack of physical exercise, and obesity (DeFronzo and Ferrannini 1991).

Several genome-wide scans for linkage with type 2 diabetes have been conducted over the past five years, and have detected linkage with many genetic loci in various populations. This illustrates either the genetic heterogeneity of type 2 diabetes or the inability to replicate linkage with defined loci. However, at least one susceptibility gene, namely CAPN10, was found using

a genome-wide scan approach (Horikawa *et al.* 2000).

While obesity is a risk factor for type 2 diabetes it is also a complex trait determined by multiple genetic and environmental factors (including physiological, behavioral, and sociocultural factors) (Perusse and Bouchard 1999). In recent years, several single-gene defects responsible for obesity have been identified in rodents and also in humans in rare instances of extended families. In addition to leptin (OMIM:164160), which is the most notable example, numerous other proteins and neuropeptides have recently been found that participate in a complex network regulating food intake and energy expenditure (Rankinen *et al*. 2002). The genetic relationship between type 2 diabetes and obesity appears complex and it is unknown how these two diseases influence each other at the genetic level. It seems unlikely that all forms of obesity will be associated with type 2 diabetes, or vice versa, and therefore any possible direct link between the two at the genetic level would probably be limited to a subset of patients.

To study further the relationship between type 2 diabetes and obesity we performed a genome-wide screen in a cohort of type 2 diabetes patients with known BMI values. The aim of the study was two-fold. The first aim was to find loci modulating BMI in type 2 diabetes patients, using variance components (VC) analysis. The second aim was to identify loci responsible for the restricted phenotype 'obesity-driven type 2 diabetes ', using affected sibpair (ASP) analysis in obese diabetes patients (mean BMI 31.9 ± 2.6).

The resulting genotypes were analyzed in two ways. First the data was used to determine linkage for BMI in type 2 diabetes patients according to a continuous scale to define Quantitative Traits Loci (QTL) using the adjusted BMI values of affected individuals for type 2 diabetes. Secondly, the data were analyzed with the ASP method to find BMI susceptibility loci involved in BMI levels (obesity) in type 2 diabetes patients.

If loci found in both analyses show overlap, these loci might possibly influence BMI and be involved in type 2 diabetes. Furthermore, if loci are found in the VC analysis but not in the ASP analysis, these loci will likely influence BMI in general, but possibly not be involved in type 2 diabetes. Vice versa, if

loci are found in the ASP analysis but not in the VC analysis, these BMI loci might be specifically involved in obese type 2 diabetes but would not influence BMI in general.

The VC analysis revealed two genomic regions showing VC LOD scores ≥ 1.0 (see Figure 1). On chromosome 1 a VC LOD score of 1.0 was obtained between markers D1S1678 and D1S549. Fine-mapping of this region with the addition of four extra markers increased the VC LOD from 1.0 to 1.5 between markers D1S1678 and D1S2141. For chromosome 11 a VC LOD score of 2.5 was obtained between markers D11S940 and D11S2000, while addition of four extra markers in this region slightly decreased the VC LOD from 2.5 to 2.3 between markers D11S1887 and D11S940 (see figure 3).

The ASP analysis revealed four genomic regions showing maximum LOD scores ≥ 1.0 (see Figure 2). On chromosome 1 a maximum LOD score of 1.0 was found near marker D1S549. Since the region found on chromosome 1 showed overlap with the region found in the VC analysis the same fine-map with additional markers was applied. Addition of four extra markers in this region decreased the maximum LOD from 1.0 to 0.7. On chromosome 11, a maximum LOD score of 1.7 was obtained between markers D11S2362 and ATA34E08; addition of three extra markers between these markers slightly decreased the maximum LOD from 1.7 to 1.5 between markers D11S2362 and D11S1999. For chromosome 12, a maximum LOD score of 1.9 was obtained near marker D12S1042; after addition of six extra markers the maximum LOD of 1.9 decreased to a maximum LOD score of 1.7 between D12S1207 and D12S398 (see figure 3). For chromosome 18, a maximum LOD score of 2.3 was obtained between markers D18S471 and D18S843 (van Tilburg *et al.* unpublished data). Chromosome 18 had already been fine-mapped to 5 cM or less, and no further genotyping was performed in this region.

Initially, the loci detected in both analyses on chromosome 1 appeared to overlap. However, after addition of extra markers the VC LOD increased from 1.0 to 1.5 between D1S1678 and D1S2141 in the whole dataset, whereas the maximum LOD score decreased from 1.0 to 0.7 in the 20% most obese pedigrees. The small peak in the 20% most obese pedigrees may represent its

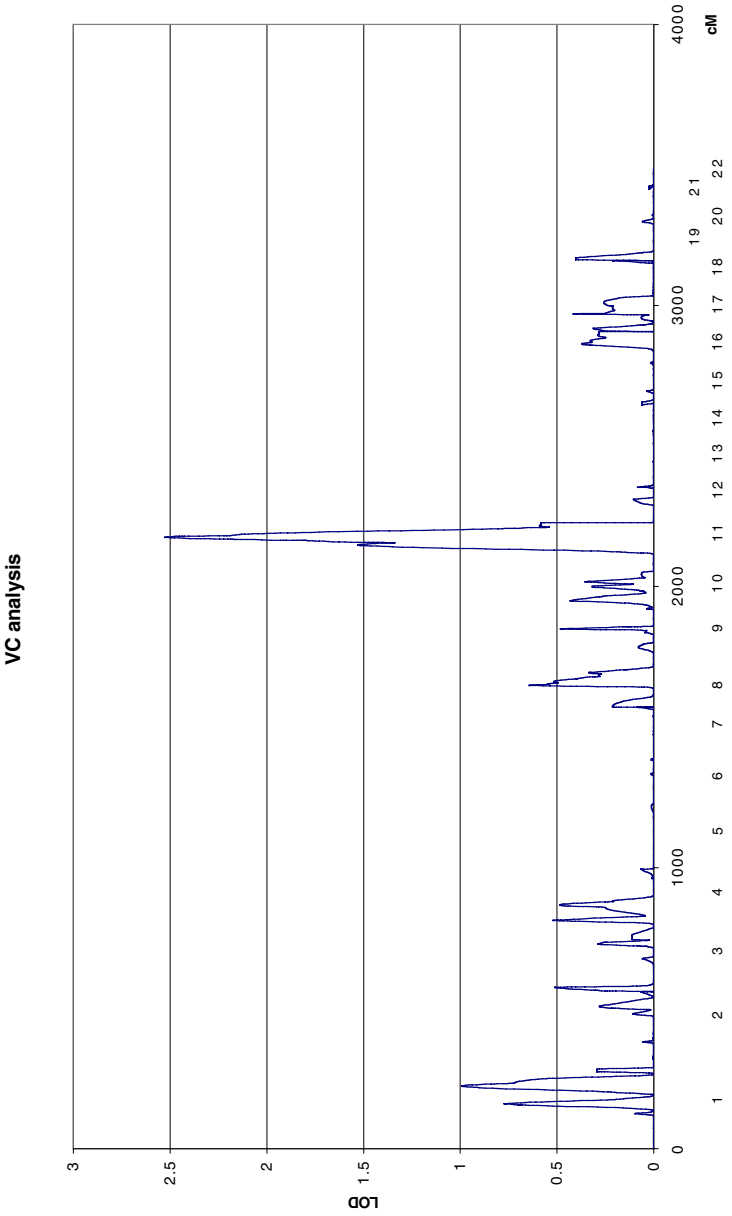


Figure 1. Multipoint variance-components analysis of BMI, with 325 autosomal microsatellite markers. Chromosome numbers on the X-axis are placed at the midpoint of the respective chromosomes; length of chromosomes adjusted according to the sex average map of the Marshfield genetic map (http://research.marshfieldclinic.org/genetics/).

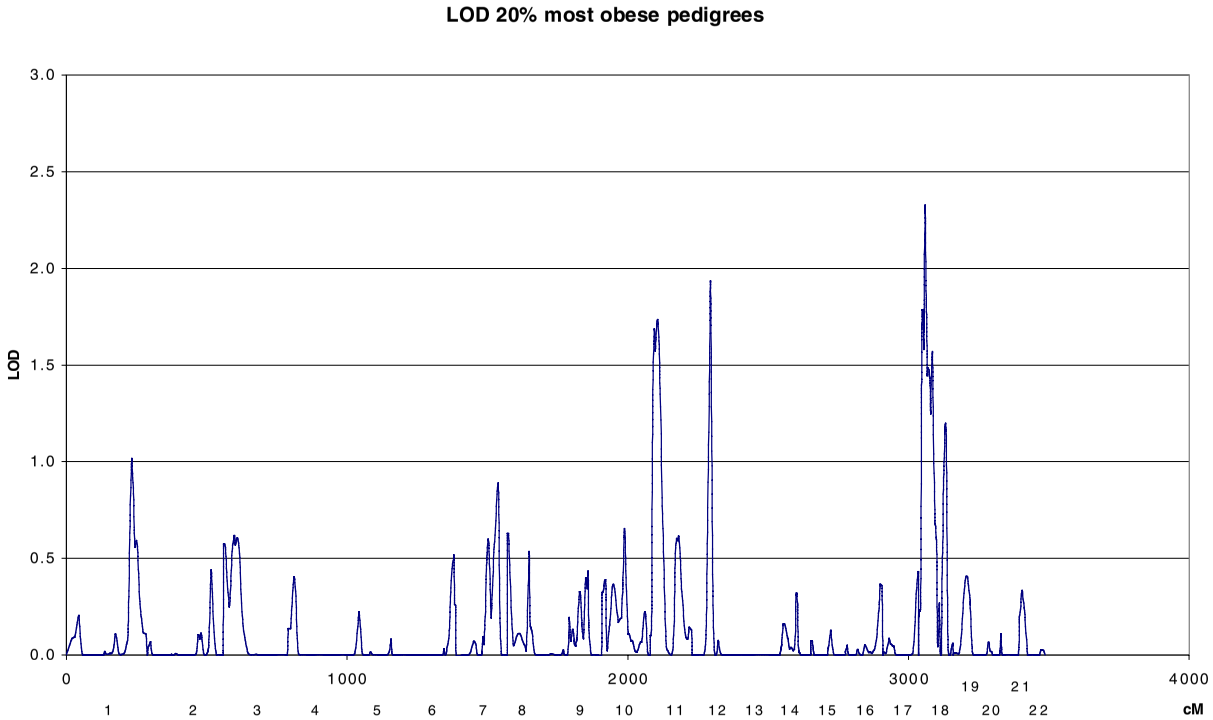


Figure 2. Multipoint non-parametric linkage analysis of the 20% most obese pedigrees, with 325 autosomal microsatellite markers. Chromosome numbers on the X-axis are placed at the midpoint of the respective chromosomes, length of chromosomes adjusted according to the sex average map of the Marshfield genetic map (http://research.marshfieldclinic.org/genetics/).

contribution to the total peak found in the whole dataset for BMI (figure 3). We therefore concluded that the evidence suggests that this locus possibly influences BMI and may also be involved in type 2 diabetes.

The QTL suggested by the linkage with BMI on the long arm of chromosome 1 (LOD =1.5) in our dataset has also been found in various other linkage

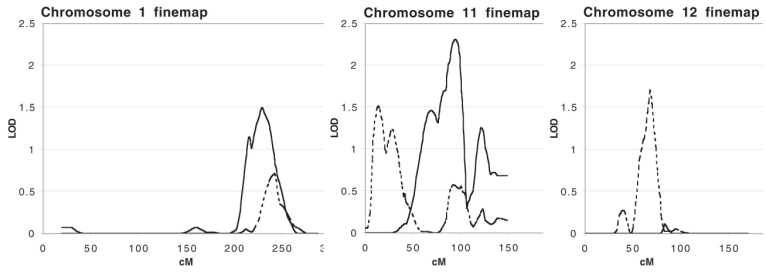


Figure 3. Fine-map results of chromosomes 1, 11 and 12. Solid lines represents the variance-components linkage analysis of BMI in type 2 diabetes patients. Dashed lines represent the non-parametric linkage analysis of the 20% most obese type 2 diabetes pedigrees.

studies on type 2 diabetes (Elbein *et al.* 1999; Hanson *et al.* 1998; Meigs *et al.* 2002; Wiltshire *et al.* 2001). The 95% confidence limits for our linkage estimate, generally assumed to be the maximum LOD value -1 (subsequently referred to in the text as LOD-1) was located in region 1q31-q42. Our findings and those of others (Elbein *et al.* 1999; Meigs *et al.* 2002; Wiltshire *et al.* 2001) focus particularly on the region between markers D1S518 and D1S179, in Caucasian subjects. This region contains CAPN2 and CAPN9, calcium-activated neutral proteases related to CAPN10, a candidate gene recently associated with type 2 diabetes in a Mexican-American population (Horikawa *et al.* 2000).

So far, no human QTL influencing obesity has been reported for this region on chromosome 1, although the syntenic region seems to be involved in various animal models for both obesity and diabetes (Brockmann *et al.* 1998; Moody *et al.* 1999; Taylor *et al.* 2001). The similarity between our finding

and these studies supports the hypothesis that a diabetes susceptibility locus is located in this region on chromosome 1q and that it may influence BMI as well.

It is of note that the loci found on chromosome 11 are over 100 cM apart. The region between markers D11S2362 and D11S1999 on the short arm of chromosome 11 (LOD =1.5; LOD-1 region 11p15) shows suggestive evidence for linkage to obese type 2 diabetes. This region harbors the insulin (INS) gene and its VNTR; it also harbors the sulphonylurea receptor 1 (SUR1) gene. Various association studies, although not conclusive, have found that the INS VNTR locus has been implicated in type 2 diabetes. More recent studies have established a significant association between the class III VNTR allele size and type 2 diabetes in Caucasian subjects in the UK (Huxtable *et al.* 2000; Ong *et al.* 1999). Thus, insulin deficiency in type 2 diabetes might depend on polymorphisms in the VNTR, affecting the expression of the insulin gene. SUR1 has been proposed as a candidate gene for type 2 diabetes, since it is a major determinant of normal glucose-induced insulin secretion in the beta-cell (t Hart *et al.* 2000), and a target for the sulfonylurea type medication. It has been shown (Hani *et al.* 1997) that an exon 18 variant of SUR1 was associated with morbid obesity and type 2 diabetes, although other sib-pair studies have failed to provide evidence for linkage in this region.

The suggestive evidence for linkage found on the long arm of chromosome 11 (LOD=2.3; LOD–1 region 11q14-q24) in our analysis for BMI in type 2 diabetes, was also found in a linkage study with Pima Indians (Hanson *et al.* 1998; Norman *et al.* 1997). Hence our finding further strengthens the indication that a locus influencing BMI in type 2 diabetes may reside on chromosome 11q. Nevertheless, to date no physiologically plausible candidate genes have been found that account for linkage to diabetes and BMI on chromosome region 11q23 (Baier *et al.* 2002).

The region found on chromosome 12 (LOD=1.7; LOD –1 region 12q12-q14) shows suggestive evidence for linkage in obese type 2 diabetes. This region has been shown to harbor the gene for vitamin D3 receptor (VDR); allelic variations in VDR were reported to modulate insulin secretion in re-

sponse to glucose (Ye *et al*. 2001). It was also found that polymorphisms in the VDR gene were associated with the susceptibility to obesity in subjects with early-onset type 2 diabetes (Ye *et al.* 2001). Subsequently, other studies also showed evidence for linkage to chromosome 12q for type 2 diabetes (Bektas *et al.* 1999; Mahtani *et al.* 1996). In combination with our data, this suggests that a susceptibility locus for type 2 diabetes in combination with obesity resides on chromosome 12q. So far, no physiologically plausible candidate gene has been proposed that might account for the described linkage to diabetes on chromosome region 12q.

In summary, our study to determine linkage of BMI in type 2 diabetes and linkage to high BMI in obese type 2 diabetes families confirmed previous findings on various chromosomes (summarized in table 1). Our findings support a susceptibility loci for type 2 diabetes residing on chromosome 1q, which may also influence BMI. We confirmed the finding of a locus on chromosome region 11q that influences BMI in type 2 diabetes. Additionally, we found evidence-suggesting linkage for type 2 diabetes on chromosome regions 11p, 12q (this report) and 18p11 (Van Tilburg *et al.* unpublished data), which confirm previous findings to the corresponding regions. However, at least in the Breda Cohort of type 2 diabetes patients, these regions are most likely to be influenced by obesity. Previous studies did not consider the role of BMI, with the exception of the 18p11 locus, where a similar BMI stratification was applied (Parker et al. 2001).

From the present data it cannot be inferred whether individual BMI loci are independently or cooperatively involved in determining diabetes status. Further studies in additional populations of obese patients as well as in type 2 diabetes patients will be necessary to provide a better insight into the interplay between loci and/or genes primarily associated with obesity and those primarily associated with type 2 diabetes. Ideally, two independent groups from the same population should be studied, one ascertained for obesity irrespective of diabetes status and, vice versa, one ascertained for type 2 diabetes independently of BMI.

Research design and methods

Subjects. The study group comprised 562 individuals from 178 families from the Breda Study Cohort (322 women and 240 men, of whom 235 and 185, respectively, were diagnosed as having type 2 diabetes). The level of obesity

Table 1. Summary of results.

Locus	Method	LOD	Previously described in literature
1q	VC-analysis/ ASP-analysis	1.5 0.7	Linkage studies in T2D (Elbein *et al.* 1999; Hanson *et al.* 1998; Meigs *et al.* 2002; Wiltshire *et al.* 2001)
11p	ASP-analysis	1.5	Implicated in candidate gene analysis for T2D (Hani *et al.* 1997; Huxtable *et al.* 2000; Ong *et al.* 1999)
11q	VC-analysis	2.3	Linkage studies in T2D (Hanson *et al.* 1998; Norman *et al.* 1997)
12q	ASP-analysis	1.7	Linkage studies in T2D (Bektas *et al.* 1999; Mahtani *et al.* 1996)
18p11*	ASP-analysis	2.3	Linkage study in T2D (Parker *et al.* 2001)

T2D = type 2 diabetes, * van Tilburg *et al.* unpublished data

in each individual was given by the body mass index (BMI), defined as weight (in kilograms) divided by height (in meters) squared. The relationship between BMI, age and gender was determined via multiple linear regression analysis. The raw BMI values were adjusted for age and gender in all family members according to the obtained regression coefficients, and normalized percentile values obtained following natural log transformation. For the ASP analysis family members were classified as affected only if they had both type 2 diabetes and a high BMI level (the highest 20% of the adjusted BMI distribution within the study group). The 20% most obese pedigree group comprised 30 families with 44 sib-pairs having type 2 diabetes (range adjusted BMI 28.5-43.5 kg/m2). Selection and ascertainment of the Breda Cohort has been reported elsewhere (http://humgen. med.uu.nl/research/diabetes/ BredaCohort.html).

Genotyping. A modified version of the Weber set 6 containing 325 markers from 22 autosomes with at an average spacing of 11 cM was used for the genome-wide screen. Complementary marker information can be found at: http://humgen.med.uu.nl/publications/jonathan2002_2/index.html. The

markers were analyzed as described by van Tilburg *et al.* (unpublished data). ***Statistical analysis.*** QTL analysis was performed using the multipoint variance-component (VC) method implemented in GENEHUNTER 2.0 software (Pratt *et al.* 2000), assuming an additive model and applying all-possible-sib-pairs (unweighted) analysis option. The VC method assumes that the expected genetic covariance between relatives for a trait is a function of the estimated proportion of alleles shared identically by descent (IBD) at a linked marker locus. The IBD probabilities were estimated using a multipoint approach that considers all available genotypes. The likelihood-ratio test was applied to test the null hypothesis of no additive genetic variance due to a QTL at a particular location. Linkage analysis in type 2 diabetes patients with high BMI was performed using MAPMAKER/SIBS software 2.0 package (Kruglyak and Lander 1995) and is described elsewhere (van Tilburg *et al*., unpublished data).

Acknowledgements

We thank the Dutch families for their support and willingness to participate in this study. We are also grateful to the Diabetes Service Breda and the nurses for collecting the blood samples, and to the general practitioners for encouraging their patients to participate. This work was financially supported by the Dutch Diabetes Research Foundation (DFN grant 97.114 to CW and TWvH). We would also like to thank Ms. Jackie Senior for improving the manuscript.

References

Baier L, Kovacs P, Wiedrich C, Cray K, Schemidt A, Shen GQ, Sutherland J, et al. (2002) Positional Cloning of an Obesity/Diabetes Susceptibility Gene(s) on Chromosome 11 in Pima Indians. Ann N Y Acad Sci 967:258-264.

Bektas A, Suprenant ME, Wogan LT, Plengvidhya N, Rich SS, Warram JH, Krolewski AS, et al. (1999) Evidence of a novel type 2 diabetes locus 50 cM centromeric to NIDDM2 on chromosome 12q. Diabetes 48:2246-51.

Brockmann GA, Haley CS, Renne U, Knott SA, Schwerin M (1998) Quantitative

trait loci affecting body weight and fatness from a mouse line selected for extreme high growth. Genetics 150:369-81.

DeFronzo RA, Ferrannini E (1991) Insulin resistance. A multifaceted syndrome responsible for NIDDM, obesity, hypertension, dyslipidemia, and atherosclerotic cardiovascular disease. Diabetes Care 14:173-94.

Elbein SC, Hoffman MD, Teng K, Leppert MF, Hasstedt SJ (1999) A genome-wide search for type 2 diabetes susceptibility genes in Utah Caucasians. Diabetes 48:1175-82.

Hani EH, Clement K, Velho G, Vionnet N, Hager J, Philippi A, Dina C, et al. (1997) Genetic studies of the sulfonylurea receptor gene locus in NIDDM and in morbid obesity among French Caucasians. Diabetes 46:688-94.

Hanson RL, Ehm MG, Pettitt DJ, Prochazka M, Thompson DB, Timberlake D, Foroud T, et al. (1998) An autosomal genomic scan for loci linked to type II diabetes mellitus and body-mass index in Pima Indians. Am J Hum Genet 63:1130-8.

Horikawa Y, Oda N, Cox NJ, Li X, Orho-Melander M, Hara M, Hinokio Y, et al. (2000) Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. Nat Genet 26:163-75.

Huxtable SJ, Saker PJ, Haddad L, Walker M, Frayling TM, Levy JC, Hitman GA, et al. (2000) Analysis of parent-offspring trios provides evidence for linkage and association between the insulin gene and type 2 diabetes mediated exclusively through paternally transmitted class III variable number tandem repeat alleles. Diabetes 49:126-30.

Kruglyak L, Lander ES (1995) Complete multipoint sib-pair analysis of qualitative and quantitative traits. Am J Hum Genet 57:439-54.

Mahtani MM, Widen E, Lehto M, Thomas J, McCarthy M, Brayer J, Bryant B, et al. (1996) Mapping of a gene for type 2 diabetes associated with an insulin secretion defect by a genome scan in Finnish families. Nat Genet 14:90-4.

Meigs JB, Panhuysen CI, Myers RH, Wilson PW, Cupples LA (2002) A genome-wide scan for loci linked to plasma levels of glucose and HbA(1c) in a community-based sample of Caucasian pedigrees: The Framingham Offspring Study. Diabetes 51:833-40.

Moody DE, Pomp D, Nielsen MK, Van Vleck LD (1999) Identification of quantitative trait loci influencing traits related to energy balance in selection and inbred lines of mice. Genetics 152:699-711.

Norman RA, Thompson DB, Foroud T, Garvey WT, Bennett PH, Bogardus C, Ravussin E (1997) Genomewide search for genes influencing percent body fat in Pima Indians: suggestive linkage at chromosome 11q21-q22. Pima Diabetes Gene Group. Am J Hum Genet 60:166-73.

Ong KK, Phillips DI, Fall C, Poulton J, Bennett ST, Golding J, Todd JA, et al. (1999) The insulin gene VNTR, type 2 diabetes and birth weight. Nat Genet 21:262-3.

Parker A, Meyer J, Lewitzky S, Rennich JS, Chan G, Thomas JD, Orho-Melander M, et al. (2001) A gene conferring susceptibility to type 2 diabetes in conjunction with obesity is located on chromosome 18p11. Diabetes 50:675-80.

Perusse L, Bouchard C (1999) Genotype-environment interaction in human obesity. Nutr Rev 57:S31-7; discussion S37-8.

Pratt SC, Daly MJ, Kruglyak L (2000) Exact multipoint quantitative-trait linkage analysis in pedigrees by variance components. Am J Hum Genet 66:1153-7.

Rankinen T, Perusse L, Weisnagel SJ, Snyder EE, Chagnon YC, Bouchard C (2002) The human obesity gene map: the 2001 update. Obes Res 10:196-243.

t Hart LM, Dekker JM, van Haeften TW, Ruige JB, Stehouwer CD, Erkelens DW, Heine RJ, et al. (2000) Reduced second phase insulin secretion in carriers of a sulphonylurea receptor gene variant associating with Type II diabetes mellitus [In Process Citation]. Diabetologia 43:515-9.

Taylor BA, Wnek C, Schroeder D, Phillips SJ (2001) Multiple obesity QTLs identified in an intercross between the NZO (New Zealand obese) and the SM (small) mouse strains. Mamm Genome 12:95-103.

Valsania P, Micossi P (1994) Genetic epidemiology of non-insulin-dependent diabetes. Diabetes Metab Rev 10:385-405.

Wiltshire S, Hattersley AT, Hitman GA, Walker M, Levy JC, Sampson M, O'Rahilly S, et al. (2001) A genomewide scan for loci predisposing to type 2 diabetes in a U.K. population (the Diabetes UK Warren 2 Repository): analysis of 573 pedigrees provides independent replication of a susceptibility locus on chromosome 1q. Am J Hum Genet 69:553-69.

Ye WZ, Reis AF, Dubois-Laforgue D, Bellanne-Chantelot C, Timsit J, Velho G (2001) Vitamin D receptor gene polymorphisms are associated with obesity in type 2 diabetic subjects with early age of onset. Eur J Endocrinol 145:181-6.

Chapter 5

Candidate gene analysis
Part I

The exon 16-3t variant of the sulphonylurea receptor gene is not a risk factor for type II diabetes mellitus in the Dutch Breda cohort

J.H.O. van Tilburg[1], L.B. Rozeman[1], H. van Someren[1], C.A.E. Rigters-Aris[1], J.P. Freriks[4], P.L. Pearson[1], L.A. Sandkuijl[1,3], T.W. van Haeften[2], and C. Wijmenga[1]

[1]Department of Medical Genetics, and [2]Department of Internal Medicine, University Medical Center, Utrecht, the Netherlands. [3]Departments of Medical Statistics and of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands. [4]Diabetes Service Breda, Breda, the Netherlands.

Diabetologia (2000) 43;681-684 reprinted with permission from Springer-Verlag

Introduction

Genes involved in insulin signalling and insulin response are excellent candidate genes for type 2 diabetes mellitus (T2D). Generally, genes are chosen on the basis of the existing knowledge of glucose homeostasis, and often dysfunction of the gene would be predicted to lead to either insulin resistance or impaired insulin secretion or both. The approach would then consist of finding a significant association between diabetes mellitus and a functional polymorphism in the proposed candidate gene. Alternatively, a finding could consist of an association between a polymorphism and a dysfunction in one (or more) of the mechanisms involved in glucose homeostasis (e.g. insulin sensitivity or secretion). In general, this is achieved by comparing a significant number of unrelated T2D patients with a control group. To date, over 250 candidate genes have been studied for their role in T2D (DeFronzo 1997). The majority of these studies have failed to uncover any association. However during recent years it has also become clear that this method may be powerful especially if it is used to uncover the impact of rather common variants that exert relatively small effects on the phenotype.

Studies involving polymorphisms of the insulin receptor, insulin receptor substrates, the glucagon receptor, and the sulfonylurea receptor (SUR), which is a constituent of the K_{ATP} channel of the pancreas β-cell (John *et al.* 1998), have been somewhat more successful.

The sulphonylurea receptor 1 (SUR1) has been proposed as a candidate gene for type 2 diabetes mellitus. The SUR1 is a major determinant of normal glucose induced insulin secretion in the beta-cell, and is target for the sulfonylurea type medication. Previous studies showed that in some populations an association could be found between a single nucleotide polymorphism (SNP) of exon 16 (SNP16) of SUR1 with type 2 diabetes mellitus. 't Hart *et al.* (1999) reported an association of the SNP16-3t variant with T2D, showing that the genotype frequencies between controls and Dutch Caucasian T2D patients, from two different cohorts (Rotterdam and Hoorn), differed significantly from controls ($p<0.05$). This effect was even stronger when allele frequencies of the t-variant, rather than genotype frequencies, were compared

between T2D patients (0.42) and controls (0.48) (p=0.01).

Materials and methods

The patient materials used for the candidate gene analysis are described in chapter 2 of this thesis. The DNA was extracted from 10 ml of blood using standard procedures (Miller *et al.* 1988). The described variant SNP16-3t of SUR1 was examined by a PCR-RFLP based method as described by Hansen *et al.* (1998). All 542 DNA samples were analysed twice, in independent assays. The correctness of the alleles was confirmed by sequencing of 9 samples, three samples, one from each genotype group, were used as control-standards in every assay. Apart from the 542 T2D patients, 150 control subjects were included in the analysis of the SUR1 polymorphism. Examples of the assay are depicted in figure 1.

Statistical analysis

Differences between groups were compared with a χ^2 test with two degrees of freedom, for both the genotypes and the allelic frequencies. A p-value < 0.05 would be considered to be significant. As a measure of the relative risk of the different alleles the Odds ratio (OR) was calculated, together with the 95% confidence interval (95% CI). The SNP16 polymorphism was analyzed using 542 type 2 diabetes patients from the Breda Study Cohort. The characteristics of this cohort are described in chapter 2 of this thesis. To determine the normal frequencies of the variations in the Dutch general population 150 control subjects were screened. Due to PCR failure of certain samples or samples only scored in one of the in independent assays, only 523 T2D patients of the 542 T2D patients were fully typed. The genotype and allele frequencies for SNP16-3t of the 523 T2D patients are presented in table 1, next to the results of 't Hart *et al.* (1999). No association was observed between the t-variant of SNP16 and T2D in our cohort of patients (p=0.10).

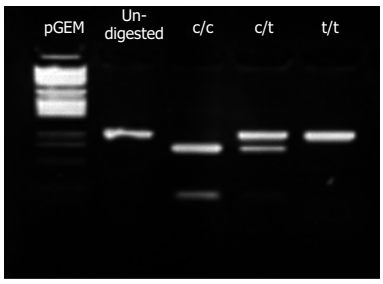


Figure 1. PCR-RFLP analysis of the SUR1 exon 16-3t polymorphism, PCR products were digested with PstI. c/c = wild type, c/t = heterozygotes, t/t = mutant type

Discussion

This study was undertaken in order to investigate whether an association can be found between a relatively common polymorphism of the sulphonylurea receptor 1 (SUR1) and type 2 diabetes mellitus. The studies indicate that in our cohort, no association could be found between the polymorphism and type 2 diabetes mellitus. The frequency of SNP16-3t of the SUR1 did not differ significantly (p=0.42) between our control population and the control populations used by 't Hart *et al.* (frequencies 0.46 and 0.41, respectively). Whereas this SNP16-3t variant was significantly increased (p=0.01) in cases compared to controls in the studies of 't Hart *et al.*'s study, such an increased frequency was not observed in our Breda study; if anything a slight decrease was found (p=0.10).

There are several possible explanations for this difference. Firstly, the patients involved in these studies might differ from patients in other studies. The t-variant of the SUR1 gene does not create irregular splice variants (Inoue *et al.* 1996) but may be associated with functional changes in the ATP-sensitive K^+ channel in insulin-secreting pancreatic β-cells. Since the oral agents mostly used for type 2 diabetes (the socalled sulphonylurea derivatives) act via this SUR1, one might speculate that carriers of the polymorphism would have worse metabolic control than other diabetic subjects without the polymorphism. Hence, carriers would be more prone to be insulin users. In our Breda cohort, which is a general practitioner driven cohort, patients are mainly managed with diet or SU medication, and only 9.4% of our patients receive

Table 1. SUR1 exon 16 variants in Dutch type II diabetes patients both in the currently described Breda cohort and in the Rotterdam and Hoorn cohort of 't Hart *et al.*

	Breda Study		't Hart *et al.* (1999)		Rotterdam cohort[#]		Hoorn cohort[#]	
Genotype	Controls n=150	Patients n=523[¶]	Controls n=336	Patients n=388	Controls n=170	Patients n=196	Controls n=166	Patients n=192
-3 c/c	0.29	0.36	0.33	0.25	0.31	0.22	0.35	0.28
-3 c/t	0.50	0.46	0.51	0.54	0.53	0.57	0.49	0.51
-3 t/t	0.21	0.17	0.16	0.21	0.16	0.21	0.16	0.21
χ^2	2.84		6.81		3.48		3.42	
p-value	0.24		0.03		0.18		0.18	
Allele								
-3c	0.54	0.60	0.59	0.52	0.57	0.51	0.60	0.53
-3t	0.46	0.40	0.41	0.48	0.43	0.49	0.40	0.47
χ^2	2.72		6.24		2.90		3.36	
p-value	0.10		0.01		0.09		0.07	
OR t-variant	0.80		1.30		1.29		1.32	
95% CI-OR-t-variant	0.62-1.04		1.06-1.61		0.96-1.73		0.98-1.78	

[#] The different cohorts from Rotterdam and Hoorn used by 't Hart *et al.* (1999).

[¶] Due to PCR failure of certain samples or samples only scored in one of the independent assays, 523 T2D patients were fully typed.

insulin. It cannot be excluded that the cohort used by 't Hart *et al.* comprised more patients using insulin, giving rise to a higher presence of the t variant in patients than in controls. Assuming that the t-variant of the SUR1 gene is associated with functional changes in the ATP-sensitive K^+ channel in insulin-secreting pancreatic β-cells, functional studies could provide further insight into the effect of this t-variant on insulin-secreting pancreatic b-cells. Indeed, in recent studies 't Hart *et al.* showed that the carriers of the polymorphism showed less insulin secretion during a standard hyperglycemic clamp than controls (t Hart *et al.* 2000).

Secondly, the use of different populations may possibly make a difference. 't Hart *et al.* sampled from two different populations, namely Rotterdam and Hoorn. The difference between our results and those of 't Hart *et al.* resides almost entirely in the Rotterdam patient group, with no contribution from the Hoorn cohort. Amongst the patients from the Rotterdam cohort there is an increased frequency of the c/t genotypes (0.57), as compared to the frequency in patients from the Hoorn cohort and the Breda cohort (0.51 and 0.46 respectively).

Another possible explanation may lie in the stratification or population admixture. Positive association can also arise as an artefact of population admixture. This problem has afflicted many association studies performed in inhomogeneous populations ranging from the population of metropolitan Los Angeles to Native American tribes (Lander and Schork 1994; Pritchard and Rosenberg 1999). Rotterdam is home to the world's largest port and has a mixed population arising from its long-standing migrant and immigrant populations.

A combination of all three points may explain the difference between the results reported by 't Hart et al. and our results from the Breda cohort.

In conclusion, it appears that the SUR1 exon 16 t-allele is not a risk factor for developing type II diabetes in the Breda cohort of 523 patients with an apparently low insulin use. It will be of great interest to see if this t-allele is associated with type II diabetes mellitus and insulin use.

Acknowledgements

This study has been made possible by a grant from the Dutch Foundation for Diabetes Research (DFN97.114) to TWvH and CW.

References

DeFronzo RA (1997) *Pathogenesis of type 2 diabetes: metabolic and molecular implications for identifying diabetes genes*. Diabetes Reviews 5:177-269.

Hansen T, Echwald SM, Hansen L, Moller AM, Almind K, Clausen JO, Urhammer SA, et al (1998) *Decreased tolbutamide-stimulated insulin secretion in healthy subjects with sequence variants in the high-affinity sulfonylurea receptor gene*. Diabetes 47:598-605.

Inoue H, Ferrer J, Welling CM, Elbein SC, Hoffman M, Mayorga R, Warren-Perry M, et al (1996) *Sequence variants in the sulfonylurea receptor (SUR) gene are associated with NIDDM in Caucasians*. Diabetes 45:825-31.

John SA, Monck JR, Weiss JN, Ribalet B (1998) *The sulphonylurea receptor SUR1 regulates ATP-sensitive mouse Kir6.2 K+ channels linked to the green fluorescent protein in human embryonic kidney cells (HEK 293)*. J Physiol 510:333-45.

Lander ES, Schork NJ (1994) *Genetic dissection of complex traits*. Science 265:2037-48

Miller SA, Dykes DD, Polesky HF (1988) *A simple salting out procedure for extracting DNA from human nucleated cells*. Nucleic Acids Res 16:1215.

Pritchard JK, Rosenberg NA (1999) *Use of unlinked genetic markers to detect population stratification in association studies*. Am J Hum Genet 65:220-8.

t Hart LM, de Knijff P, Dekker JM, Stolk RP, Nijpels G, van der Does FE, Ruige JB, et al (1999) *Variants in the sulphonylurea receptor gene: association of the exon 16- 3t variant with Type II diabetes mellitus in Dutch Caucasians*. Diabetologia 42:617-20.

t Hart LM, Dekker JM, van Haeften TW, Ruige JB, Stehouwer CD, Erkelens DW, Heine RJ, et al (2000) *Reduced second phase insulin secretion in carriers of a sulphonylurea receptor gene variant associating with Type II diabetes mellitus*. Diabetologia 43:515-9.

Chapter 6

Candidate gene analysis
Part II

Relationship of Beta2-Adrenergic Receptor polymorphism with obesity in type 2 diabetes mellitus

Jonathan H van Tilburg,[1] Cisca Wijmenga,[1] Leida Roseman,[1] Eric Strengman,[1] Peter L Pearson[1] and Timon W van Haeften[2]

[1]Department of Medical Genetics,
[2]Department of Internal Medicine,
University Medical Center, Utrecht, the Netherlands.

In Press, Diabetes Care (2002)

Abstract

OBJECTIVE: To investigate the possible association of the Glu27 variant of the beta2-adrenergic receptor (B2ADR) and obesity in type 2 diabetes mellitus subjects from a Dutch cohort; to investigate the influence of this polymorphism on metabolic control and on plasma lipids in these type 2 diabetes subjects.

RESEARCH DESIGN AND METHODS: The Glu27Gln polymorphism (SNP27) of B2ADR was examined by a PCR-RLFP based method in 542 DNA samples of type 2 diabetes mellitus patients. Body Mass Indexes, HbA1c, and plasma cholesterol, HDL-cholesterol, and triglyceride levels were compared between the three genotype groups using ANOVA with age and gender as covariates.

RESULTS: ANOVA of the three genotype groups showed no statistically significant differences in BMI, age at diagnosis of diabetes, HbA1c, or in plasma lipids (all $p>0.10$). After exclusion of the insulin treated subjects, once more the differences were not statistically significant. Assessment of women and men separately did not alter the results.

CONCLUSION: The Glu27Gln polymorphism has not important effect on BMI in type 2 diabetes mellitus, neither in men nor in women. It has also no appreciable effect on HbA1c, or plasma lipids.

Since type 2 diabetes mellitus is strongly associated with obesity, putative obesity related genes are good candidate genes for type 2 diabetes.

Adrenergic receptor genes are candidate genes for obesity because they regulate lipid mobilization, energy expenditure and are involved in glycogen breakdown. It has previously been demonstrated that the beta2-adrenergic receptor (B2ADR: OMIM 109690) is involved in the transmission of adrenergic stimuli in the vasculature and in bronchioli (Barbe *et al.* 1996; Szefler *et al.* 1991). In adipose tissue, this receptor is notably involved in the adrenergic signal leading to increase in lipolysis, and has therefore been suggested to play a role in obesity (Barbe *et al.* 1996). A small number of studies have indicated an association between a single nucleotide polymorphism (SNP) in

codon 27 (SNP27) of B2ADR with obesity in non-diabetic (Swedish) subjects (Ehrenborg *et al.* 2000; Large *et al.* 1997). In a group women (Large *et al.* 1997), but not in men (Hellstrom *et al.* 1999), the Gln27 variant of SNP27 was markedly associated with obesity, with homozygotes having an average fat mass excess of 20 kg. Since the beta-2 adrenergic receptor is involved in lipolysis, various authors have addressed whether the polymorphism has a relationship with plasma lipid levels. Ukkola *et al*. (2001) found an influence of the variant on plasma total cholesterol in non-diabetic subjects. Carlsson *et al*. (2001) reported an association of the Gln variant with plasma Non-Esterified Fatty Acids (NEFA) in type 2 diabetes. In apparent contrast with this, others have reported an association of the Glu variant with plasma triglycerides (Rosmond *et al.* 2000), and cholesterol (Ehrenborg *et al.* 2000) in non-diabetic subjects.

We performed a study in a large group of type 2 diabetes subjects, to assess the possible relationship of the SNP27 of B2ADR with obesity, and with plasma lipids, in our Dutch Breda Cohort of type 2 diabetes mellitus patients. Detailed knowledge about this possible association may not only have an impact on our appreciation of the development of type 2 diabetes mellitus, but may potentially also have implications for its management.

Research design and methods

Subjects with type 2 diabetes mellitus were recruited in collaboration with their general practitioners and the Diabetes Service of the city of Breda, which is the only regional laboratory service for the western part of the North Brabant county in the Netherlands. Since 1990 the Diabetes Service Breda has collected clinical and biochemical data on more than 13,000 patients with type 2 diabetes mellitus. Initially, 4,000 possible participants were asked at random to take part in this study, if they had at least one sibling affected by type 2 diabetes. DNA was obtained from 542 type 2 diabetes mellitus patients, the majority of whom had at least one affected sibling; all were diagnosed with diabetes mellitus after the age of 35 years according to WHO criteria. Of each family only one subject took part in the studies. The Medical Ethics

Committee of the University Medical Center in Utrecht approved our study protocol and all the participants signed an informed consent. The participants' clinical parameters are shown in Table 1.

Table 1. Clinical characteristics of all participating subjects with type 2 diabetes mellitus (total of 542)

	Women (n=296)	Men (n=246)
Age (y)	70 ± 9	69 ± 9
Age at diagnosis (y)	61 ± 9	60 ± 9
Years of disease (y)[#]	10 ± 6	10 ± 6
Height (m)	1.63 ± 0.06	1.75 ± 0.07
Weight (kg)	75.5 ± 14.1	83.7 ± 12.2
BMI (kg/m^2)	28.5 ± 4.8	27.2 ± 3.4
HbA1c (%)	7.3 ± 1.1	7.3 ± 1.3
HDL cholesterol (mmol/l)	1.2 ± 0.3	1.1 ± 0.3
Total cholesterol (mmol/l)	5.5 ± 1.1	5.1 ± 1.0
Triglycerides (mmol/l)	1.9 ± 0.9	1.8 ± 1.0
Insulin use (yes/no/unknown)	30/264/2	22/222/2

Data is given as mean ± SD
[#]Duration in years of the disease from age at diagnosis

HbA1c and fasting plasma cholesterol, HDL cholesterol and triglycerides were obtained, and measured in one laboratory (Diabetes Service Breda).

The DNA was extracted from 10 ml of blood using standard procedures (Miller *et al.* 1988). The single nucleotide polymorphism 27 (SNP27), which leads to a Glu to Gln substitution, of the beta-2 adrenergic receptor (B2ADR) gene was examined by a PCR-RLFP based method as described by Large *et al.* (1997). However, the assay by Large *et al.* used the restriction enzyme ItaI whereas we used the restriction enzyme Fnu4H1, which had no influence on the assay because both enzymes recognise the same restriction site. All 542 DNA samples were analysed twice, in independent assays. The alleles were confirmed by sequence analysis of 20 samples; three of these (previously sequenced) samples, one from each genotype group, were used as control-standards in every assay.

Statistical analysis

ANOVA of BMI, age at diagnosis of diabetes, HbA1c, and plasma lipids was performed on all subjects with age and gender as covariates (table 2). Since insulin treatment is associated with weight gain, ANOVA was repeated after exclusion of the subjects who used insulin treatment (9.4%). Since the data of Large *et al.* (1997), and Hellström et al (1999) would suggest that this polymorphism may have a different effect in women as compared to men, the analyses were also performed for women and men separately.

If a population is in Hardy Weinberg equilibrium (HWE), the observed genotype frequencies will conform to $p^2 + 2pq + q^2=1$, where p^2 = freq (Gln/Gln), 2pq = freq (Gln/Glu), and q^2 = freq (Glu/Glu). To determine whether the population used is in HWE, the p and q values of the population (observed data) were compared with the expected genotype frequencies (if the population were in HWE) using a χ^2 goodness of fit test, with two degrees of freedom.

Table 2. Comparison of age, age at diagnosis of diabetes, BMI, HbA1c, and plasma lipids in 502 type 2 subjects subdivided according to the beta2-adrenergic receptor gene Glu27Gln polymorphism.

	Gln/Gln	Gln/Glu	Glu/Glu
Number (men/women)	109 (58/51)	225 (101/124)	168 (72/96)
Age (y)	70.5 ± 9.9	70.3 ± 9.0	70.0 ± 8.7
Age at onset (y)	60.6 ± 10.1	60.3 ± 9.1	60.6 ± 9.0
BMI (kg/m^2)	27.2 ± 3.8	27.9 ± 4.4	28.2 ± 4.0
HbA1c (%)	7.2 ± 1.4	7.4 ± 1.3	7.2 ± 1.2
HDL cholesterol (mmol/l)	1.2 ± 0.3	1.1 ± 0.3	1.2 ± 0.3
Total cholesterol (mmol/l)	5.3 ± 1.2	5.3 ± 1.0	5.3 ± 1.0
Triglycerides (mmol/l)	1.8 ± 1.1	1.9 ± 0.9	1.9 ± 0.9

Results

The SNP27 of B2ADR was investigated in 542 type 2 diabetes mellitus patients in order to study if there was an association between the polymorphism and obesity.

After exclusion of PCR failures, and of samples in which only scores in one of the two independent assays had been obtained, 502 type 2 diabetes mellitus patients were fully typed. Our study population was in Hardy-Weinberg equilibrium.

ANOVA of age of diagnosis of type 2 diabetes, BMI, HbA1c, and plasma lipids showed no statistically significant differences between the three genotype groups (with age and gender as covariates) (Table 3). After exclusion of the insulin treated subjects, once more the differences were not statistically significant.

Since the data of Large (1997) and Hellström (1999) would suggest that this polymorphism may have different effects regarding obesity in women as opposed to men, we also performed separate analyses of men and women (Table 3). However, again, no statistically significant differences were found.

Table 3. P-values for the assessment (ANOVA) of the influence of the beta2-adrenergic receptor gene Glu27Gln polymorphism on BMI, HbA1c, and plasma lipids in type 2 diabetes subjects. Data is given for all subjects and for the non-insulin-treated subjects separately.

	All subjects (including insulin use)			Subgroup (without insulin use)		
	Total	Men	Women	Total	Men	Women
Number	502	231	271	454	210	245
Age at diagnosis (y)	0.59	0.15	0.71	0.79	0.45	0.79
BMI (kg/m^2)	0.26	0.12	0.73	0.29	0.12	0.63
HbA1c (%)	0.27	0.38	0.59	0.18	0.38	0.45
HDL cholesterol (mmol/l)	0.40	0.16	0.34	0.60	0.16	0.47
Total cholesterol (mmol/l)	0.75	0.88	0.50	0.84	0.88	0.66
Triglycerides (mmol/l)	0.96	0.67	0.51	0.77	0.67	0.86

ANOVA was performed with age and/or gender as the covariate(s).

Conclusions

This study was undertaken to assess whether we can find an association between SNP27 of the B2ADR gene and obesity in a large group of type 2 diabetes mellitus. We also addressed the possible relationship of this polymorphism with age of diagnosis of diabetes, metabolic control, and plasma lipids.

BMI values between the three genotype groups were compared using ANOVA with age and gender as covariates. No statistically significant difference was observed between the various groups, this was also seen after exclusion of the insulin treated subjects. This implies that, at least in our cohort of subjects with type 2 diabetes mellitus, this polymorphism of the beta2-adrenergic receptor has no important effect on body mass index. We also found no effect of the polymorphism on age of diagnosis of diabetes, on metabolic control or plasma lipids.

So far, an association of the polymorphism has only been found in Swedish (non-diabetic) subjects (Ehrenborg *et al.* 2000; Hellstrom *et al.* 1999; Large *et al.* 1997), while others found no such association in Caucasians (Evans *et al.* 2001; Oberkofler *et al.* 2000; Rosmond *et al.* 2000). This polymorphism is less prevalent in Korean and Japanese ethnic groups than in Caucasians, and was not found to be associated with obesity in them (Hayakawa *et al.* 2000; Iwamoto *et al.* 2001; Kawamura *et al.* 2001; Kim *et al.* 2002).

Various explanations are possible for the differences found between the various studies. First, Large *et al.* (1997) and Hellström *et al.* (1999) studied healthy women and men in whom diseases such as diabetes mellitus, or hypertension were excluded. In contrast, our subjects were diagnosed with type 2 diabetes mellitus and were substantially older. It is therefore possible that the association found by Large *et al.* is specific for non-diabetic obese women. Whether some of these relatively young subjects (average age around 40 years) would later develop diabetes is, of course, uncertain. The female cohort from the Breda Study is (almost) twice the size of the one studied by Large *et al.* (1997). Hellström *et al.* (1999) found a protective association for the Glu27 variant of SNP27 in men; we were not able to confirm such an association in

the Dutch type 2 diabetes mellitus men. However, the men in the studies of Hellström *et al.* were (on average) 27 years younger and they were non-diabetic. Whether the protective effect of the Glu27 polymorphism as observed by Hellström *et al.* is an age-dependent effect, with other age-related effects overruling this potentially protective effect in our cohort, remains speculative.

We also addressed the possible relationship of the polymorphism with plasma lipids in our type 2 diabetes subjects, and found no such relationship. In a study involving 284 Swedish men, Rosmond (2000) found an association of the homozygous Glu27Glu genotype with elevations in plasma triglyceride, while Ehrenborg (2000) reported an association of the Glu variant with elevations in cholesterol and triglyceride levels in a study involving 180 healthy men. In the only study involving the relationship of plasma lipids with the polymorphism in siblings with type 2 diabetes, Carlsson *et al.* (2001) found no such relationship for total and HDL-cholesterol and triglycerides. However, in apparent contrast to the previous data involving the Glu variant (Ehrenborg *et al.* 2000; Rosmond *et al.* 2000), Carlsson observed a small effect of the Gln variant on plasma NEFA (Carlsson *et al.* 2001). We have not measured NEFA, but also found no effect of the polymorphism on the other plasma lipids.

In conclusion, the current data can only lead to the conclusion that they exclude large differences in BMI between carriers of the Glu27 variant as compared to the carriers of the Gln27 variant of the B2ADR gene in patients with type 2 diabetes mellitus. This polymorphism has also no appreciable effects on metabolic control or plasma total and HDL-cholesterol and triglycerides in our cohort.

Acknowledgements

We thank the families for their support and willingness to participate in this study. We are also grateful to the Diabetes Service Breda (The Netherlands) and the nurses for collecting the blood samples, and to the general practitioners for encouraging their patients to participate. This work was financially

supported by the Dutch Diabetes Research Foundation (Amersfoort, The Netherlands, DFN grant 97.114 to CW and TWvH). We would also like to thank Ms. Jackie Senior for improving the manuscript.

References

Barbe P, Millet L, Galitzky J, Lafontan M, Berlan M, *In situ assessment of the role of the beta 1-, beta 2- and beta 3- adrenoceptors in the control of lipolysis and nutritive blood flow in human subcutaneous adipose tissue.* Br J Pharmacol, 1996. 117(5): p. 907-13.

Carlsson M, Orho-Melander M, Hedenbro J, Groop LC, *Common variants in the beta2-(Gln27Glu) and beta3-(Trp64Arg)— adrenoceptor genes are associated with elevated serum NEFA concentrations and type II diabetes.* Diabetologia, 2001. 44(5): p. 629-36.

Ehrenborg E, Skogsberg J, Ruotolo G, Large V, Eriksson P, Arner P, Hamsten A, *The Q/E27 polymorphism in the beta2-adrenoceptor gene is associated with increased body weight and dyslipoproteinaemia involving triglyceride-rich lipoproteins.* J Intern Med, 2000. 247(6): p. 651-6.

Evans D, Wolf AM, Nellessen U, Ahle S, Kortner B, Kuhlmann HW, Beisiegel U, *Association between polymorphisms in candidate genes and morbid obesity.* Int J Obes Relat Metab Disord, 2001. 25 Suppl 1: p. S19-21.

Hayakawa T, Nagai Y, Kahara T, Yamashita H, Takamura T, Abe T, Nomura G, et al, *Gln27Glu and Arg16Gly polymorphisms of the beta2-adrenergic receptor gene are not associated with obesity in Japanese men.* Metabolism, 2000. 49(9): p. 1215-8.

Hellström L, Large V, Reynisdottir S, Wahrenberg H, Arner P, *The different effects of a Gln27Glu beta 2-adrenoceptor gene polymorphism on obesity in males and in females.* J Intern Med, 1999. 245(3): p. 253-9.

Iwamoto N, Ogawa Y, Kajihara S, Hisatomi A, Yasutake T, Yoshimura T, Mizuta T, et al, *Gln27Glu beta2-adrenergic receptor variant is associated with hypertriglyceridemia and the development of fatty liver.* Clin Chim Acta, 2001. 314(1-2): p. 85-91.

Kawamura T, Egusa G, Fujikawa R, Okubo M, *Gln27Glu variant of the beta2-adrenergic receptor gene is not associated with obesity and diabetes in Japanese-Americans.* Metabolism, 2001. 50(4): p. 443-6.

Kim SH, Kim DJ, Seo IA, Min YK, Lee MS, Kim KW, Lee MK, *Significance of beta2-adrenergic receptor gene polymorphism in obesity and type 2 diabetes mellitus in Korean subjects.* Metabolism, 2002. 51(7): p. 833-7.

Large V, Hellstrom L, Reynisdottir S, Lonnqvist F, Eriksson P, Lannfelt L, Arner P, *Human beta-2 adrenoceptor gene polymorphisms are highly frequent in obesity and associate with altered adipocyte beta-2 adrenoceptor function.* J Clin Invest, 1997. 100(12): p.

3005-13.

Miller SA, Dykes DD, Polesky HF, *A simple salting out procedure for extracting DNA from human nucleated cells.* Nucleic Acids Res, 1988. 16(3): p. 1215.

Oberkofler H, Esterbauer H, Hell E, Krempler F, Patsch W, *The Gln27Glu polymorphism in the beta2-adrenergic receptor gene is not associated with morbid obesity in Austrian women.* Int J Obes Relat Metab Disord, 2000. 24(3): p. 388-90.

Rosmond R, Ukkola O, Chagnon M, Bouchard C, Bjorntorp P, *Polymorphisms of the beta2-adrenergic receptor gene (ADRB2) in relation to cardiovascular risk factors in men.* J Intern Med, 2000. 248(3): p. 239-44.

Szefler SJ, Ando R, Cicutto LC, Surs W, Hill MR, Martin RJ, *Plasma histamine, epinephrine, cortisol, and leukocyte beta-adrenergic receptors in nocturnal asthma.* Clin Pharmacol Ther, 1991. 49(1): p. 59-68.

Ukkola O, Perusse L, Weisnagel SJ, Bergeron J, Despres JP, Rao DC, Bouchard C, *Interactions among the glucocorticoid receptor, lipoprotein lipase, and adrenergic receptor genes and plasma insulin and lipid levels in the Quebec Family Study.* Metabolism, 2001. 50(2): p. 246-52.

Chapter 7

General Discussion

Introduction

Type 2 diabetes mellitus or non-insulin-dependent diabetes mellitus (NIDDM) accounts for approximately 90% of all diabetes mellitus cases worldwide and arises from three separate causes. These are: resistance of insulin action on glucose uptake in peripheral tissue, particularly skeletal muscle and adipose tissue; impairment of insulin action to inhibit hepatic glucose production, and dysregulation of insulin secretion (DeFronzo 1997).

Type 2 diabetes mellitus is a very common disease, rising to an incidence of more than 15% in persons older than 65 years; it has serious complications and reduces life expectancy by an average of 8-10 years. It is rapidly becoming one of the major diseases within the European Union (EU), while over 150 million people are affected worldwide. It has been estimated that the overall prevalence will rise to 40% by 2010 (McCarthy and Zimmet 1994). Type 2 diabetes mellitus is difficult to treat and expensive to manage. Patients have a high risk of developing a range of complications leading to disability and premature death. These complications are the main cause of end stage renal disease, blindness, and lower limb amputation in the elderly.

Type 2 diabetes mellitus appears to be a disease with a complex inheritance patterns and is considered a multifactorial disease due to the interplay of genetic factors and external factors (DeFronzo 1997; Kahn *et al.* 1996). The genetic factors raise the risk approximately 3.5 times above the general population risk for first-degree relatives within families in which one or more type 2 diabetes mellitus patients are already present. As a consequence, the disease frequency within families is much lower than for fully penetrant Mendelian disorders so that large pedigrees segregating the disease are hard to find. Furthermore, the current rapid changes in disease prevalence worldwide cannot be due to changes in genetic predisposition but must be environmental in origin. These changes may reflect an extant genetic predisposition being challenged by changing or new environmental and life-style factors, the most important of which are diet and exercise. However, demographic factors such as better access to health care and an aging population are also involved in the disease aetiology. Nonetheless, the strongest risk factor is the ethnic back-

ground of a population (e.g. Pima Indians have a 50% risk of developing type 2 diabetes mellitus) (for excellent reviews see Marx 2002; Zimmet *et al.* 2001; Zimmet 1999).

An elucidation of the molecular background of the pathology of type 2 diabetes mellitus using a genetic approach will help us to focus on the underlying causes of the disease, and may also provide new insights for modifying diagnostic treatment and improving prevention.

Common variants and rare variants

In contrast to Mendelian traits, the expected patterns of genetic variation in the genes underlying complex traits are far more blurred. Although a vast amount of literature is available on various genetic models for complex traits, relatively little is known about the specific genetic variants that underlie these traits (Cargill *et al.* 1999; Chakravarti 1999; Lander 1996). Two major models make opposing predictions about the nature of genetic variation underlying complex traits, namely the *"common disease–rare variant"* (CD/RV) model versus the *"common disease–common variant"* (CD/CV) model. Common and rare variants are expected to be detected at different rates. If most variation is principally maintained by recent mutations, then they are likely to be rare, comparable with Mendelian traits, but there may be many of them. If most of the genetic variation is principally maintained by some form of balancing selection, each allele contributing to that variation is likely to be common but the actual number of these alleles is likely to be relatively small.

The CD/RV model predicts that phenotypic variation in complex traits will be caused by numerous, individually rare, genetic variants at multiple loci (Collins *et al.* 1997; Lander 1996; Pritchard 2001). Most populations, however, will harbour many distinct genetic variants, and these variants taken together may have a total frequency that is substantial. Although because any individual variant is rare, these variants are unlikely to be shared among subpopulations. Risch (2000) stated that rare alleles (<5% frequency) are most likely to be population-specific, and that common alleles (>10% frequency) are more likely to be found globally.

The CD/CV model predicts that the number of variant sites will be few at any particular locus, but have relatively common allele frequencies in the overall population and be shared across multiple subpopulations (Cargill *et al.* 1999; Chakravarti 1999; Lander 1996).

In practice neither of these extremes is likely to hold sway and a mixture of the two is the more likely scenario, with the rare alleles defining differences between populations and the common alleles resulting in a general discrete genetic disposition. Extensive discussion of the accuracy of these models must await the definition of enough rare and common disease-causing alleles to be able to assess the magnitude of their relative contribution.

It is recognised that the chances of detecting a significant association (linkage) between a variant (allele) and the disease is much greater if the allele is rare, particularly in situations where the allele frequency approximates the disease prevalence and few "healthy" persons carry the allele unlike the majority of the diseased persons. In contrast, although some common alleles may make a greater individual genetic contribution than rare alleles, this is probably more difficult to demonstrate due to the large number of "healthy" persons carrying the allele. In addition, both rare and common genetic variants are probably also affected by environmental risk factors.

What are the implications of these two different models in the search for genes in type 2 diabetes mellitus? As described in **chapter 1**, two common genetic approaches can be used to find the genetic variants (alleles) involved in the aetiology of type 2 diabetes mellitus, namely association studies between the disease phenotype and sequence variants in defined candidate genes and a genome-wide scan using highly polymorphic markers to identify chromosome regions harbouring disease-risk genes. In the CD/RV model, it is unlikely that association studies can be used, because they would probably miss most of the variants involved (see section: *The correct variant*). However, using a linkage approach (e.g. a genome-wide scan, see *section: Genome-wide scan*) it should be possible to identify the genes involved, even if the variants are rare. A consequence of the latter approach is that the population under study needs to be homogenous and preferably isolated (e.g. described by Vaessen

(2001)), so that the rare variants are more common in the selected population. On the other hand, the CD/CV model opens the door to genome-wide association studies for type 2 diabetes mellitus and other "common" complex diseases.

Association studies

Advances in our understanding of the physiology of nutrient regulation and of diabetes pathogenesis are generating a constantly expanding list of candidate genes that play a potential role in the pathways involved in insulin control, glucose homeostasis, adipose tissue metabolism or in the development of the pancreatic β-cell. The screening of these genes for sequence variants that may associate with type 2 diabetes mellitus is an important component of diabetes research. The function of the majority of human genes (>70%) is still unknown, so that statistically it is much more likely that a gene of unknown function will prove to be the "correct" candidate (see *section: The correct variant*).

The sequence variants can be of various types. Firstly, specific changes in coding sequences give rise to various influences on the protein level. A premature stop can be introduced causing dysfunction of the protein, although this is very unlikely to be the cause in complex traits such as type 2 diabetes mellitus. It is more likely that the change is subtle, and involves a missense mutation giving rise to molecular changes such as incorrect folding of the protein and thereby altering the function. Secondly, specific changes in regulatory sequences give rise to modified expression levels of the protein that contribute to the disease phenotype. Thirdly, the changes in coding or adjacent non-coding sequences may be neutral in their contribution to disease phenotype but may be in linkage disequilibrium (LD) with a yet unrecognised functional variant. Generally, an association study involves comparing the allele frequency of a given variant between a random sample of unrelated type 2 diabetes mellitus patients with a matched control group. Statistically significant differences in allele frequency between control and patient groups may indicate a contribution of the given variant to the disease phenotype.

Susceptibility effects have been claimed for variants in some of the gene products involved in insulin secretion or insulin action, such as insulin receptor substrate-1 (IRS-1) (Almind *et al.* 1993; 1996; Porzio *et al.* 1999), the glucagon receptor (Hager *et al.* 1995; Hansen *et al.* 1996; Lok *et al.* 1994), the sulfonylurea receptor (SUR) (Hani *et al.* 1997; Inoue *et al.* 1996; t Hart *et al.* 1999), the peroxisome proliferator-activated receptor-γ (PPARγ) (Altshuler *et al.* 2000; Hegele *et al.* 2000) and the mitogen-activated protein kinase 8-interacting protein 1 (MAPKBIP1) (Waeber *et al.* 2000).

The role for these candidate genes seems to be limited to a small percentage of type 2 diabetes mellitus patients or to specific populations (So *et al.* 2000; Velho and Froguel 1997). This limited success may result from two situations, as discussed in the next section.

The correct variant

Following discovery of an association between a given variant and the disease, the major question to be addressed is whether it is the correct variant. The sequence variant may only present in one or a limited number of populations, even though the candidate gene concerned contributes to disease susceptibility in all populations (see *section: Common variants and rare variants*). So multiple sequence variants within the same gene are associated with the disease (allelic heterogeneity, see also *section: Complexity of type 2 diabetes mellitus*), and the correct variant in one population may therefore not be the correct variant in another population. Thus, it is also possible that the correct variant is in linkage disequilibrium (LD) with the variant under study.

Furthermore, most studies focus only on the coding region of the candidate gene, although the variant could be located in a regulatory gene sequence region. These regulatory sequences may be the location of the true variants concerned in type 2 diabetes mellitus. However, if there is strong LD between certain variants in the gene and the correct variant in the regulatory sequence, it should still be possible to find an association. On the other hand, it is known that LD may vary between populations (see reviews (Ardlie *et al.* 2002a; Pritchard and Przeworski 2001; Wright *et al.* 1999)), which may be due to

differences in recombination, to a different historical set of events giving rise to the LD in the first place, or to differences in mutation frequency.

The correct gene

The other problem limiting success of the candidate approach in type 2 diabetes mellitus is that the researchers have failed to study the "correct" candidates genes because they are involved in, as yet, unknown pathways. As mentioned above, the function of the majority of human genes is still unknown, and it is much more likely that an unknown gene is involved in type 2 diabetes mellitus than a known gene.

This may be true for the genes considered in **chapters 5 and 6**, for which no association was found between the candidate genes described and type 2 diabetes mellitus.

Replication of a reported association

A complication in the interpretation of candidate gene studies in type 2 diabetes mellitus has been the proliferation of small studies often resulting in isolated reports of positive associations that have proved difficult or impossible to replicate (Ardlie *et al.* 2002b; McCarthy and Froguel 2002). Explanations for this lack of reproducibility can be found in small sample size or incorrect assumptions about the underlying genetic architecture, leading to inappropriate subgroup analysis and multiple testing, and often the use of poorly matched control groups (as mentioned in **chapter 6**, e.g. the control group used by Large et al. (1997) and (Cardon and Bell 2001).

This lack of reproducibility has led various authors to suggest that future association studies with candidate genes should meet a minimum of five criteria for the study design. These criteria include: (1) a justifiable biological rationale; (2) appropriate selection and sampling of both cases and controls; (3) rigorous and well documented phenotyping and genotyping procedures; (4) large sample sizes; and (5) physiologically meaningful evidence supporting a functional role of the variant, by means of a transgenic study model that produces the phenotypic effect, or a functional assay for that particular candi-

date gene (Cardon and Bell 2001; Hegele 2002; Lander and Schork 1994; McCarthy and Froguel 2002; Tabor *et al.* 2002). However, the last criterion, involving the development of proper functional assays or the development of the correct animal models, will slow down the research drastically.

Still, despite these known limitations, it has been claimed that the power of association studies to detect a genetic contribution to a complex disease can be much greater than that of linkage studies if the appropriate candidate genes and gene variants are investigated (Risch 2000). This "correct" candidate can also be chosen on the basis of the positional location found in linkage studies. An even better approach would be to pick the correct candidate from more elegant and comprehensive assessments of biological candidacy, e.g. through expression profiling using microarrays involving significant up- and down-regulation of transcription. However, it has not yet been demonstrated whether this approach is feasible.

Genome-wide scan

The major drawback of the candidate gene approach is that it may not lead to the identification of entirely new genes or pathways involved in type 2 diabetes mellitus. In order to identify new genes for the disease, genome-wide scans using evenly spaced polymorphic markers need to be performed. Furthermore, if the phenotypic variation in type 2 diabetes mellitus is caused by numerous, individually rare, genetic variants at multiple loci, it will be impossible to find all these variants using the candidate gene approach.

The classic 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 individuals can be reliably used for linkage studies because of the reduced and age-dependent penetrance.

It is therefore not possible to use parametric linkage analysis and non-parametric analysis methods need to be applied, since these require no prior

knowledge of the mode of inheritance of the disease, the disease allele (gene) frequencies, or the disease penetrance (Lander and Schork 1994). Currently, the most commonly used non-parametric approach is that of the affected sibpair (ASP), which ideally requires pairs of affected siblings and parents (Holmans and Craddock 1997). The ASP approach is the only convenient method of analysis for type 2 diabetes mellitus because only one or two generations in a family with this disease are normally available.

The basis of the ASP analysis is that individuals concordant for a given genetic trait should show greater than expected concordance for marker alleles that are closely linked to the disease. The most frequently used measure of concordance for two siblings at a locus is the number of alleles they share *identical-by-descent* (IBD). This method is more extensively described in box 1 in **chapter 1.**

To date, various research groups have completed genome-wide scans for type 2 diabetes mellitus using affected sibpairs (Ehm *et al.* 2000; Ghosh *et al.* 1999; Ghosh *et al.* 2000; Hanis *et al.* 1996; Hegele *et al.* 1999; Ji *et al.* 1997; Watanabe *et al.* 2000; Zouali *et al.* 1997) or small multi-generational families (Duggirala *et al.* 1999; Elbein *et al.* 1999; Hanson *et al.* 1998; Ji *et al.* 1997; Mahtani *et al.* 1996). Combined with the findings from our Breda study, described in **chapters 3 and 4**, the conclusion from all these studies is that there is no common susceptibility locus for type 2 diabetes mellitus. The genes involved are presumed to make individually small contributions to the development of the disease aetiology (see below).

Replication of reported linkage

Due to the limited power of linkage analysis based on IBD (identity-by-descent) mapping, a genomic region linked to a complex disease is generally very broad (often 20-40 cM) and often only a suggestive statistical significance is achieved (i.e., $1 < LOD < 3.6$ (Lander and Kruglyak 1995)). Putative linkage results in this range of significance should be confirmed using a separate group of affected sibling pairs before being considered indicative of localisation of a disease susceptibility locus. The lack of replication with ge-

nome-wide scans may be one of the reasons why no common susceptibility locus has been found for type 2 diabetes mellitus.

There are various reasons why the majority of putative loci may prove hard to replicate: firstly, the ethnic heterogeneity, secondly, the complexity of the disease itself, and thirdly, differences in diagnostic criteria or ascertainment of the patients (use of subphenotypes).

Ethnic heterogeneity

It may be that disease-susceptibility genes are so numerous and their interaction so varied and context-dependent that there is a unique profile of disease alleles for each population (see *section: Common variants and rare variants)*. As a result, identification of disease genes in one population may be difficult to replicate in another population and false-positive findings will be hard to resolve if effects are population-specific. Furthermore, some loci may have a higher contribution to the disease aetiology in one particular population than others (Horikawa *et al.* 2000). This can be best illustrated by the findings of Helege et al. (Hegele 1999; Hegele *et al.* 1999), who showed that a single nucleotide polymorphism (SNP) in the hepatic nuclear factor 1 alpha (HNF-1-_) gene is associated with type 2 diabetes mellitus in the Oji-Cree Indians. This finding has not been replicated in other type 2 diabetes mellitus populations. This gene is also implicated in a relatively rare form of diabetes mellitus, maturity-onset diabetes mellitus of the young (MODY) type 3, which is characterised by monogenic, autosomal dominant segregation and early age of onset. Similarly, Horikawa et al. (2000) found a SNP associated with type 2 diabetes mellitus in the calpain 10 (CAPN10) gene in Mexican Americans, but various studies in other populations could not confirm this association (Elbein *et al.* 2002; Evans *et al.* 2001; Hegele *et al.* 2001).

Complexity of type 2 diabetes mellitus

Most genomic regions linked to type 2 diabetes mellitus are generally very broad (often 20-40 cM, see also **chapters 3 and** 4 for examples). These regions may contain hundreds of genes, each with its own variants (common

or rare), giving rise to a vast genetic human diversity. However the relevant human diversity is also quite limited, in that most candidate genes have only a handful of common variants in their coding regions, while the vast majority of alleles are exceedingly rare. Mutational diversity at each locus is high; each mutation is rare, having occurred in recent human history (no earlier than 2,000 years ago) (Chakravarti 1999; Lander 1996).

However, much less is known about the allelic spectrum for genes underlying common disorders such as diabetes or asthma (Reich and Lander 2001). It is suspected that the mutations that lead to a complex phenotype occur at multiple genes. There are a number of possible models of allelic architecture to be considered in complex diseases like type 2 diabetes mellitus (Terwilliger and Weiss 1998).

Model 1: allelic homogeneity

The simplest model for the allelic complexity of genetic disease is model 1; all disease-predisposing alleles at a given locus are identical-by-descent in the population, having been derived from a common ancestor. In this model, it is assumed that in a given gene there is one – and only one – disease-predisposing allele and this allele has an identical etiological effect in all individuals, whether related or not (Terwilliger and Weiss 1998). To employ this model for a complex disease such as type 2 diabetes mellitus, multiple genes must be involved. The model can be tested directly in an association study, particularly using single nucleotide polymorphism (SNP) analysis to look for these alleles (as previously described).

Model 2: allelic heterogeneity

A more representative model for type 2 diabetes would be model 2, a common gene where multiple unique, but functionally equivalent, alleles give rise to the disease. Furthermore, multiple modifier genes may also influence this common gene. This model can also be tested directly in an association study, particularly using single nucleotide polymorphism (SNP) analysis to look for linkage disequilibrium of the various alleles in such a gene. A more

complex and more realistic variation of model 2 is that various alleles in the gene might have different quantitative effects on the phenotype. Additionally, there may also be various variants in the modifier genes leading to variable expression of the common gene. However, if the latter is true, it will be more difficult to find association between the variants and the disease unless variants in regulatory or modifier gene sequences are also studied. To date no such detailed studies have been published. The use of an association study design in this model is therefore not advisable. On the other hand, using a linkage approach it might be possible to pick up at least the common gene.

Model 3: locus heterogeneity

An even more complex model, model 3, that is probably closer to the reality in type 2 diabetes mellitus, is the combined effect of a collection of alleles (variants) in a set of key genes, plus environmental factors, which together determine whether an individual will suffer from type 2 diabetes mellitus (Chakravarti 1999; Lander and Schork 1994; Terwilliger and Weiss 1998; Weiss 1998). The potential level of complexity for type 2 diabetes mellitus and other 'complex diseases' could be enormous. According to the model, the number of key genes could vary between a few, tens, or even hundreds (oligogenic or polygenic). This has implications on the success of finding loci in a genome-wide scan. The relative risk for type 2 diabetes mellitus is ~ 3.5 (λ_r, risk of a relative compared with the risk of the general population), which means that if tens or hundreds of genes are involved, the individual contribution of these genes will be very small and it may therefore be impossible to find these genes at all. On the other hand, the multiplicity of genes underlying the complex phenotype of type 2 diabetes mellitus allows genetic mapping of key genes with a relatively large contribution in a genome-wide scan (see *section: Genome-wide scan*).

Furthermore, it is suspected that the mutations that lead to a complex phenotype occur at multiple genes. It is also plausible that these genes have pervasive interactions with each other, that is, dominance and epistasis. The first is a genetic interaction between two alleles at a locus, such that the phe-

notype of heterozygotes deviates from the average of two homozygotes. The second, refers to any genetic interaction in which the combined phenotypic effect of two or more loci is less than (negative epistasis) or greater than (positive epistasis) the sum of effects at individual loci (Barton and Keightley 2002).

The mixture of genetic and environmental factors increases the complexity of type 2 diabetes mellitus even further. It will take a long time (if ever, if the contributions of the genes involved are very small) before all the possible variants are found, together with the various possible inter-actions. Linking all the data to a suitable model will be a daunting challenge.

Subphenotypes

A possibility of increasing the power of a genome-wide scan approach is the use of subphenotypes. As mentioned in **chapter 1**, type 2 diabetes mellitus often occurs together with obesity and hypertension, but each may have its own genetic origin.

It is often debated whether it is better to attempt to perform positional cloning of complex disease susceptibility genes using the total disease phenotype or a subphenotype such as insulin resistance or obesity. Analyses of subphenotypes are potentially advantageous because the study group will become more homogeneous and potentially fewer genetic determinants will be involved than in the full disease phenotype. (See figure 1).

Most observable variation between individuals in disease susceptibility is quantitative, with population variation often approximating a statistical normal distribution, such as body mass index (BMI). Subphenotypes are generally quantitative traits and may be more informative than the dichotomy of affected versus unaffected persons. However, in contrast to traits controlled by one or a few loci with large effects, variations in quantitative traits are caused by segregation at multiple quantitative trait loci (QTL) with individually small effects that are sensitive to the environment (model 3). For complex traits, such as type 2 diabetes mellitus, the relationship between genotype and phenotype is not simple, and QTL genotypes cannot be determined from segregation of phenotypes in defined crosses within human pedi-

grees (Mackay 2001).

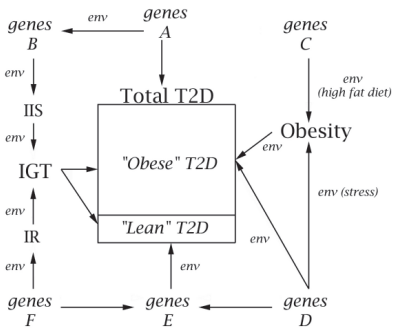


Figure 1. A possible schematic representation of the complexity of type 2 diabetes mellitus (T2D). Genes A may have a direct influence on the disease aetiology, or may together with environmental factors (env) influence other genes. Genes B may be involved in impaired insulin secretion (IIS), which in the end will lead to impaired glucose tolerance (IGT). IGT will in the end lead to T2D (obese driven or lean driven). This is similar for genes F, which may be involved in insulin resistance (IR). Genes C, together with environmental factors (such as a high fat diet) are involved in obesity, one of the highest risk factors for developing "obese" T2D. Other genes, under the influence of different environmental factors (stress) may also cause obesity, but may also directly have an impact on "obese" T2D. Genes E may have a direct involvement in'"lean" T2D. Furthermore, Genes D, E and F may have a combined effect (epistasis) on the disease aetiology.

Since the effects of individual QTLs are too small to be tracked by segregation in pedigrees, QTLs are mapped in a similar way as described above using the genome-wide scan approach. The principle of QTL mapping is simple and was noted 80 years ago by Sax (1923). If a QTL is linked to a marker locus, there will be a difference in mean values of the quantitative trait (e.g. BMI) among individuals with different genotypes at the marker locus. If the QTL and marker locus are unlinked, the mean value of the quantitative trait will be the same for each of the marker genotypes. The closer the QTL and marker locus, the larger the difference in trait phenotype between the marker genotypes. The marker in a local region exhibiting the greatest difference in the mean value of the trait is thus the one closest to the QTL (Mackay 2001).

Various subphenotypes (QTL) may influence type 2 diabetes mellitus (e.g. obesity, hyperinsulinemia, insulin resistance and fasting C-peptide/glucose). However, several factors make it difficult to estimate the true numbers and effects of loci that influence a quantitative trait. Closely linked QTLs with opposite effects tend to be missed. Similarly, closely linked QTLs with effects in the same direction tend to give the appearance of a single QTL of larger

effect. There is a lower limit for the phenotypic contribution of a QTL that can be detected, which will vary according to the size of the experiment (number of sib-pairs used) and the properties of the trait; real QTL with effects below this lower limit nearly always remain undetected (Barton and Keightley 2002).

It would be very interesting to map these multiple QTLs or various subgroups and see if they are linked and/or interact with each other. It is therefore necessary to study the genetics of QTLs in normal controls and in twin studies to find the right cut-off point for the QTLs under investigation (e.g. the variation of insulin resistance in normal individuals is unknown). Interacting QTLs are of particular interest as they indicate regions of the genome that might not otherwise be associated with the disease using a one-dimensional search. Although the concept of locating multiple, interacting QTLs is straightforward, implementation is quite difficult due to the huge number of potential QTLs and their interactions within these subgroups. This will lead to innumerable statistical models and heavy demand on computational facilities.

In **chapters 3 and 4**, it was observed that the Breda study cohort consisted of various arbitrarily defined subgroups of type 2 diabetes mellitus patients (a lean group and an obese group). One approach may therefore be to compare genome-wide scans of patients with obesity as well as type 2 diabetes mellitus with scans of patients having "only" one of these diseases, preferably in the same ethnic population (Parker *et al.* 2001; Perusse *et al.* 2001).

Study design

Often the initial study design does not take into account the various subphenotypes that can be analysed. This may indicate a possible disadvantage of the subphenotype approach. A clear subphenotype description is needed in which clear-cut parameters and standardised norms are used (non-affected versus affected). For example, some authors may define obesity as BMI > 27 kg/m^2, while others may use BMI >30 kg/m^2 as a threshold, whereas others may propose using measurements of visceral adiposity. A way to circumvent arbitrarily defined cut-offs is to perform a genome-wide scan in at least 1,000 healthy subjects and to search for the "true" cut-off.

Not incorporating a clear subphenotype in the study design often greatly limits the numbers of relative pairs that can be obtained for a genomic scan and linkage analyses for that particular subphenotype, and has consequences for the power and the homogeneity of the obtained subphenotype. Given the limited power of linkage analysis to localize genes for non-Mendelian traits, having fewer relative pairs for analyses is a major drawback (Bogardus *et al.* 2002). On the other hand, obtaining a more homogeneous group will also limit the variants involved (see *section: Common variants and rare variants*) and therefore increase the power to detect linkage.

Locus position

It has also become clear that when two or more studies find suggestive evidence for a particular region of the genome, there may be a large degree of variation in the specific position that gives maximum evidence for linkage. A priori, this variation might represent chance variation around a single genetic signal, the presence of multiple genetic signals, or one or more false-positive signals (Roberts *et al.* 1999).

It is suggested that the variability in maximum LOD position is substantial for loci of complex disorders, such as type 2 diabetes mellitus, with 95% confidence intervals covering tens of cM in samples consisting of relatively large numbers of families. Notably, most studies use sample sizes of less than 200 families (Roberts *et al.* 1999), similar to the Breda study cohort described in this thesis. This broad interval means that some studies which have claimed to detect a unique locus within a chromosome region have not done so, but they have detected the same locus as others, albeit with a different position for their maximum LOD.

Alternatively, multiple positive replications for closely linked regions might indicate that the findings are true and that the regions actually are multiple susceptibility loci close to each other that should be further explored.

Fine-mapping

What should be done if a putative linkage is replicated in additional studies? The first step is to narrow down the genomic region harbouring the putative susceptibility gene by means of fine mapping, which can be performed by saturating the region with additional markers. The maximum power in an affected sibpair approach (with no available parents) in type 2 diabetes mellitus was reached with an average marker spacing of 5 cM; adding more markers did not increase the power, because the maximum information was already extracted from the markers.

Another method for narrowing the interval in which the disease gene may lie is to use linkage disequilibrium (LD), or association testing, between genetic markers and the disease. If most affected individuals in a population share the same mutant allele at a causative locus, it is possible to narrow down the genetic interval around the disease locus by detecting disequilibrium between nearby markers and the disease locus (see *section: Common variants and rare variants*) (Jorde 1995; Lander and Schork 1994). However, the early applications of LD mapping were limited to rare diseases in a few favourable populations.

LD mapping is performed using single nucleotide polymorphisms (SNPs) since they are an abundant form of genome variation, occurring about every 1000 base pairs, and are mutably very stable (Gray *et al.* 2000; Wang *et al.* 1998). They are also mostly biallelic and thus easy to assay. More importantly, SNPs allow the unification of a candidate gene approach and association-based fine mapping to identify gene(s) of interest. They also aid in association of linkage analysis to the phenotypic and genotypic data giving rise to haplotype analysis. Automated genotyping would be a major advantage since localization of a susceptibility allele that has only a small phenotypic effect will require genotyping of a large number of SNPs in a large number of individuals (Brookes 1999; Lai 2001).

The SNP analysis strategy can potentially narrow a broad region of linkage (commonly 10-40 cM) to a physically small region of association (<1Mb) for intensive and thorough analysis for possible candidate genes. The final

step is to find functional variants within these genes to be associated with the disease to pinpoint the right candidate gene.

The general sibpair approach followed by LD mapping to identify candidate genes for a complex disease was successfully used to find the diabetes mellitus susceptibility gene CAPN10 (Horikawa *et al.* 2000; Wang *et al.* 1998) and more recently, to identify polymorphisms in the NOD2/CARD15 gene causing Crohn's disease (Hugot *et al.* 2001) and to identify polymorphisms in ADAM33, a putative asthma susceptibility gene (Van Eerdewegh *et al.* 2002).

SNPs at a frequency of less than 20% are of particular interest to researchers in complex diseases such as diabetes because it is presumed that the variant causing the disease should be at a frequency prevalent enough to represent the frequency of the disease in the population.

SNPs can now be obtained for most genomic regions from public databases (Sachidanandam *et al.* 2001) such as the SNP consortium (TSC). However, only a small percentage of the SNPs have been validated, that is, tested on a significant number of chromosomes in order to derive an accurate allele frequency for a given population. At this time the majority SNPs do not have valid population frequency data but this will improve soon we hope. For example, in January 2002, in a chromosomal region of 12.4 Mb, 5,714 unique SNPs were analysed of which only 129 (2%) had been validated (had confirmed frequency information in a population sample). The same chromosomal region analysed in July 2002 contained 7000+ SNPs of which 494 (~7%) had some frequency information. So for the time being, far fewer than 10% of the SNPs in the database have been validated (M. Erdos, The National Human Genome Research Institute, NIH, Bethesda, USA, personal communication).

Also, gaps remain in the human genome draft sequence, leaving some genomic regions with no known SNPs. To fill in these gaps with novel SNPs for a particular positional cloning project is very time-consuming and costly.

How many SNPs need to be genotyped to definitively identify a disease-associated locus is still under debate (Kruglyak 1999; Weiss and Terwilliger 2000). In some genes, there are multiple SNPs in varying degrees of linkage

disequilibrium with one another, while in other areas, a high degree of linkage disequilibrium extends over large genomic regions (Daly *et al.* 2001). Thus, it remains unclear exactly what SNP density is required to find a disease-associated SNP (or SNPs). LD maps need to be established to see how many SNPs are needed for a certain region (e.g. a large LD block needs fewer SNPs). Another important question is whether the patterns of LD, obtained by SNP analysis, found in one population will be replicated in other populations with a different population history (Gabriel *et al.* 2002; Reich *et al.* 2001; Zavattari *et al.* 2000).

However, to improve the rate of finding new susceptibility genes for complex diseases, the construction of a reliable SNP linkage map will be necessary. The volume of genotyping in LD studies will be enormous, particular if the number of cases is large. Pooling the DNA of the cases and controls and then estimating marker allele frequencies in each of the pooled samples could reduce the workload and costs (Barcellos *et al.* 1997). However, the pooling strategy does not allow haplotype analysis, so other techniques need to be applied to find all the possible haplotypes, such as large family studies or the use of somatic cell hybrids.

New and improved technologies, such as microarrays or mass spectroscopy that can type thousands of SNPs in a single assay will be of great importance. Easy-to-use bioinformatic tools will also need to be developed to process the vast amount of data generated by these types of studies.

Hopefully, combining these technologies, will narrow the list of potentially biologically relevant genes to a relatively small number of candidates involved in type 2 diabetes mellitus.

Future prospects

The semi-completion of the human genome project (HGP) has produced a vast amount of new information, which will lead to annotation of all human genes. The HGP will also circumscribe human variation and the extent and distribution of linkage disequilibrium. Other follow-ups to the HPG will include the initiation of functional studies on a genome-wide basis. These

will give insight into the function of all human genes, the possible interactions between various genes, and their involvement in various pathways associated with different diseases. Such studies will also provide insight into transcriptional regulation. It is not yet possible to efficiently integrate the various types of data, so new and more sophisticated computational tools will be required to achieve this. Such tools should permit incorporation of expression data from specific tissues with the possible interactions between genes and with the possible interactions between genes and the environment.

In addition, it is important that much more wide-scale collaboration should be established by sharing the available patient resources. This is particularly acute in the field of diabetes research. Sharing patient samples would reduce the cost of many studies because the collection and characterisation of patient material is time consuming and expensive. For example, in our own case, we were obliged to collect our own samples, despite the fact that the Dutch Diabetes Fund had previously financed the collection of diabetes samples for the Hoorn and Rotterdam studies. There may, of course, be compelling reasons to resample, such as inappropriate study design or not having DNA samples from sufficient family members to carry out identity-by-descent estimates. However, a critical evaluation should be made before yet another patient collection is initiated.

A useful model for tackling this problem would be to adopt the existing guidelines used by NIH, where patient collections are made available to other studies for the purposes of confirming results or analysing new aspects, so that resources are not wasted. This approach requires, of course, adjustment of the initial aims in setting up studies so that the study criteria are broadly established and samples can be used for several projects rather than only one. Such an enlightened policy should be seriously considered by funding agencies, which could implement such guidelines in the selection and assignment of new grants. Such a policy would also promote more collaboration, both at national and international levels.

For example, in the Netherlands although a large proportion of the available financial resources for research into type 2 diabetes mellitus has been put

in to work on candidate genes in well-characterised population cohorts, there are no possibilities for making the samples readily available to other researchers. The scientific boards of the funding agencies could decide on the appropriateness of allowing other individuals or groups to make use of the same resources. A consequence of such a policy would be the need to standardise the storage of patient information and biological samples, whilst always paying due care and attention to protecting the privacy of participants.

In conclusion, combining the results of genome-wide scans, followed by association studies with genes found in the regions reported by the genome-wide scans, will provide further significant new insights into disease aetiology. Our own results demonstrate that a genome-wide scan can identify previously unknown loci and confirm previously describe regions. Hopefully, subsequent association studies will be effective in defining new candidate genes and pathways involved in type 2 diabetes mellitus.

Together with previously reported results, the results described in this thesis will accelerate the efforts to identify the correct susceptibility genes located in the regions described. Combined with new developments in the fields of bioinformatics, genomics and proteomics, this will lead to a greater understanding of the pathogenesis of type 2 diabetes mellitus. Identifying new pathways involved in the disease aetiology may help determine new therapeutic targets and direct efforts to target therapies at relevant tissues.

There is little doubt that a detailed genetic dissection of type 2 diabetes mellitus will lead to an improved classification of type 2 diabetes. Together with new insights into pharmacogenetics, such genetic information will form the basis for the design and development of new drug therapies based on individual specificities rather than on a concept of type 2 diabetes mellitus as a global, homogeneous disease. Only then will pharmacotherapy be able to lead to an effective cure and/or prevention of type 2 diabetes mellitus.

References

Almind K, Bjorbaek C, Vestergaard H, Hansen T, Echwald S, Pedersen O, *Aminoacid polymorphisms of insulin receptor substrate-1 in non-insulin-dependent diabetes mellitus.* Lancet, 1993. 342(8875): p. 828-32.

Almind K, Inoue G, Pedersen O, Kahn CR, *A common amino acid polymorphism in insulin receptor substrate-1 causes impaired insulin signaling. Evidence from transfection studies.* J Clin Invest, 1996. 97(11): p. 2569-75.

Altshuler D, Hirschhorn JN, Klannemark M, Lindgren CM, Vohl MC, Nemesh J, Lane CR, et al, *The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes.* Nat Genet, 2000. 26(1): p. 76-80.

Ardlie KG, Kruglyak L, Seielstad M, *Patterns of linkage disequilibrium in the human genome.* Nat Rev Genet, 2002a. 3(4): p. 299-309.

Ardlie KG, Lunetta KL, Seielstad M, *Testing for population subdivision and association in four case-control studies.* Am J Hum Genet, 2002b. 71(2): p. 304-11.

Barcellos LF, Klitz W, Field LL, Tobias R, Bowcock AM, Wilson R, Nelson MP, et al, *Association mapping of disease loci, by use of a pooled DNA genomic screen.* Am J Hum Genet, 1997. 61(3): p. 734-47.

Barton NH, Keightley PD, *Understanding quantitative genetic variation.* Nat Rev Genet, 2002. 3(1): p. 11-21.

Bogardus C, Baier L, Permana P, Prochazka M, Wolford J, Hanson R, *Identification of susceptibility genes for complex metabolic diseases.* Ann N Y Acad Sci, 2002. 967: p. 1-6.

Brookes AJ, *The essence of SNPs.* Gene, 1999. 234(2): p. 177-86.

Cardon LR, Bell JI, *Association study designs for complex diseases.* Nat Rev Genet, 2001. 2(2): p. 91-9.

Cargill M, Altshuler D, Ireland J, Sklar P, Ardlie K, Patil N, Shaw N, et al, *Characterization of single-nucleotide polymorphisms in coding regions of human genes.* Nat Genet, 1999. 22(3): p. 231-8.

Chakravarti A, *Population genetics—making sense out of sequence.* Nat Genet, 1999. 21(1 Suppl): p. 56-60.

Collins FS, Guyer MS, Charkravarti A, *Variations on a theme: cataloging human DNA sequence variation.* Science, 1997. 278(5343): p. 1580-1.

Daly MJ, Rioux JD, Schaffner SF, Hudson TJ, Lander ES, *High-resolution haplotype structure in the human genome.* Nat Genet, 2001. 29(2): p. 229-32.

DeFronzo RA, *Pathogenesis of type 2 diabetes: metabolic and molecular implications for identifying diabetes genes.* Diabetes Reviews, 1997. 5(3): p. 177-269.

Duggirala R, Blangero J, Almasy L, Dyer TD, Williams KL, Leach RJ, O'Connell P, et al, *Linkage of type 2 diabetes mellitus and of age at onset to a genetic location on chromo-*

some 10q in Mexican Americans. Am J Hum Genet, 1999. 64(4): p. 1127-40.

Ehm MG, Karnoub MC, Sakul H, Gottschalk K, Holt DC, Weber JL, Vaske D, et al, *Genomewide Search for Type 2 Diabetes Susceptibility Genes in Four American Populations.* Am J Hum Genet, 2000. 66(6): p. 1871-1881.

Elbein SC, Chu W, Ren Q, Hemphill C, Schay J, Cox NJ, Hanis CL, et al, *Role of calpain-10 gene variants in familial type 2 diabetes in Caucasians.* J Clin Endocrinol Metab, 2002. 87(2): p. 650-4.

Elbein SC, Hoffman MD, Teng K, Leppert MF, Hasstedt SJ, *A genome-wide search for type 2 diabetes susceptibility genes in Utah Caucasians.* Diabetes, 1999. 48(5): p. 1175-82.

Evans JC, Frayling TM, Cassell PG, Saker PJ, Hitman GA, Walker M, Levy JC, et al, *Studies of association between the gene for calpain-10 and type 2 diabetes mellitus in the United Kingdom.* Am J Hum Genet, 2001. 69(3): p. 544-52.

Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, et al, *The structure of haplotype blocks in the human genome.* Science, 2002. 296(5576): p. 2225-9.

Ghosh S, Watanabe RM, Hauser ER, Valle T, Magnuson VL, Erdos MR, Langefeld CD, et al, *Type 2 diabetes: evidence for linkage on chromosome 20 in 716 Finnish affected sib pairs.* Proc Natl Acad Sci U S A, 1999. 96(5): p. 2198-203.

Ghosh S, Watanabe RM, Valle TT, Hauser ER, Magnuson VL, Langefeld CD, Ally DS, et al, *The Finland-United States Investigation of Non-Insulin-Dependent Diabetes Mellitus Genetics (FUSION) Study. I. An Autosomal Genome Scan for Genes That Predispose to Type 2 Diabetes.* Am J Hum Genet, 2000. 67(5): p. 1174-1185.

Gray IC, Campbell DA, Spurr NK, *Single nucleotide polymorphisms as tools in human genetics.* Hum Mol Genet, 2000. 9(16): p. 2403-8.

Hager J, Hansen L, Vaisse C, Vionnet N, Philippi A, Poller W, Velho G, et al, *A missense mutation in the glucagon receptor gene is associated with non-insulin-dependent diabetes mellitus.* Nat Genet, 1995. 9(3): p. 299-304.

Hani EH, Clement K, Velho G, Vionnet N, Hager J, Philippi A, Dina C, et al, *Genetic studies of the sulfonylurea receptor gene locus in NIDDM and in morbid obesity among French Caucasians.* Diabetes, 1997. 46(4): p. 688-94.

Hanis CL, Boerwinkle E, Chakraborty R, Ellsworth DL, Concannon P, Stirling B, Morrison VA, et al, *A genome-wide search for human non-insulin-dependent (type 2) diabetes genes reveals a major susceptibility locus on chromosome 2.* Nat Genet, 1996. 13(2): p. 161-6.

Hansen LH, Abrahamsen N, Hager J, Jelinek L, Kindsvogel W, Froguel P, Nishimura E, *The Gly40Ser mutation in the human glucagon receptor gene associated with NIDDM results in a receptor with reduced sensitivity to glucagon.* Diabetes, 1996. 45(6): p. 725-30.

Hanson RL, Ehm MG, Pettitt DJ, Prochazka M, Thompson DB, Timberlake D, Foroud

T, et al, *An autosomal genomic scan for loci linked to type II diabetes mellitus and body-mass index in Pima Indians.* Am J Hum Genet, 1998. 63(4): p. 1130-8.

Hegele RA, *Genetic prediction of coronary heart disease: lessons from Canada.* Scand J Clin Lab Invest Suppl, 1999. 230: p. 153-67.

Hegele RA, *SNP judgments and freedom of association.* Arterioscler Thromb Vasc Biol, 2002. 22(7): p. 1058-61.

Hegele RA, Cao H, Harris SB, Zinman B, Hanley AJ, Anderson CM, *Peroxisome proliferator-activated receptor-gamma2 P12A and type 2 diabetes in Canadian Oji-Cree.* J Clin Endocrinol Metab, 2000. 85(5): p. 2014-9.

Hegele RA, Harris SB, Zinman B, Hanley AJ, Cao H, *Absence of association of type 2 diabetes with CAPN10 and PC-1 polymorphisms in Oji-Cree.* Diabetes Care, 2001. 24(8): p. 1498-9.

Hegele RA, Sun F, Harris SB, Anderson C, Hanley AJ, Zinman B, *Genome-wide scanning for type 2 diabetes susceptibility in Canadian Oji- Cree, using 190 microsatellite markers.* J Hum Genet, 1999. 44(1): p. 10-4.

Holmans P, Craddock N, *Efficient strategies for genome scanning using maximum-likelihood affected-sib-pair analysis.* Am J Hum Genet, 1997. 60(3): p. 657-66.

Horikawa Y, Oda N, Cox NJ, Li X, Orho-Melander M, Hara M, Hinokio Y, et al, *Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus.* Nat Genet, 2000. 26(2): p. 163-75.

Hugot JP, Chamaillard M, Zouali H, Lesage S, Cezard JP, Belaiche J, Almer S, et al, *Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease.* Nature, 2001. 411(6837): p. 599-603.

Inoue H, Ferrer J, Welling CM, Elbein SC, Hoffman M, Mayorga R, Warren-Perry M, et al, *Sequence variants in the sulfonylurea receptor (SUR) gene are associated with NIDDM in Caucasians.* Diabetes, 1996. 45(6): p. 825-31.

Ji L, Malecki M, Warram JH, Yang Y, Rich SS, Krolewski AS, *New susceptibility locus for NIDDM is localized to human chromosome 20q.* Diabetes, 1997. 46(5): p. 876-81.

Jorde LB, *Linkage disequilibrium as a gene-mapping tool.* Am J Hum Genet, 1995. 56(1): p. 11-4.

Kahn CR, Vicent D, Doria A, *Genetics of non-insulin-dependent (type-II) diabetes mellitus.* Annu Rev Med, 1996. 47: p. 509-31.

Kruglyak L, *Prospects for whole-genome linkage disequilibrium mapping of common disease genes.* Nat Genet, 1999. 22(2): p. 139-44.

Lai E, *Application of SNP technologies in medicine: lessons learned and future challenges.* Genome Res, 2001. 11(6): p. 927-9.

Lander E, Kruglyak L, *Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results.* Nat Genet, 1995. 11(3): p. 241-7.

Lander ES, *The new genomics: global views of biology.* Science, 1996. 274(5287): p. 536-9.

Lander ES, Schork NJ,Ä*Genetic dissection of complex traits.* Science, 1994. 265(5181): p.

2037-48.

Large V, Hellstrom L, Reynisdottir S, Lonnqvist F, Eriksson P, Lannfelt L, Arner P, *Human beta-2 adrenoceptor gene polymorphisms are highly frequent in obesity and associate with altered adipocyte beta-2 adrenoceptor function.* J Clin Invest, 1997. 100(12): p. 3005-13.

Lok S, Kuijper JL, Jelinek LJ, Kramer JM, Whitmore TE, Sprecher CA, Mathewes S, et al, *The human glucagon receptor encoding gene: structure, cDNA sequence and chromosomal localization.* Gene, 1994. 140(2): p. 203-9.

Mackay TF, *The genetic architecture of quantitative traits.* Annu Rev Genet, 2001. 35: p. 303-39.

Mahtani MM, Widen E, Lehto M, Thomas J, McCarthy M, Brayer J, Bryant B, et al, *Mapping of a gene for type 2 diabetes associated with an insulin secretion defect by a genome scan in Finnish families.* Nat Genet, 1996. 14(1): p. 90-4.

Marx J, *Unraveling the causes of diabetes.* Science, 2002. 296(5568): p. 686-9.

McCarthy D, Zimmet P, *Diabetes 1994 to 2010: global estimates and projections.* 1994. Leverkusen: Bayer AG: p. 1-46.

McCarthy MI, Froguel P, *Genetic approaches to the molecular understanding of type 2 diabetes.* Am J Physiol Endocrinol Metab, 2002. 283(2): p. E217-25.

Parker A, Meyer J, Lewitzky S, Rennich JS, Chan G, Thomas JD, Orho-Melander M, et al, *A gene conferring susceptibility to type 2 diabetes in conjunction with obesity is located on chromosome 18p11.* Diabetes, 2001. 50(3): p. 675-80.

Perusse L, Rice T, Chagnon YC, Despres JP, Lemieux S, Roy S, Lacaille M, et al, *A genome-wide scan for abdominal fat assessed by computed tomography in the Quebec Family Study.* Diabetes, 2001. 50(3): p. 614-21.

Porzio O, Federici M, Hribal ML, Lauro D, Accili D, Lauro R, Borboni P, et al, *The Gly972—>Arg amino acid polymorphism in IRS-1 impairs insulin secretion in pancreatic beta cells.* J Clin Invest, 1999. 104(3): p. 357-64.

Pritchard JK, *Are rare variants responsible for susceptibility to complex diseases?* Am J Hum Genet, 2001. 69(1): p. 124-37.

Pritchard JK, Przeworski M, *Linkage disequilibrium in humans: models and data.* Am J Hum Genet, 2001. 69(1): p. 1-14.

Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, Richter DJ, Lavery T, et al, *Linkage disequilibrium in the human genome.* Nature, 2001. 411(6834): p. 199-204.

Reich DE, Lander ES, *On the allelic spectrum of human disease.* Trends Genet, 2001. 17(9): p. 502-10.

Risch NJ, *Searching for genetic determinants in the new millennium.* Nature, 2000. 405(6788): p. 847-56.

Roberts SB, MacLean CJ, Neale MC, Eaves LJ, Kendler KS, *Replication of linkage studies of complex traits: an examination of variation in location estimates.* Am J Hum Genet, 1999. 65(3): p. 876-84.

Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, Sherry S, et al, *A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms.* Nature, 2001. 409(6822): p. 928-33.

Sax K, *The association of size differences with seed-coat pattern and pigmentation in Phaseolus vulgaris.* Genetics, 1923. 8: p. 552-60.

So WY, Ng MC, Lee SC, Sanke T, Lee HK, Chan JC, *Genetics of type 2 diabetes mellitus.* Hong Kong Med J, 2000. 6(1): p. 69-76.

t Hart LM, Stolk RP, Dekker JM, Nijpels G, Grobbee DE, Heine RJ, Maassen JA, *Prevalence of variants in candidate genes for type 2 diabetes mellitus in The Netherlands: the Rotterdam study and the Hoorn study.* J Clin Endocrinol Metab, 1999. 84(3): p. 1002-6.

Tabor HK, Risch NJ, Myers RM, *Opinion: Candidate-gene approaches for studying complex genetic traits: practical considerations.* Nat Rev Genet, 2002. 3(5): p. 391-7.

Terwilliger JD, Weiss KM, *Linkage disequilibrium mapping of complex disease: fantasy or reality?* Curr Opin Biotechnol, 1998. 9(6): p. 578-94.

Vaessen N (2001) *Genetic determinants of diabetes and vascular complications.* Ph.D. thesis, Department of Epidemiology & Biostatistics; Erasmus Medical Center, Rotterdam, pp 167.

Van Eerdewegh P, Little RD, Dupuis J, Del Mastro RG, Falls K, Simon J, Torrey D, et al, *Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness.* Nature, 2002. 418(6896): p. 426-30.

Velho G, Froguel P, *Genetic determinants of non-insulin-dependent diabetes mellitus: strategies and recent results.* Diabetes Metab, 1997. 23(1): p. 7-17.

Waeber G, Delplanque J, Bonny C, Mooser V, Steinmann M, Widmann C, Maillard A, et al, *The gene MAPK8IP1, encoding islet-brain-1, is a candidate for type 2 diabetes.* Nat Genet, 2000. 24(3): p. 291-5.

Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, Ghandour G, et al, *Large-scale identification, mapping, and genotyping of single- nucleotide polymorphisms in the human genome.* Science, 1998. 280(5366): p. 1077-82.

Watanabe RM, Ghosh S, Langefeld CD, Valle TT, Hauser ER, Magnuson VL, Mohlke KL, et al, *The Finland-United States Investigation of Non-Insulin-Dependent Diabetes Mellitus Genetics (FUSION) Study. II. An Autosomal Genome Scan for Diabetes-Related Quantitative-Trait Loci.* Am J Hum Genet, 2000. 67(5): p. 1186-1200.

Weiss KM, *In search of human variation.* Genome Res, 1998. 8(7): p. 691-7.

Weiss KM, Terwilliger JD, *How many diseases does it take to map a gene with SNPs?* Nat Genet, 2000. 26(2): p. 151-7.

Wright AF, Carothers AD, Pirastu M, *Population choice in mapping genes for complex diseases.* Nat Genet, 1999. 23(4): p. 397-404.

Zavattari P, Deidda E, Whalen M, Lampis R, Mulargia A, Loddo M, Eaves I, et al, *Major factors influencing linkage disequilibrium by analysis of different chromosome re-*

gions in distinct populations: demography, chromosome recombination frequency and selection. Hum Mol Genet, 2000. 9(20): p. 2947-57.

Zimmet P, Alberti KG, Shaw J, *Global and societal implications of the diabetes epidemic.* Nature, 2001. 414(6865): p. 782-7.

Zimmet PZ, *Diabetes epidemiology as a tool to trigger diabetes research and care.* Diabetologia, 1999. 42(5): p. 499-518.

Zouali H, Hani EH, Philippi A, Vionnet N, Beckmann JS, Demenais F, Froguel P, *A susceptibility locus for early-onset non-insulin dependent (type 2) diabetes mellitus maps to chromosome 20q, proximal to the phosphoenolpyruvate carboxykinase gene.* Hum Mol Genet, 1997. 6(9): p. 1401-8.

Chapter 8

English Summary

Chapter 8

Little is known about the nature of genetic variation underlying complex diseases in humans. The recognition that susceptibility to type 2 diabetes mellitus has a strong inherited component provides a mechanism for developing the molecular understanding of the pathogenesis of type 2 diabetes mellitus through various genetic approaches. The main aim of the Breda study described in this thesis was to identify genetic factors involved in type 2 diabetes mellitus in a defined Dutch population.

Chapter 1 gives an overview of the approaches that can be used to identify genetic factors in type 2 diabetes mellitus and, in particular, the role of candidate gene analysis and genome-wide scanning is emphasised. **Chapter 2** describes the collection of the Breda Cohort together with the clinical data obtained from all participants. The objective to collect at least 250 affected sibpairs with type 2 diabetes mellitus was met, with a total of 312 affected sibpairs from 178 families being sampled. These families were used for the genome-wide scans described in **chapters 3 and 4**. The Breda cohort also contained 542 independent patients with type 2 diabetes mellitus, which were subsequently used in the candidate gene analyses, as described in **chapters 5 and 6**.

Chapter 3 describes the results of a genome-wide scan performed for type 2 diabetes mellitus in a defined Dutch population. The genome-wide scan was carried out using identity-by-descent analysis in affected sibpairs. Since obesity and type 2 diabetes mellitus are inter-related, the data set was stratified for the sub-phenotype body mass index (BMI), corrected for age and gender. This resulted in a suggestive maximum multi-point LOD score of 2.3 (p value 9.7×10^{-4}) for the most obese 20% pedigrees of the data set in the region flanked by marker loci D18S471 and D18S843 (chromosome region 18p11). We hereby confirmed the presence of a susceptibility locus on chromosome 18, reported earlier from a Finnish/Swedish population. This finding provided solid and independent evidence that the chromosome 18p11 locus is of definite interest for type 2 diabetes mellitus in connection with obesity in the Breda study cohort. In addition, we demonstrated that in the lowest 80% obese pedigrees ("lean" type 2 diabetes mellitus) two interesting loci on chro-

mosomes 2 and 19 were found with LODs of 1.5 and 1.3, respectively (p-values 7.5 x 10^{-3} and–11.2 x 10^{-3}).

Chapter 4 describes the analysis for linkage to loci influencing BMI (using quantitative trait locus (QTL) mapping) in 420 type 2 diabetes mellitus patients from the Breda cohort for which BMI values where available. Subsequently, the genotype data from the 20% most obese type 2 diabetes mellitus pedigrees ("obese" type 2 diabetes mellitus) was also analysed for linkage to type 2 diabetes mellitus using the ASP analysis. The quantitative (QTL mapping) approach was completely different to the categorical clinical definition used in the ASP analysis described in chapter 3. The QTL results support previous findings of a susceptibility locus (QTL) influencing BMI in type 2 diabetes mellitus residing on chromosome region 11q. In addition, suggestive evidence was found and previous findings confirmed for linkage with type 2 diabetes mellitus on chromosome regions 1q, 11p, and 12q. In general, it appears that the linkage found for type 2 diabetes mellitus in the present cohort is strongly influenced by obesity. This supports the notion that a genetic predisposition to obesity is closely intertwined with one predisposition to type 2 diabetes mellitus. However, our study fails to determine to what extent obesity and type 2 diabetes mellitus are genetically unique entities in their own right.

Chapter 5 describes the results of an association study with the sulphonylurea receptor 1 (SUR1) gene and type 2 diabetes mellitus. The SUR1 is a major determinant of normal glucose-induced insulin secretion in the pancreatic β-cell, and is a target for sulphonylurea type medication. Moreover, the SUR1 gene is located in the chromosomal region 11p, which we identified in the 20% most obese type 2 diabetes mellitus pedigrees (described in chapter 4). Previous studies by others had shown an association between a single nucleotide polymorphism (SNP) of exon 16 (SNP16) of SUR1 with type 2 diabetes mellitus in some populations. In our cohort, no such significant association was found and the frequency of the SNP16 variant of the SUR1 gene was only slightly lower in patients than controls, although not significantly so ($p = 0.10$).

Finally, **Chapter 6** describes the results of an association study with the beta-adrenergic receptor-2 (B2ADR) gene and obesity in type 2 diabetes mellitus subjects from the Breda study cohort. The adrenergic receptor genes are candidate genes for obesity because they regulate lipid mobilisation, energy expenditure and glycogen breakdown. A few studies have indicated an association between a SNP in codon 27 (SNP27) of B2ADR with obesity in non-diabetic (Swedish) subjects; in a group of women, but not men, the Gln27 variant was markedly associated with obesity.

The genotypes of SNP27 in the B2ADR gene were matched to the BMI values of 542 patients in our cohort and were compared using ANOVA with age and gender as covariates. No statistically significant difference was observed between the various groups, implying that, in our cohort, this polymorphism has no important effect on body mass index. We also found no effect from the polymorphism on either the age of diagnosis of diabetes or on plasma lipids levels.

The results of this thesis will, together with previously reported results, help accelerate the efforts to identify susceptibility genes for type 2 diabetes mellitus located in the regions described above. Combined with new developments in the fields of bioinformatics, genomics and proteomics, this will lead to a greater understanding of the pathogenesis of type 2 diabetes mellitus. Identifying new pathways involved in the disease aetiology may help identify new therapeutic goals, and direct efforts to target therapies to relevant tissues. By improving the classification of type 2 diabetes, together with new insights in pharmacogenetics, this genetic information may form the basis for the development of new drug therapies and hopefully, in the future, will lead to the prevention of type 2 diabetes mellitus.

Chapter 9

Nederlandse samenvatting

Type 2 diabetes mellitus komt vaak binnen families voor en is een multifactoriële aandoening, d.w.z. dat meerdere erfelijke factoren samen met omgevingsfactoren een rol spelen in het ontstaan van type 2 diabetes. Er zijn tot dusver enkele zeldzame vormen van erfelijke diabetes ontdekt (MODY en MIDD). De genetische oorzaak van de meest voorkomende vorm van type 2 diabetes is echter volledig onbekend.

Het doel van de Breda studie, beschreven in dit proefschrift, was het vinden van genen die betrokken zijn bij het verkrijgen van type 2 diabetes in een gekarakteriseerde Nederlandse populatie. Voor dit onderzoek hebben we gebruik gemaakt van families waarin bij broers en/of zussen diabetes voorkomt (zgn. aangedane sibparen). Eventuele niet zieke broers of zussen zijn ook betrokken bij het onderzoek om de overerving beter te volgen in deze families.

In **hoofdstuk 1** wordt een overzicht gegeven over de mogelijk genetische onderzoek methoden om de genetische factoren op te sporen bij complexe ziekten zoals type 2 diabetes mellitus. Zo wordt er aandacht besteed aan de twee meest gebruikte methoden, de kandidaat gen aanpak en de genome wijde scan aanpak. Vervolgens wordt in **hoofdstuk 2** beschreven hoe het Breda Studie Cohort werd samengesteld. Het doel om tenminste 250 aangedane sibparen te verzamelen werd ruimschoots gehaald met uiteindelijk 312 aangedane sibparen uit 178 families.

Van alle aangedane sibparen en de extra broers/zussen werd DNA verzameld (± 1000 personen). Het DNA werd vervolgens onderzocht met een groot aantal DNA markers (± 300) die verspreid liggen over alle chromosomen. Vervolgens werd het genetisch materiaal van de aangedane sibparen met elkaar vergeleken. Sibparen zijn voor gemiddeld 50% van hun erfelijk materiaal aan elkaar gelijk. Als blijkt dat ze voor bepaalde delen van het erfelijk materiaal meer dan de verwachte 50% met elkaar gemeen hebben, is dit een aanwijzing dat zich op die plaats een gen kan bevinden dat betrokken is bij het ontstaan van type 2 diabetes. Hiermee kan hopelijk worden vastgesteld welke stukjes van de chromosomen samenhangen met het ontstaan van type 2 diabetes.

Dit deel van het onderzoek wordt beschreven in **hoofdstuk 3 en** 4. Het

Breda Studie Cohort bevat naast de 312 aangedane sibparen ook 542 onafhankelijke (geen familie van elkaar) patiënten met type 2 diabetes mellitus. Het DNA van deze 542 onafhankelijke patiënten werd gebruikt om naar twee verschillende kandidaat genen te kijken, dit deel van het onderzoek wordt beschreven in **hoofdstuk 5 en 6**.

Zoals gezegd, in **hoofdstuk 3** worden de resultaten beschreven van de genome wijde scan aanpak in onze populatie van 312 aangedane sibparen. Het is gebleken, uit ander onderzoek, dat zwaarlijvigheid en type 2 diabetes mellitus veel met elkaar te maken hebben en ook vaak te gelijkertijd voorkomen bij patiënten. Het cohort van 178 families werd gestratificeerd op het subfenotype body mass index (BMI = een maat om zwaarlijvigheid mee te bepalen), deze waarden werden gecorrigeerd voor leeftijd en geslacht. Dit resulteerde, voor een gebied op chromosoom 18, in een LOD score (een maat om aan te geven of er werkelijk een gen ligt dat betrokken is bij de ziekte) van 2.3 in de 20% meest zwaarlijvige families in ons Cohort. Het blijkt dat dit resultaat voor chromosoom 18 overeenkomt met een zelfde gebied wat gevonden is in zwaarlijvige Zweden en Finnen, en al eerder is beschreven. Ons resultaat geeft dus een onafhankelijk bewijs dat er mogelijk een gen ligt op chromosoom 18 dat betrokken is bij het verkrijgen van type 2 diabetes mellitus. Naast dit resultaat hebben we ook nog twee andere gebieden gevonden op de chromosomen 2 en 19, alleen nu in de magere families met type 2 diabetes mellitus.

In **hoofdstuk 4** wordt het onderzoek beschreven naar chromosoom gebieden die BMI beïnvloeden in 420 type 2 diabetes patiënten van het Breda Studie Cohort (zgn. Quantitative Trait Locus-analyse). Tevens is er ook gekeken welke gebieden nog meer betrokken zijn bij de zwaarlijvige type 2 diabetes mellitus patiënten.

De resultaten van de QTL-analyse onderschrijven eerdere bevindingen van gebieden die BMI beïnvloeden in type 2 diabetes mellitus patiënten, zoals een gebied op chromosoom lange arm van chromosoom 11. Naast het al beschreven gebied op chromosoom 18 (**hoofdstuk 4**) hebben we tevens nog drie andere gebieden gevonden op chromosomen 1, 11 (korte arm) en 12. Het

blijkt echter dat de gebieden gevonden voor type 2 diabetes mellitus in het Breda Studie Cohort zeer sterk beïnvloed worden door zwaarlijvigheid. Dit onderschrijft de notie dat een genetische gevoeligheid voor type 2 diabetes mellitus zeer sterk verbonden is met een genetische gevoeligheid voor zwaarlijvigheid. Helaas is het in onze studies niet gelukt om een duidelijk onderscheid te maken in hoever zwaarlijvigheid en type 2 diabetes mellitus twee genetisch onafhankelijke unieke ziekten zijn.

In **hoofdstuk 5 en 6** worden de resultaten beschreven van twee onderzoeken naar mogelijke kandidaat genen die betrokken kunnen zijn bij type 2 diabetes mellitus.

In **hoofdstuk 5** wordt het resultaat beschreven van een associatie studie tussen variaties in het SUR1 gen (sulphonylurea receptor 1 gen) en type 2 diabetes mellitus. Het SUR1 gen is een determinant van normaal geïnduceerde insuline secretie in de pancreas (alvleesklier), het is ook het doel van de sulphonylurea type medicijnen. Het blijkt ook dat dit gen ligt op de korte arm van chromosoom 11 in het zelfde gebied wat wordt gevonden in zwaarlijvige type 2 diabetes mellitus patiënten (**hoofdstuk 4**).

Uit eerdere studies is gebleken dat een bepaalde variant van het gen (een single nucleotide polymorphism (SNP)) geassocieerd is met type 2 diabetes mellitus in verschillende populaties. De SNP, gelegen in exon 16 van het SUR1 gen, blijkt in onze populatie niet geassocieerd te zijn met type 2 diabetes mellitus.

In **hoofdstuk 6** wordt het resultaat beschreven van een associatie studie tussen een variant van het beta adrenergic receptor 2 gen (B2ADR) en zwaarlijvigheid in type 2 diabetes patiënten. Dit type adrenergic receptor genen zijn goede kandidaat genen voor zwaarlijvigheid omdat zijn de vet mobilisatie, energie verbruik en glycogeen afbraak reguleren. Enkele studies hebben aangegeven dat er een mogelijke associatie is tussen een SNP in codon 27 van het gen en zwaarlijvigheid in niet diabetes patiënten. Helaas was deze SNP van het B2ADR gen niet geassocieerd in onze Breda Studie Cohort en type 2 diabetes patiënten.

De resultaten beschreven in dit proefschrift zullen hopelijk, samen met

eerder beschreven resultaten, bijdragen aan het vinden en identificeren van genen betrokken bij type 2 diabetes mellitus in de beschreven chromosoom gebieden. Gecombineerd met nieuwe ontwikkelingen op het gebied van bio-informatica, genomics, en proteomics zal dit uiteindelijk leiden naar een beter inzicht in het ontstaan en de ziekte-ontwikkeling van type 2 diabetes mellitus. Het identificeren van nieuwe routes betrokken bij de ziekte kunnen mogelijk helpen bij het ontwikkelen van nieuwe medicijnen.

Door het beter indelen van type 2 diabetes mellitus in diverse subgroepen en een beter inzicht in pharmacogenetica, kan deze genetische informatie een basis vormen voor het ontwikkelen van nieuwe therapieën een hopelijk, in de toekomst leiden tot het voorkomen van type 2 diabetes mellitus.

Tot slot

Tja, tot slot, is het dus tijd om een aantal mensen te bedanken voor de hulp met het volbrengen van dit proefschrift in het wel bekende dankwoord.
En ja, waar te beginnen …

Allereerst wil ik alle families bedanken voor hun deelname aan het onderzoek en voor hun bereidwilligheid om bloed af te staan. Natuurlijk ook de 80 huisartsen uit de omgeving Breda die hun patiënten hebben aangemoedigd om mee te doen aan het onderzoek. Niet te vergeten natuurlijk Marjan de Jong en Erdtsienk Ernste, de onderzoek coördinatoren van de Stichting Huisartsen Laboratorium Breda zonder wie het onmogelijk was geweest om de bloed afnamen zo efficiënt te laten verlopen, dit natuurlijk met de hulp van een grote hoeveelheid afnameverplegers(sters).

Peter, het was een zware bevalling maar toch, alles is op z'n pootjes terecht gekomen. Toch, bedankt voor de getoonde interesse en de vruchtbare discussies vooral met betrekking tot de general discussion.

Timon, je bent een zeer deskundige copromotor; bedankt voor je vakkundige inbreng en soms kritische noot in het gehele onderzoek. Ook voor je af en toe relativerende woorden als het tegen zat.

Cisca, hoe moet ik jou bedanken, je hebt me laten inzien dat onderzoek wel degelijk leuk kan zijn, ondanks alle tegenslagen. Ben je erg dankbaar voor alle inspirerende woorden tijdens onze gespreken en geloof in het onderzoek. De jaarlijkse bbq's en nieuwjaars borrels waren voortreffelijk, was onmogelijk om deze over te slaan door het heerlijke eten wat altijd aanwezig was. De aubergines van de barbecue en de heerlijke sushi zal ik nooit vergeten, en dat zijn maar enkele voorbeelden. Niet alleen een voortreffelijke wetenschapper, maar ook nog een keuken prinses waar vindt je dat hedentendagen. Bedankt voor alles.

Eric, waar zou ik zijn zonder ..., ja waarschijnlijk nog steeds aan het PCR-en en runnen van alle samples. Dat fijne stukje techniek bleek niet te doen wat was beloofd. De hierdoor opgelopen achterstand hebben wij toch mooi in een maand of twee weer ingehaald, wel natuurlijk door met z'n tweeën op te gaan in een goed geoliede machine en 7 dagen per week te werken. Te gek en

hartstikke bedankt voor dit alles.
Had het niet beter kunnen treffen met jou als analist en nu als paranimf.

Verder wil ik natuurlijk alle collegae bij de Divisie Medische Genetica bedanken voor de gezellige tijd samen en een paar in het bijzonder.

Harry, bedankt voor de grote hoeveelheid patiënten informatie die je hebt ingevoerd in onze database en het stickeren van alle brieven. Tineke, jij natuurlijk bedankt voor je voortreffelijke doorzettingsvermogen om toch alle mogelijke familieleden boven water te krijgen en te overreden om toch ook mee te doen zelfs als ze zelf niet ziek waren.

Jackie, jij bedankt voor al het correctie werk, van alle manuscripten en natuurlijk dit proefschrift, vraag me soms wel eens af waar we, als vakgroep, zouden zijn zonder jou Engelstalige inbreng, je bent een fantastische editor, bedankt.

Harm, bedankt voor je heerlijk relativerend vermogen, wat is het leven toch ... (je standaard zin) en voor alle gezellige sneak-avonden en discussies over alles en nog wat, gelukkig ben je niet zo uitgesproken (hum). Bart, ook jij bedankt voor de vermakelijke uurtjes in de AIO kamer, heel veel succes in de States en wie weet …

Leida bedankt voor je werk aan beide kandidaat genen, en Tjing jij bedankt voor je werk aan chromosoom 18, jullie waren twee voortreffelijke studenten en wens jullie veel succes met jullie, mogelijke, wetenschappelijke carrière.

Jos & Stephan, Marcel, Sjoerd, bedankt dat jullie er voor mij waren in die moeilijke perioden van de afgelopen 4 jaar, echte vrienden.

Toos, bedankt dat je me weer op het juiste pad hebt geholpen, je massages waren en zijn heerlijk en ontspannend, bedankt voor je liefde.

Anneke, bedankt dat je me weer gevonden heb, me alles in perspectief laat zien. Alles is compleet met jou. Lief, 'k hou van je.

Peter en Ine, hoezo cliché, zonder jullie had ik dit punt nooit bereikt in mijn leven. Dankzij jullie geloof in mij en jullie onvoorwaardelijke steun en liefde is dit allemaal tot stand gekomen, kan onmogelijk verwoorden hoe zeer ik jullie dankbaar ben voor alles wat jullie de afgelopen 32 jaar voor mij hebben gedaan.

Curriculum Vitae

Jonathan van Tilburg werd op 19 maart 1970 geboren te Schiedam. Nadat hij in 1988 zijn HAVO-diploma haalde aan Scholengemeenschap de Wegwijzer te Sleeuwijk begon hij met de opleiding tot laborant aan de Hogeschool van West-Brabant te Etten-Leur.

Na zijn afstudeer stage bij FMC Bioproducts in Rockland, Maine in de Verenigde Staten begon hij aan de studie Medische Biologie aan de Universiteit Utrecht. Gedurende deze studie heeft hij onderzoek gedaan naar mutaties bij Wilson Disease patienten bij het Klinisch Genetisch Centrum te Utrecht (hoofd DNA-lab: Dr. H.K. Ploos van Amstel). Zijn afstudeeronderwerp over mRNA expressie in de lever van genen betrokken bij de mitochondriale verbranding van vetzuren heeft hij uitgevoerd bij de vakgroep Metabole Ziekten in het Wilhelmina Kinderziekenhuis te Utrecht (hoofd: Prof. Dr. R. Berger).

In oktober1998 studeerde hij af, waarna hij onder begeleiding van Dr. Cisca Wijmenga en Dr. Timon van Haeften aan zijn promotie begon bij de divisie Medische Genetica van het Universitair Medisch Centrum Utrecht, wat heeft geresulteerd in dit proefschrift.

Per 2003 zal hij werkzaam zijn aan Harvard Institutes of Medicine (onderleiding van Dr. Ping Lu) Harvard Univeristy, en zich gaan bezig houden met telomeer regulatie en kankeronderzoek.

List of publications

Peer-reviewed publications

Van Tilburg JHO, Sandkuijl LA, Strengman E, van Someren H, Rigters-Aris T, Pearson PL, van Haeften TW, Wijmenga C (2002) A genome-wide scan in type 2 diabetes mellitus provides independent replication of a susceptibility locus on 18p11 and suggests the existence of novel loci on 2q12 and 19q13. **In press**

Van Tilburg JHO, Wijmenga C, van Bakel HHMJ, Rozeman LB, Pearson PL, van Haeften TW (2002) Relationship of beta2-adrenergic receptor polymorphism with obesity in type 2 diabetes mellitus. **In press**

Van Tilburg J, van Haeften TW, Pearson P, Wijmenga C (2001) Defining the contribution of type 2 diabetes mellitus. **J Med Genet** 38: 569-578.

Van Tilburg JHO, Rozeman LB, van Someren H, Rigters-Aris CAE, Freriks JP, Pearson PL, Sandkuijl LA, van Haeften TW, Wijmenga C (2000) The exon 16-3t variant of the sulphonylurea receptor gene is not a risk factor for type II diabetes mellitus in the Dutch Breda cohort. **Diabetologia** 43: 681-682.

Published abstracts

Wijmenga C, Van Tilburg JHO, Sandkuijl LA, Strengman E, Pearson PL, van Haeften TW (2002) A genome-wide scan in type 2 diabetes mellitus provides independent replication of a susceptibility locus on 18p11 and suggests existence of novel loci on 2q12 and 19q13. **Am J Hum Genet 71:455, Suppl.**

Van Tilburg JHO, Sandkuijl LA, Strengman E, Pearson PL, van Haeften TW, Wijmenga C (2002) Genome-wide scan in a subset of obese type 2 diabetes patiens and subsequent QTL mapping for BMI suggests linkage to two distinct loci on chromosome 11. **Am J Hum Genet 71:455, Suppl.**

Van Tilburg JHO, Strengman E, van Someren H, Rigters-Aris CAE, Freriks JP, Pearson PL, Sandkuijl LA, van Haeften TW, Wijmenga C (2001) Are genes on chromosome 20 involved in type 2 diabetes mellitus in the Netherlands? **Neth J Med** 2001:59:A1-A7

Van Tilburg JHO, Strengman E, Cheng T, van Someren H, Rigters-Aris CAE, Freriks JP, Pearson PL, Sandkuijl LA, van Haeften TW, Wijmenga C (2001) Genomic screen in type 2 diabetes mellitus sibpairs in the Dutch population. The Breda Study. **Am J Hum Genet** 69 :479, Suppl.

van Tilburg JH, van Someren H, Rigters-Aris CA, Freriks JP, Pearson PL, Sandkuijl LA, van Haeften TW, Wijmenga C (2000) Genomic screen in type 2 diabetes mellitus sibpairs in the Dutch population. **Eur J Human Genet** 8: 653.