

The genetics of type 2 diabetes: A review

Alan Taylor

Kellogg College, University of Oxford, Wellington Square, Oxford, OX1 2JA United Kingdom

Abstract

Because of the widespread distribution and increasing prevalence of Type 2 diabetes in modern society, great efforts have been made to understand the underlying causes of the disease. The fact that the cause of Type 2 diabetes is primarily of genetic origin has been known for many years. In the last five years advances in the microarray analysis of gene expression and statistical genetics had provided hope that these techniques would give a clear identification of the genes involved and an understanding of the genetic nature of the disease. An analysis of recent work in this area has been carried out and possible reasons put forward as to why the results have not been as useful as might have been hoped. (Int J Diabetes Metab 14: 76-81, 2006)

Key Words: Genetics, diabetes

Introduction

Diabetes mellitus affects approximately 150 million people around the world. It is forecasted that by 2025 this number would be doubled, with a prevalence that varies markedly from population to population.¹ Many causal factors have been identified which influence the prevalence, that is the number of cases as a percentage of the total population or the incidence which is the 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 amount of body fat, hypertension, lack of physical exercise, and ethnic background. The general definition of diabetes, as being an insulin deficiency may be divided into two disease types known as Type 1 and Type 2 diabetes. Both types of diabetes show a disposition to run in families, which is a strong indication of a genetic cause for the diseases. However, the aetiology underlying types 1 and 2 is quite different indicating that different genes are likely to be involved in the two types of diabetes mellitus.

Types of diabetes

As the major cause of Type 1 diabetes mellitus is the autoimmune destruction of the pancreatic β cells, this form of diabetes usually appears or is first diagnosed in young people and is sometimes known as "juvenile diabetes". Type 1 diabetes mellitus requires continual insulin therapy to prevent diabetic ketoacidosis, diabetic coma and ultimately death.

Type 2 diabetes which accounts for around 90% of all cases of diabetes mellitus, usually develops after the age of 40, the disease being known as "adult diabetes mellitus". Unlike Type 1, Type 2 diabetes is not caused by autoimmune destruction of the pancreatic β cells, but appears to be caused by a number of different defects in both insulin action and insulin secretion. In fact, it would be more accurate to define Type 2 diabetes as a common symptom caused by a number of different diseases. For example,

with Type 2 diabetes both insulin's inhibitory effect on liver glucose production and its stimulatory effect on peripheral glucose uptake are diminished.² It is usually possible to treat mild Type 2 diabetes by changes to the diet or with oral hypoglycaemic drugs. However, as the disease develops it may become necessary to use insulin to overcome the hyperglycaemia.

The long term complications of Type 2 diabetes include chronic complications that have serious impact on the health and quality of life of patients. These include circulatory problems that may cause blindness, lower limb gangrene and renal failure in adults. It has been shown that Type 2 diabetes mellitus is also a major risk factor for strokes and other cardiovascular diseases.

Genetics of Type 2 diabetes

There is strong evidence that Type 2 diabetes is inherited and has a genetic origin. The risk of developing Type 2 diabetes is approximately 3-4 times higher among first degree relatives of diabetic subjects compared to subjects without a family history of diabetes and similar values have been found from the offspring of diabetic subjects. The relative risk of acquiring a disease of genetic origin is called the lambda value (λ). It gives an estimate of the risk of disease in subjects related to affected individuals compared to the risk in the general population. However, the λ value cannot distinguish familial genetic from non-genetic components. In the case of Type 2 diabetes the λ value has been determined as 3.5.

If one parent has diabetes the risk that the offspring will develop the disease is about 40%, while if both parents have diabetes the risk is approximately 70%.³ This supports the hypothesis that there are familial factors that contribute to the disease and suggests that these factors are, to some extent, additive. Very high concordance rates of Type 2 diabetes have been reported in monozygotic twins. One population-based twin study suggested concordance rates of 34% among monozygotic and 16% among dizygotic twin pairs.³ Therefore, approximately 40% of the variability of the diabetic phenotype may be heritable (familial genetic).

There is also strong evidence for a significant environmental causal involvement in Type 2 diabetes. It

Received on: 13/5/06

Accepted on: 11/9/06

Correspondence to: Alan Taylor, Kellogg College, University of Oxford, Wellington Square, Oxford, OX1 2JA United Kingdom, Email ataylor@biop.ox.ac.uk, Fax: 44 1865 270314

has been shown that the prevalence of the disease is directly proportional to the degree of westernization of the population studies. However, it is more likely related to the degree of affluence rather than any specific element of western culture.⁴ The incidence of Type 2 diabetes among the Chinese in rural China is approximately 1% but among Hong Kong Chinese it is 9% and among the Chinese population of Mauritius it is 15%.⁵ In the aboriginal populations of Australia and Mexico that have adopted a “western” lifestyle it is 25% and 45%, respectively whereas among European Caucasians it is 5-10%. The explanation for these statistics is the so called, “thrifty gene” theory. In populations that in the past would have been exposed to repeated periods of famine and unpredictable food supply there would be a natural selection for alleles that predisposed towards energy conserving genotypes. If a human stored energy as fat rather than glycogen then that person would be better protected against the potentially fatal effects of starvation. Such a genotype in a prolonged period of plentiful food would then be predisposed to insulin disorders and related obesity, particularly if the transition in evolutionary time had been rapid.

Diseases (or traits) that can be caused by a combination of genetic and environmental factors are called complex or multifactorial. Many of the most common diseases are multifactorial, such as cancer, asthma, diabetes and inflammatory bowel syndrome. These diseases are heritable, but what is inherited is not the disease itself but rather the susceptibility to it. While some genetic diseases may be caused by a mutation on a single gene and are known as monogenic most complex diseases are caused by mutations on a numbers of genes and are termed polygenic. Examples of monogenic diseases are cystic fibrosis and phenylketonuria both of which are caused by a single SNP (single nucleotide polymorphism) with both having a high penetrance, that is all people with that specific mutation will develop the disease. The environmental influence on a disease may also vary with the amount that the population has been exposed to a specific environmental factor.

A person who develops a complex disease may carry several genetic factors that predispose them to the disease that can then be initiated by environmental factors. Such complex diseases are particularly difficult to study genetically because of the number of variables.

Two variables describing the size of the familial genetic component that may be determined from family data are the heritability (h^2) and the lambda (λ) value. Heritability is the variability of an allele that can be accounted for by heritable (genetic) factors. Heritability can be calculated from the difference of concordance rate among monozygotic and dizygotic twins. Since monozygotic twins share 100% genetic material and dizygotic twins share 50% of the genetic material, the concordance rate is larger among monozygotic twins if the trait/disease has considerable genetic causation. Segregation analyses investigate the method by which a genetic disease is transmitted through families. Mendelian genetic disease can be dominant,

codominant, recessive, autosomal or X-/Y-linked and maternally transmitted. However, when the process of inheritance cannot be readily determined, this usually means that the disease is not inherited in a Mendelian fashion. Non-Mendelian genetic diseases may need two or more genetic factors in order to occur ie. be polygenic and they may require environmental factors as well. Both Mendelian and non-Mendelian genetic diseases may have variable penetrance, which means that the phenotype can be variably expressed or not expressed at all in some individuals. A higher penetrance may be the result of a strong interaction with an environmental factor. A person who develops a complex disease may carry several genetic factors that predispose to the disease, and the number of predisposing genes will most likely affect the severity as well as the age at onset of the disease.

The use of gene expression microarrays to investigate complex diseases has become commonplace.^{6,7} Microarrays have also been used to identify disease subtypes including diabetes.⁸ However the use of genetic profiling to identify the genes underlying common polygenic diseases has been less than successful. Microarray data has been used in conjunction with other techniques to try to identify genetic relationships with complex traits such as diabetes.⁹⁻¹¹ Schadt¹² developed the “likelihood causality model selection” (LCMS) process which was used to identify genetic links to obesity. One gene in particular, *Tgfb2* seemed to have a strong link in mice. This gene has also been linked to both obesity and Type 2 diabetes.^{13,14} Recent work by Goodarzi *et al*¹⁵ found that a single SNP in the human lipoprotein lipase gene was linked to both obesity and insulin resistance in a human population.

The integration of genotypic and expression data may help to find target genes linked to common human diseases but caution is required because there are several opportunities for errors and variation in such sensitive techniques. Much work has been done in the field of population genetics to try to discover the genetic origins of Type 2 diabetes.

In order to find the gene or genes that may be the cause of the disease, it is logical to look for candidate genes that play a part in insulin control and then to look for a significant association between the occurrence of diabetes and polymorphisms on the candidate genes. Such an association analysis is done by comparing a random sample of subjects with type 2 diabetes with a matched control group in order to find a polymorphic allele with an increased frequency in the patient group. The idea is that a significant association would point to the disease susceptibility locus on the gene. Although there have been a number of such studies no useful associations have been found.¹⁶ There are a number of reasons why these studies have been unfruitful. It is quite possible that the genes showing the polymorphism are peculiar to the specific population being tested and only cause diabetes in that specific population or that the genetic differences are widespread but that the association with diabetes only occurs because of general differences between populations.

Any population stratification caused by non-random mating allows marker allele frequencies to vary among the population as a result of genetic drift and this frequently can give rise to false associations in the analysis of the results.¹⁷ It is possible to overcome these problems by using the Transmission Disequilibrium Test that looks at the genotypes of the parents of the patients. This test takes advantage of population level associations and is not susceptible to the errors caused by population stratification. Unfortunately, because Type 2 diabetes is a late onset disease, it is likely that the parents of the patient will not be alive to give samples of DNA. Therefore, the only way to explore the associations between specific genes and the occurrence of the disease is the analysis of the genomes of a large number of gene families in order to determine the contribution made by the genetic background.¹⁸

The analysis or mapping of the genome is done using a method known as *Linkage Disequilibrium* (LD). This maps the correlation between specific alleles at two loci such as a marker locus and a disease locus in a population. LD occurs because of the fact that apparently unrelated subjects are likely to have distant relatives. Therefore, two loci that are very closely linked on the chromosome and unlikely to be separated by recombination will likely be transmitted together through generations to seemingly unrelated individuals. LD occurs when a marker allele lies so close to the disease susceptibility allele that these alleles are inherited together over many generations. Thus, the same allele will be detected in affected subjects in multiple, but apparently unrelated families. The genetic mapping has to be followed by testing all the candidate genes from the region for their involvement in the disease and this should result in the positional cloning of a gene associated with type 2 diabetes mellitus. The actual degree of association between the disease loci and markers on the chromosome is measured on a scale called the LOD score. It is a measure of recombination frequencies that are then converted into map distances measured in units called centimorgans (cM). The LOD (logarithm of odds) score indicates the likelihood of linkage between two relevant markers. It is equal to \log_{10} of the ratio of the probability that two markers are linked with a given recombination value divided by the probability that they are unlinked. For example, a LOD value of 3.0 means that the chances are greater than 1000:1 that the markers are linked for a given recombination estimate. A high LOD value will result in smaller linkage groups and therefore a critical LOD value of 3.0 and over is taken to indicate linkage between pairs of markers.

The LOD score method of linkage analysis is very good at mapping Mendelian disease genes and DNA markers but in order to calculate a LOD score it is necessary for the allele frequencies and penetrance values for all the loci involved to be known. It is possible to analyse diseases with complex (non-Mendelian) inheritance if the values for these parameters are known. However, when transmission is non-Mendelian it can be extremely difficult to estimate the penetrance values and the allele frequency of the disease mutation. Sometimes it is possible that different mutations

at different loci have different kinds of effect on susceptibility, some major and some minor, some dominant and some recessive. If there is a different mode of transmission in different families, or if different loci interact in the same family, then no single transmission model may be valid. Because of these problems, a variety of methods have been developed to test for linkage without the need to specify values for the parameters defining the transmission model, and these methods are termed nonparametric.

A non-parametric analysis method can overcome the problems caused by the lack of Mendelian inheritance and the lack of parental DNA, since this requires no knowledge of the mode of inheritance of the disease, the disease allele (gene) frequencies, or the penetrance.¹⁹

A commonly used non-parametric genetic mapping approach is the affected sib pair (ASP) approach using randomly spaced polymorphic markers usually every 10 centimorgans for all chromosomes. The basis of the ASP analysis is the fact that subjects that are concordant for a particular genetic trait should have greater than expected concordance for marker alleles that are closely linked to the disease. The most frequently used measure of concordance in two sibs at a locus is the number of shared alleles that are identical by descent. If a marker is not linked to a disease locus, then the probability of sib pair sharing 0, 1, and 2 alleles is 0.25, 0.5 and 0.25. With the mean being 0.5, if a marker is found to have more than 0.5 ASP alleles identical by descent, then it may be assumed that the marker is linked to the disease locus. Using ASPs in genome wide scans generally requires large numbers of ASPs in order to obtain sufficient power for detecting linkage for a given value of λ_r (relative risk for a sib). However, this strategy is very expensive and can be extremely time consuming.

The most effective way of performing a genome-wide scan using ASP is by using the method known as "staged searching". The initial genome scan is carried out with a sparse marker set with an average spacing of 20cM. Regions of interest should exceed a LOD score threshold of 1.0. It has been shown by Risch and Merikangas²⁰ 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. Subsequently, the regions of interest are investigated with a denser marker set with an average spacing of 5 cM. The threshold for significant linkage would be a LOD score of 3.3. Brown and Gorin²¹ have recommended that a three stage strategy, with increasing thresholds at each stage, is the best approach to adopt for a genome wide scan.

The alternative is to use different types of cohorts either as nuclear families or multigenerational families. In 1996, a genome wide significance for Type 2 diabetes was found on chromosome 2q37 in a combined data set of 330 Mexican-American ASPs from Starr County, Texas. This locus was designated NIDDM1. Subsequently in a population from Botnia in Western Finland, a small number of selected pedigrees with high 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 NIDDM2.²²

Elbein and Hoffman²³ and Hanson et al²⁴ have completed genome scans for Type 2 diabetes mellitus using multigenerational families. Both types of genome scans (using ASPs or multigenerational families), yield varying levels of evidence.

In 1998, Hanson²⁴ published the results of a genome scan in Pima Indians that showed strong evidence that chromosome 11q contains a susceptibility locus influencing both Type 2 diabetes mellitus and obesity. Chromosomes 1q and 7q also showed some evidence of additional Type 2 diabetes mellitus susceptibility loci but these were not linked with obesity. The familial aggregation of diabetes in the Pima Indians was found to occur in a manner that is not linked directly to familial obesity, thereby suggesting that other genetic determinants are involved in the susceptibility to Type 2 diabetes. A further study of the Pima Indians by Wolford using linkage disequilibrium found a relationship between Type 2 diabetes and SNP's on the KCNJ9 gene on chromosome 1q21.²⁵ Chromosome 1q was also linked to Type 2 diabetes in the study by Elbein²³ of 42 multigenerational families with northern European ancestry from Utah. A significant linkage was also found under a model of recessive inheritance on chromosome 1q21-23. Unfortunately, attempts to replicate the linkage in a further 20 families was unsuccessful. In addition, when certain subjects with a suspected diagnosis were eliminated from the results the LOD score fell which might indicate that this was a false positive association.

Vionnet et al²⁶ carried out a similar genome wide scan on a white European French population and went to great lengths to obtain a homogenous sample. Some 677 associated sib pairs from 402 families were analysed and non-parametric and multipoint analyses were carried out using the program MAPMAKER-SIBS. The first stage genome scan gave strong evidence of a linkage between Type 2 diabetes and markers on chromosomes 3q27 and 1q21-24. Further markers were added in these regions to give a marker density of 2 centimorgans. The second stage scan confirmed the results found in the first scan. A simulation study confirmed the linkage area on the 3q27 chromosome but the area on chromosome 1q21-24 only gave a linkage to a selected population of non-obese subjects. It was argued that this linkage was likely to indicate a true Type 2 diabetes loci compared to loci having only a strong linkage to both obesity and diabetes. An even larger genome wide scan was carried out by Wiltshire.²⁷ Non parametric linkage analysis gave evidence of a linkage to seven different chromosomes some of which had been identified in other studies but interestingly a linkage with a LOD score of 2 was found on chromosome 1q24-25. Non parametric linkage analysis and the use of the ALLEGRO program to compute the LOD score found seven regions with a LOD score of 1.2 or more but after dense mapping the LOD score for 1q24-25 increased to 2. No evidence was found for any linkage to loci on chromosomes 2q,12q and 20q all of which had

shown linkage in other studies but some linkage was found on chromosome 8p as reported by Hegele et al¹ and also to 10q and 5q similar to the findings of Vionnet et al.²⁶

In the USA a large population assembled for research into heart disease, was divided by Meigs²⁹ into a subpopulation with symptoms of Type 2 diabetes. The severity of the diabetes was then corrected for age, smoking, alcohol and estrogen use, body mass and exercise. While a logical approach, this introduced many opportunities for the data to be misinterpreted. Nevertheless, a linkage was found to chromosome 1q in the same region as that found by both Hanson²⁴ and Wiltshire.²⁷ A linkage was also found to a region associated with the production of the protease calpain 10 by the gene CAPN10. This same linkage had also been found by Harikawa.³⁰ Like Wiltshire²⁷ they also found a linkage to chromosomes 3p,11p and 21p but in regions far distant from those found in previous studies.

In Western Pennsylvania there exists a rather isolated puritan population called the "Amish people" or Pennsylvania Dutch. Originally from Holland they have remained in self imposed isolation for almost 300 years. Because of this isolation they have large sibships with pedigrees that are well defined, and their adherence to a puritan lifestyle means that they have a relatively homogenous environment and lifestyle. Unlike the Pima Indians, they are an almost ideal population for genetic linkage analysis. However the incidence of Type 2 diabetes in the Amish is 5%, very similar to the wider population in the USA and Europe. Hsueh³¹ carried out a genome wide scan using a 10 cM map to find genes linked to Type 2 diabetes. Like others a linkage was found to chromosome 1q21, but the study also found a similar LOD score for a locus on chromosome 14 and a less intense linkage on eight other chromosomes. More recently Xiang³² looked at a population of Han Chinese that was far less homogenous than the Amish population. Nevertheless a linkage to Type 2 diabetes was found on chromosome 1q24 and a similar intensity linkage was found on chromosome 6 and a number of suggested linkages on several other chromosomes. Most authors found a linkage to chromosome 1q within a range of 160-210 centimorgans but there is no way of knowing if it is the same gene in slightly different positions in the different populations or if in fact there are more than one Type 2 diabetes related genes in this part of the chromosome.

It appears that there are several possible conclusions that may be drawn from these studies. It is generally accepted that in order to get a genome-wide significance level of 0.05 or less for evidence of a linkage to a disease locus, the allele sharing LOD score should be 3.0 or better.³³ None of the studies reported showed that level of significance for any linkage. On that basis, no clear-cut linkage has been shown between any specific locus and Type 2 diabetes. Linkages on some genes have been reported by some authors, which cannot be found by others in different populations. In one case a linkage was made which could not be confirmed in a subset of the same population. Even when a linkage has been found on the same chromosome by more than one study it is frequently in a different place on the

chromosome. Only in the case of chromosome 1q21-24 has a linkage, albeit a weak linkage to Type 2 diabetes, been found by repeated studies in different populations. It is possible that this mutation on chromosome 1 is a very old mutation in humans and that all the other suggested linkages are of more recent origin. It may well be that it requires the mutation on chromosome 1q plus other mutations and the change in life style associated with westernisation for Type 2 diabetes to occur.

A problem can arise in any branch of science when the goal and the potential benefits of a study, such as determining the cause of Type 2 diabetes, are so big that it becomes a search for the "Holy Grail". With 150 million people worldwide affected by Type 2 diabetes, finding the genetic link to the "disease" would appear to be such a goal. Does the size of the goal merit the use and reporting of "suggested linkages" and weak linkages below the LOD score shown to be acceptable for a genome wide significance? It is obvious that Type 2 diabetes is a complex disease and it is known that problems can occur in defining the disease, the genetic variability of populations is well documented and the exact nature of the environmental factors is unknown. If the original question and the experimental data are inadequate then no matter what the power of the program or statistical analysis no useful conclusions will be obtained. The current techniques follow a simplistic approach to the gene networks associated with a complex disease such as diabetes. The true situation is far more complex than identifying single genes in a causal-reactive sense because these genes exist in a larger genetic network and will be subject to both positive and negative feedback control.

It could well be that the answer to the question "Is there an understandable genetic linkage to Type 2 diabetes", is NO, but if the question were to be rephrased as, "Is there a genetic linkage to problems with the production of enzymes in a metabolic pathway that gives rise to the symptoms of hyperglycaemia" then the answer would be YES. If we consider the scenario where we may have at least 20 enzymes that are essential in the insulin regulation cascade or in different parts of the glucose metabolic pathway in the body or in the production of hormones relevant to the metabolism of glucose, then any disruption to the transcription of any of the enzymes will reduce the effectiveness of the glucose homeostasis mechanism in the body. The impaired synthesis of one, or any combination of these enzymes with or without the added stress of changing lifestyles will then give rise to the symptoms of type 2 diabetes. This scenario could explain why we see suggested linkages on almost a third of the total number of chromosomes but individually no concrete association emerges.

Future research will require more knowledge of the biochemistry of the disease and how the variation in susceptibility alleles affects individuals. In order to understand how the genetic variation in populations contributes to the conditions leading to the symptoms of Type 2 diabetes, the analysis of much larger sample

populations with detailed knowledge not only of the pedigree genetics but also of the lifestyles of the subjects within the population will be required. In the light of the rapid development of facile biochemical tests for metabolites etc in the blood it would appear to make more sense to identify which enzyme/enzymes are deficient in patients showing the symptoms of Type 2 diabetes. It would then be possible to create populations that show a similar deficiency and by using linkage analysis find the candidate genes for the production of those specific enzymes. The present trend of seeking a general genetic linkage for such an ill-defined complex "symptom" as Type 2 diabetes could well be a complete waste of time.

References

1. O'Rahilly S, Barroso I, Wareham NJ. Genetic factors in type 2 diabetes. *Science* 2005; 307: 370-373.
2. Gloyn AL. The search for type 2 diabetes genes. *Ageing Res Rev* 2003; 2: 111-127.
3. van Tilburg J, Haeflgen TW, Pearson P. Defining the genetic contribution of type 2 diabetes. *J Mol Genet* 2001; 38: 569-578.
4. Gerich JE. The genetic basis of type 2 diabetes, impaired insulin secretion versus impaired insulin sensitivity. *Endocrine Rev* 1998; 19: 491-503.
5. Fujimoto WY. The importance of insulin resistance in the pathogenesis of type 2 diabetes. *Am J Med* 2000; 108: 9-14.
6. Karp CL, Grupe A, Schadt EE, Cuomo P, Germer S. Identification of complement factor 5 as a susceptibility locus for experimental allergic asthma. *Nat Immunol* 2000; 1: 221-226.
7. Schadt EE, Monks SA, Drake T, et al. Genetics of gene expression surveyed in maize, mouse and man. *Nature* 2003; 422: 297-302.
8. Mootha VK, Lindgren C, Laureta S, et al. PGC-1-alpha-responsive genes involved in oxidative phosphorylation are coordinately down regulated in human diabetes. *Nat Genetics* 2003; 34: 267-273.
9. Zhu L. An integrative genomics approach to the reconstruction of gene networks in segregating populations. *Cytogenet Genome Res* 2004; 105: 363-374.
10. Luscombe NM, Babu MM, Yu H, et al. Genomic analysis of regulatory network dynamics reveals large topological changes. *Nature* 2004; 431: 308-312.
11. Jiang C, Zeng ZB. Multiple trait analysis of genetic mapping for quantitative trait loci. *Genetics* 1995; 140: 1111-1127.
12. Schadt EE, Lamb J, Yang X, et al. An integrative genomics approach to infer causal associations between gene expression and disease. *Nat Genetics* 2005; 37: 710-717.
13. Romano J. Association of inflammation markers with impaired insulin sensitivity and coagulation activation in obese healthy women. *J Clin Endocrinol Metab* 2003; 88: 5321-5326.
14. Rosmond R, Chagnon M, Bouchard C, Bjorntorp P. Increased abdominal obesity, insulin and glucose levels in non diabetic subjects with a T290 polymorphism of

- the transforming growth factor beta1 gene. 2003; *Horm Res* 2003; 59: 191-194.
15. Goodarzi MO, Guo X, Taylor KD, et al. Lipoprotein lipase is a gene for insulin resistance in Mexican Americans. *Diabetes* 2004; 53: 214-220.
 16. van Tilburg J, Haeflén TW, Wijmenga C. Defining the genetic contribution of type 2 diabetes mellitus. *J Med Genetics* 2001; 38: 569-578.
 17. Slatkin M. Inbreeding coefficients and coalescent times. *Genet Res* 2001; 58: 167-175.
 18. Pritchard JK, Rosenberg NA. Use of unlinked genetic markers to detect population stratification in association studies. *Am J Hum Gen* 1999; 65: 220-228.
 19. Lander ES, Schork NJ. Genetic dissection of complex traits. *Science* 1994; 273: 1516-1517.
 20. Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996; 273: 1516-1517.
 21. Brown DL, Gorin MB, Weeks DE. Efficient strategies for genomic searching using the affected-pedigree-member method of linkage analysis. *Am J Hum Genet* 1994; 54: 544-552.
 22. Mahtani MM, Widen E, Lehto M, Thomas J, et al. Mapping of a gene for type 2 diabetes associated with insulin defect in Finnish families. *Nat Genet* 1996; 14: 90-94.
 23. Elbein SC, Hoffman MD, Teng K, et al. A genome-wide search for type 2 diabetes susceptibility genes in Utah Caucasians. *Diabetes*. 1999; 48:1175-1182.
 24. Hanson RL, Ehm MG, Pettitt DJ. An autosomal genetic scan for loci linked to type 2 diabetes and body mass index in Pima Indians. *Am J Hum Genet* 1998; 63: 1124-1132.
 25. Wolford JK, Hanson RL, Bogardus C, Prochazka M. Analysis of linkage disequilibrium between polymorphisms in the KCNJ9 gene with type 2 diabetes mellitus in Pima Indians. *Mol Genet Metab* 2001; 73: 97-103.
 26. Vionnet N, Hani EH, Dupont S. Genome wide search for type 2 diabetes susceptibility genes in French whites. *Am J Hum Genet* 2000; 67: 1470-1480.
 27. Wiltshire S, Hattersley AT, Hitman GA, et al. A genome-wide scan for loci predisposing to type 2 diabetes in a UK population. *Am J Hum Genet* 2001; 69: 553-569.
 28. Hegele RA, Sun F, Harris SB, et al. Genome wide scanning for type 2 diabetes susceptibility in Canadian Oji-Cree using 190 microsatellite markers. *J Hum Genet* 1999; 44: 10-14.
 29. Meigs JB, Panhuysen CI, Myers RH, et al. A genome-wide scan for loci linked to plasma levels of glucose and HbA in a community based sample of Caucasian pedigrees. *Diabetes* 2002; 51: 833-840.
 30. Harikawa Y. Genetic variation in the gene encoding Calpain 10 is associated with type 2 diabetes. *Nature Genet* 2000; 26: 163-175.
 31. Hsueh WC, St. Jean P, Mitchell BD, et al. Genome-wide and fine mapping linkage studies of type 2 diabetes and glucose traits in the old order Amish. *Diabetes* 2002; 52: 550-557.
 32. Xiang K, Wang Y, Zheng T, et al. Genome-wide search for type 2 diabetes/impaired glucose homeostasis susceptibility genes in the Chinese. *Diabetes* 2004; 53: 228-234.
 33. Halliburton R. *An Introduction to Population Genetics*. Pearson, 2004