THE UNIVERSITY OF CALGARY

The Role of the Insulin Gene Region in Type I Insulin-Dependent Diabetes Mellitus: Possible Interactions with Candidate Loci.

by

Jane Yumiko Nagatomi

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF MEDICAL SCIENCES

CALGARY, ALBERTA

NOVEMBER, 1995

© Jane Yumiko Nagatomi 1995

THE UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "The Role of the Insulin Gene Region in Type I Insulin-Dependent Diabetes Mellitus: Possible Interactions with Candidate Loci" submitted by Jane Yumiko Nagatomi in partial fulfillment of the requirements for the degree of Master of Science.

Dr. L.L. Field, Supervisor
Department of Medical Science

Dr. R. Bell
Department of Medical Science

Dr. C. Johnson
Department of Medical Science

Dr. F. Snyder
Department of Biochemistry/Molecular Biology

November 14, 1995

Abstract

Insulin-dependent diabetes mellitus (IDDM) is an autoimmune disorder which involves both genetic and environmental factors. Three loci (tyrosine hydroxylase [TH], insulin-like growth factor 1 receptor [IGF1R], and insulin receptor) were examined for linkage and association to IDDM. Strength of association between markers and IDDM, depending on second locus allele sharing, age of onset, and HLA DR4 status in affected siblings, as well as geographic origin of the families, was examined. Interactions between two marker loci were investigated by joint sharing analysis.

Significant linkage was not demonstrated for any locus. However, associations were seen for TH ($p=0.013$) and IGF1R ($p=0.022$). No compelling evidence for interactions between loci, nor for maternal/paternal dependent association, was observed. Contrary to previously published reports, no evidence for a paternally dependent DR4 association between TH and IDDM, nor an association between the insulin receptor and IDDM, was seen.

ACKNOWLEDGEMENTS

I would like to thank Dr. Leigh Field for her support and allowing me the opportunity to work in her lab. It was a rich and rewarding experience.

I would also like to thank Elzbieta Swiergala for her help with technical aspects of this work.
Both she and Martha Hughes contributed to the typing of the multiplex families prior to my starting this project.

I would also like to thank Rose Tobias for her help and many suggestions.

DEDICATION

I dedicate this thesis in loving memory of my mother.

I wish you could have been here.

I miss you always.

To my family: my Dad, my sister and my brother.

For their support (the laser printer) and for not falling asleep on me

when I told them of my project.

To my friends April Wettig, Janet Tong, Michele Goh,

and especially Brigitte Dion, Nathalie Jourdain,

Elzbieta Swiergala and Norma Schmill de French.

For their support and their faith that I could finish this.

Finally, to my cat.

For reminding me that HE was more important than the computer,

and that 4 a.m. was well and truly past time to go to bed!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFBAC	Affected Family Based Controls
APS	ammonium persulfate
BDA	British Diabetic Association Families
bp	base pair
BSA	bovine serum albumin
CDN	Canadian Multiplex Families
cM	centi-Morgan
DNA	deoxyribonucleic acid
EDTA	ethylene-diamine-tetraacetic acid
HLA	human leukocyte antigen
IAA	insulin autoantibodies
ICA	islet cell antibodies
ICSA	islet cell surface antibodies
IDDM	insulin dependent diabetes mellitus
IGF1R	insulin-like growth factor 1 receptor
INS	insulin
INSR	insulin receptor
kb	kilobase
low TE	10 mM Tris (pH 7.0), 1.0 mM EDTA

M	molar
min	minute
µL	microlitre
mL	millilitre
mM	millimolar
NA	North American Families
ng	nanogram
nm	nanometer
OD	optical density
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PH	Philadelphia families
pH	-log[H+]
RD	random diabetics
RFLP	restriction fragment length polymorphism
rpm	revolutions per minute
SDS	sodium dodecyl sulfate
sec	second
SSCP	single strand conformation polymorphism
TEMED	N, N, N', N', tetramethylethylenediamine
V	Volt

VNTR	variable number tandem repeats
vs	versus

1. Introduction

1.1 What is Diabetes?

Diabetes is a metabolic disorder in which the body is no longer able to fully utilize carbohydrates. The pancreas was found to be involved in diabetes in the 19th century when it was found that individuals with severe pancreatic damage often had diabetes. The body's inability to use its fuel results in the accumulation of glucose in the bloodstream which then spills over into the urine (hence diabetes mellitus for the sweetness of the urine). The glucose in the urine then results in the excretion of large quantities of water (polyuria) which in turn causes a great thirst in individuals (polydipsia). The excretion of large amounts of glucose then causes individuals to become extremely hungry (polyphagia) since they are excreting all of the body's fuel [Foster 1988].

Type I insulin-dependent diabetes mellitus (IDDM) is characterized by early and rapid onset typically before age 18, although adult onset is not uncommon (approximately 33%) [Laakso and Pyörälä 1985]. These individuals are insulinopenic due to the loss of 90-95% of the insulin producing pancreatic β cells and are thus reliant on exogenous insulin for continued survival. They are often prone to ketosis (build up of ketone bodies due to partial digestion of fatty acids) and often have anti-insulin antibodies as well as anti-islet cell antibodies in their circulation.

1.2 Complications

Complications of IDDM include reduced life expectancy. Prior to the discovery and isolation of insulin, patients with juvenile diabetes were not expected to live 1 year beyond diagnosis [Bliss 1982]. Since that discovery however, IDDM patients are able to partake in normal daily activities despite having a shorter lifespan, and complications which may arise later in their life. Patients require daily insulin injections and are susceptible to widespread tissue damage possibly leading to kidney failure, blindness, amputations and premature heart disease [Cornall 1993].

1.3 Treatment

Before the discovery of insulin, treatment of IDDM was decidedly unsuccessful and often rather harmful. Initially, treatment consisted of trying to replace the fuel the body was excreting by putting patients on high carbohydrate diets. However, this often only exacerbated the problem eventually leading to patient death. Gradually however, it became clear that the opposite should be done as the body could not handle the load of food it was given, so food intake was greatly reduced. Dietary treatment consisted of determining the maximum amount of calories the body could tolerate before showing glucosuria resulting in slow starvation of the patient. Current treatment for IDDM includes both dietary as well as insulin treatment [Bliss 1982].

1.4 Incidence

The incidence of IDDM varies with geography from a high in Finland at

28/100,000 per year to a low of 1/100,000 per year in Japan [Maclaren 1992]. The concordance rate of IDDM in monozygotic twins is approximately 30-40% [Kumar et al 1993, Palmer and McCulloch 1991] with increased concordance with earlier age of diagnosis. Discordance in monozygotic twins may reflect differences in T cell repertoires between the two twins [Davey et al 1994]. While IDDM is most common in juveniles, approximately 33-37% of IDDM diagnoses are after 19 years of age [Laakso and Pyörälä 1985].

1.5 Etiology

IDDM is a multifactorial disorder in which both genetic and environmental factors are involved. Sibling and twin studies show that twins are more likely to be concordant for diabetes than are siblings. Studies of monozygotic twins have shown a concordance rate of approximately 30-40% which suggests that there is a large environmental factor involved [Barnett et al 1981, A'Hern et al 1988]. Dizygotic twin studies show a concordance rate of approximately 10% [Barnett et al 1981].

The actual mechanism of pathogenesis is currently unknown but is thought to involve a self to non-self transformation of the immune system such that normal self antigens expressed by cells are seen as foreign. This results in T cell infiltration of the pancreatic islets (insulitis) eventually leading to antibody production against self antigens. Normally, self reactive T cells are clonally deleted in the thymus or become anergic (ie. unable to respond to immunological stimuli) in the periphery. In IDDM, these cells are either released into the periphery or become activated such

that they are able to stimulate B cells into producing insulin autoantibodies (IAA's), islet cell antibodies (ICA's) and islet cell surface antibodies (ICSA's). ICA's are a heterogeneous population of antibodies which recognize peptides of different sizes; one of the more common being against a 65 kDa peptide thought to be glutamic acid decarboxylase (GAD) [Hagopian et al 1993]. Other ICA's include antibodies to 33 kDa, and 38 kDa peptides [Honeyman et al 1993, Roep et al 1990, Tun et al 1994].

While IAA's and ICA's are common in newly diagnosed type I diabetics, they are not present in all patients [Doberson et al 1980]. In one study [Toguchi et al 1985], 35% of IDDM patients had ICA's compared to 2% of controls. As well, 5% of non-diabetic parents and 14% of non-diabetic siblings had ICA's present in their sera. In monozygotic twins, IAA's are often found in the non-diabetic twin despite having been discordant for a long period of time [Wilkin et al 1985]. Since IAA's are not always found in IDDM patients, three possible explanations exist for their function. IAA's and ICA's may result from a predisposition to IDDM, they may arise due to cross-reactivity with foreign peptides due to environmental insults (eg. viral infection) or they may result as a consequence of pancreatic damage [Doberson et al 1980].

Other non-genetic factors also play a role including early exposure to bovine serum albumin, as well as viral infections. It is thought, particularly with respect to viral infections, that certain antigenic peptides in the viral particle may cause autoreactive antibody formation due to cross-reactivity. This will be discussed in greater detail in the environmental influences section (Section 1.5.2.1).

1.5.1 Genetic Susceptibility

Genetically, IDDM was first shown to be influenced by the HLA region (IDDM1) with strongest association to the class II region including the DR3 and/or DR4 haplotypes. 95% of diabetics are DR3 and/or 4 positive compared to approximately 50% of the general population. An association was shown in 1984 between the insulin gene region (IDDM2) [Bell et al 1984] specifically the insulin 5' variable number of tandem repeats (VNTR) region and IDDM. It was seen that the class I allele at the VNTR was present at a frequency of 0.83 in IDDM patients compared to a frequency of 0.75 in normal control populations. However, actual linkage between the insulin gene and IDDM was never shown [Cox et al 1988]. Recent studies have shown that at least another 6 genes are involved in human IDDM. Linkage was shown between IDDM and a chromosome 15 marker D15S107 (IDDM3) [Field et al 1994] as well as chromosome 11q (IDDM4) [Field et al 1994, Hashimoto et al 1994] and chromosome 6q (IDDM5) [Davies et al 1994], chromosome 2q near HOX D8 (IDDM7)[Owerbach and Gabbay 1995, Copeman et al 1995], and on chromosome 6q below IDDM5 (IDDM8) [Luo et al 1995]. IDDM6 has yet to be published. Currently, over 10 susceptibility loci have been discovered in the NOD (non-obese diabetic) mouse [Ghosh et al 1993].

1.5.1.1 Autoimmunity

The actual initiating mechanism of the immune response in IDDM is not known. The destruction of the pancreatic β cells, however, is thought to involve the

CD4+ helper T cells and the CD8+ cytotoxic T cells, eventually resulting in B lymphocyte involvement and the production of IAA's and ICA's [Yagi et al 1992].

Both CD4+ and CD8+ T cells are required for the immune response [Yagi et al 1992]. While CD4+ cells are able to enter the pancreas, they alone cannot mediate pancreatic β cell destruction. Conversely, CD8+ T cells cannot enter the pancreas, but are able to cause cellular destruction. Thus, it is thought that the CD4+ T cells first infiltrate the islets, secrete a signal that allows the CD8+ T cells to enter the pancreas, eventually resulting in cellular damage. Studies show that concomitant transfer of both CD4+ and CD8+ T cells are required to transfer diabetes to healthy neonatal NOD mice [Yagi et al 1992]. Antibody production is thought to occur secondarily to pancreatic β cell damage [Bach 1991, Lernmark et al 1991]. Another mechanism by which the autoimmune reaction could occur would be by abnormal class II expression by the islet β cells. One study reported the induction of HLA class II molecule expression on the islet β cells by interferon-γ [Pujol-Borrell et al 1989]. Thus, the pancreatic β cells may themselves cause the activation of the CD4+ T cells by acting as antigen-presenting cells by presenting the islet cell antigens directly to the T cells in the context of HLA class II resulting in the autoimmune response.

1.5.1.2 The HLA System

The earliest studies of the HLA region involved the class I antigens. An association between IDDM with HLA class I molecules, particularly HLA B8 and

B15 haplotypes was found [Nerup 1974]. Also, from this data, it was determined that type I and type II diabetes were in fact genetically distinct diseases, since the latter showed no HLA associations.

A stronger association was then found with HLA class II alleles DR3 and DR4. An excess of DR3/DR4 heterozygotes was seen in the diabetic population versus the normal population [Rubinstein 1991]. This increase in DR3/DR4 heterozygotes may occur only in the first born affected child in multiplex families [Rubinstein et al 1977]. It was found that DR3 and/or DR4 was present in 90-95% of IDDM patients, while it was seen in only 50% of controls [Field 1988]. Initially inheritance of the HLA-associated susceptibility was thought to involve a recessive locus with reduced penetrance [Rubinstein et al 1977]. However, analysis of the frequency of homozygous DR3 and DR4 IDDM compared to DR3/DR4 heterozygotes rejected this hypothesis and instead suggested that either multiple alleles or multiple loci were involved [Rotter et al 1983]. Inheritance of the DR4 allele was examined and found to be transmitted in a dominant fashion, with DR3 appearing to increase DR4 susceptibility such that they acted synergistically [Macdonald et al 1986]. The HLA DR2 allele is generally under-represented in the diabetic population and may therefore be protective against type I diabetes.

The HLA class II molecule is a heterodimer consisting of an α chain (M_R 32 000) and a β chain (M_R 29 000) which together bind foreign antigens [Erlich et al 1990, Nepom 1990]. HLA class II molecules are present on B cells, monocytes, macrophages, and dendritic cells [Pujol-Borrell et al 1989]. Tight linkage

disequilibrium exists across the HLA region such that extended haplotypes tend to be inherited as a single genetic unit, with certain combinations occurring in the genetic population more/less frequently than expected by chance.

Molecular studies have since determined that the HLA associations with IDDM susceptibility rather than being due to HLA-DR, may instead be due to the DQ locus [Gill and Haskins 1993]. An arginine residue at position 52 of the DQα chain is thought to encode for susceptibility. At position 57 of the DQβ chain molecule, aspartic acid is thought to confer resistance, whereas non-asp amino acids (eg. serine or alanine) are thought to encode for susceptibility [Morel et al 1988]. It may be the charge of the amino acid at position 57 which confers susceptibility [Nepom 1989]. However, this negative association of the asp-57 haplotype with IDDM is not seen in the Japanese IDDM population, since in this group asp-57 is generally positively associated with susceptibility [Jenkins et al 1992]. Thus, a different mechanism may be involved in IDDM susceptibility in the Japanese versus Caucasian population. Sheehy et al 1989 has found that the HLA susceptibility to IDDM may best be defined by a combination of HLA DR and DQ haplotypes. They found that the DR4 allele alone could not define a susceptibility as there were certain subtypes of DR4 which did not confer susceptibility to IDDM. Instead they suggest that particular DR4 subtypes in conjunction with DQ3.2 lead to susceptibility [Sheehy et al 1989].

In summary, HLA susceptibility loci include DQA1, DQB1, DRB1 and probably DPB1 [Sheehy 1992]. The HLA region is thought to encode less than 50%

of the genetic component of susceptibility to IDDM. The overall increased risk to siblings of IDDM patients is 15 x the population risk (approximately 6% of siblings of IDDM patients become diabetic compared to a population frequency of approximately 0.004. Thus 0.06/0.004=15). A total of 7.3% of affected siblings share no HLA alleles by descent compared to an expected value of 0.25. Thus, when estimating the HLA portion of the risk in siblings, a value of approximately 3.42 (0.25/0.073) is obtained. When one then estimates the possible risk due to a non-HLA locus (or loci) a value of 4.39 (15/3.42) is obtained, which suggests that most of the genetic susceptibility to IDDM occurs in a non-HLA dependent manner. [Risch 1987]. This is a maximal estimate for non-HLA genetic susceptibility, since it is theoretically possible (although no supporting evidence exists), that some of the non-HLA susceptibility is due to familial environmental factors.

1.5.1.3 The Insulin Gene Region

Since more than 50% of the genetic component of susceptibility to IDDM is thought to be non-HLA dependent, other candidate loci were sought. The insulin gene was considered a candidate locus in susceptibility to diabetes due to the specific expression of insulin by the β cells of the Islets of Langerhans. The insulin gene region consists of three genes in close proximity; tyrosine hydroxylase (TH), insulin (INS) and insulin-like growth factor II (IGF2). These three genes together span approximately 45 kb of DNA with TH and INS approximately 2.7 kb apart and INS and IGF2 separated by approximately 2kb [Lucassen et al 1993]. (See Fig. 1)

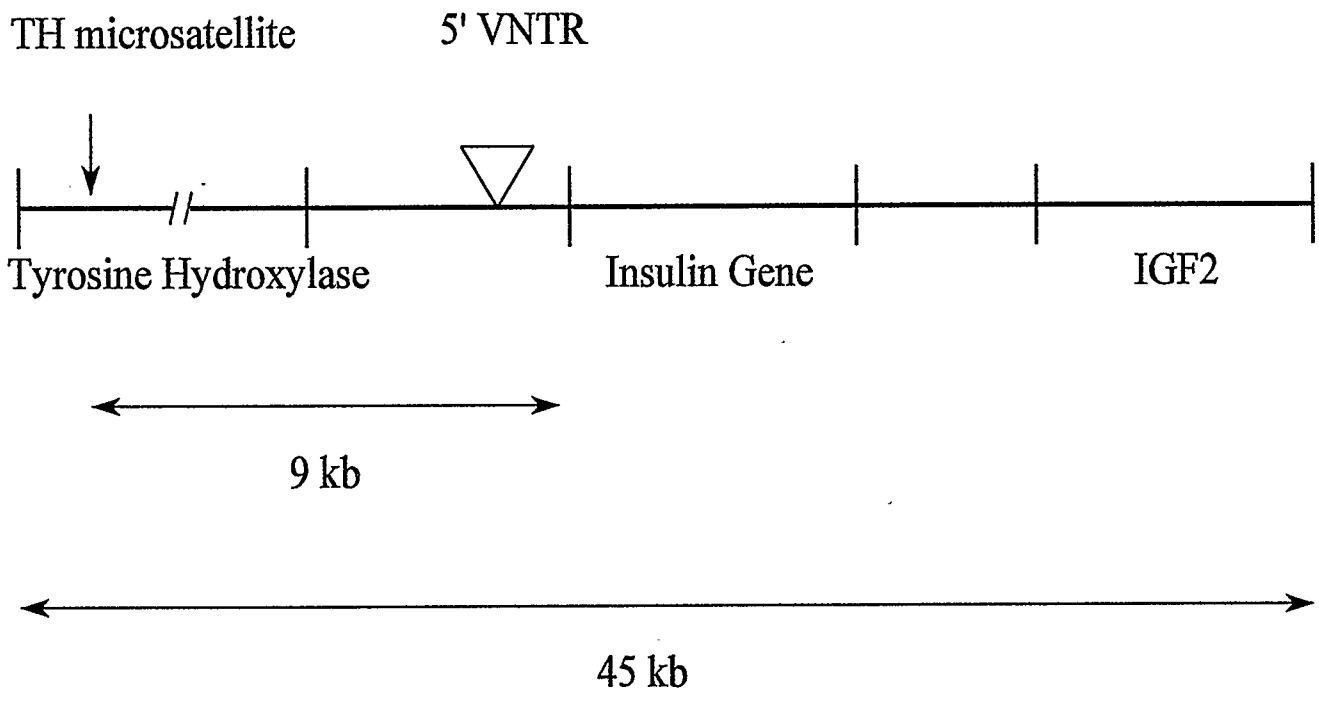


Figure 1. The insulin gene region and associated markers. (Figure not to scale).

Restriction fragment length polymorphism (RFLP) studies of the insulin gene region have shown that there is an IDDM association with the short, or class I, allele at the variable number of tandem repeats (VNTR) locus (TCTGGGGAGAGCGG) [Docherty 1992] located -363 upstream of INS (see Fig. 1) [Bell et al 1981]. The VNTR has three groups of alleles of approximately 40, 95 and 175 repeats encoding class I, II and III alleles respectively. One *in vitro* study has shown that the class I allele can form an unusual quadruplex structure that allow interactions via G residues on the top strand [Hammond-Kosack et al 1992]. The actual function of this structure is unknown. However, it may be that the small number of repeats of the class I allele of the VNTR may influence DNA folding possibly resulting in susceptibility to IDDM through altered transcription [Bell et al 1981]. A recent study looking at the class I VNTR has found that *in vitro* the class I VNTR binds poorly to the transcription factor Pur1 [Kennedy et al 1995]. Also, they have found that the class I VNTR tend to display reduced transcription of the insulin gene [Kennedy et al 1995], although others [Bennett et al 1995] have suggested that class I VNTR alleles are associated with increased insulin gene transcription. How the INS VNTR would affect susceptibility is currently not known but may be due to altered transcription for one of the genes in this region (ie. TH, INS or IGF2).

According to one study, the association between the insulin gene and IDDM has been narrowed down to a 4.1 kb region spanning the insulin gene and its regulatory sequences, excluding TH and IGF2 [Lucassen et al 1993]. This study looked at the presence of polymorphisms within the insulin gene region and

compared the frequency of these polymorphisms between IDDM and control populations. They found that within a 4.1 kb region, the IDDM group was very likely to be homozygous for the presence of a given polymorphism whereas the control group showed lower homozygosity (approximately 80% in IDDM compared to 55% in controls). Also, they found that in HLA DR4+ IDDM diabetics, the insulin polymorphisms showed greater association with IDDM than in non-DR4+ diabetics and may therefore be indicative of interactions between the HLA and INS region susceptibility genes.

Another study reported an association between insulin and HLA DR4 inheritance in multiplex families, where the IDDM associated insulin region alleles were transmitted preferentially to children with paternally inherited DR4+ [Julier et al 1991]. However, these results have not been verified in other studies [Cornall 1993, Field 1991].

Studies done on the risk of transmission from type I diabetic mothers versus type I diabetic fathers to their offspring found that the DR3 allele was inherited maternally in 62% of patients with non-diabetic patients compared to 38% which were transmitted paternally [Deschamps et al 1990]. Also, children of male type I diabetics were about 4x more likely to have IDDM than children of type I diabetic mothers [Rjasanowski et al 1993, Warram et al 1984]. Possible explanations for this difference in transmission include preferential loss of IDDM fetuses in utero, failure for such fetuses to implant, or a protective effect from maternal education of the fetal immune system [Clerget-Darpoux et al 1991, Warram et al 1984]. Since the IGF2

gene is immediately 3' to insulin, genomic imprinting may play a role in IDDM by influencing susceptibility depending on whether the diabetogenic allele was inherited maternally or paternally. IGF2 is known to be imprinted and this chromosomal region is thought to be involved in Beckwith Wiedemann Syndrome, an imprinted disorder in which both islet cell hyperplasia and hyperinsulinemia occur [Julier et al 1991].

1.5.1.4 IGF1R

A recently discovered susceptibility locus, IDDM3, has recently been found on chromosome 15q26 near the microsatellite marker D15S107 [Field et al 1994]. Subsequently, this locus has been detected in an independent data set [Luo et al 1995] Studies also indicate that affected siblings tend to share less HLA haplotypes when sharing more haplotypes at D15S107. The insulin-like growth factor 1 receptor (IGF1R) is a gene close to D15S107 (approximately 3.5 cM away) and may therefore be a candidate for the gene involved in susceptibility to IDDM for this group of patients.

1.5.1.5 The Insulin Receptor and Other Loci

The insulin receptor (INSR) has been reported as being associated with IDDM [Raffel et al 1990]. However, little has been published on this area with respect to disease association. Thus, the insulin receptor and its possible role with the insulin gene region will be examined.

Recently, new IDDM susceptibility loci have been found in humans. These include IDDM4 on chromosome 11q near FCER1 [Field et al 1994, Davies et al 1994, Hashimoto et al 1994], IDDM5 on 6q near the ESR gene [Davies et al 1994], as well as IDDM7 on 2q31 near HOX D8 [Owerbach and Gabbay 1995, Copeman et al 1995] and IDDM8 on 6q [Luo et al. 1995].

1.5.1.6 Microsatellite and SSCP Markers

Microsatellites are simple sequence repeats of DNA consisting of approximately 15-30 repeats. Most of these repeats are CA dinucleotide repeats although trinucleotide and tetranucleotide repeats are not uncommon. There are approximately 50 000 - 100 000 microsatellite repeats throughout the human genome, which if evenly spaced would be approximately every 30-60 kb of DNA sequence. These sequences are generally highly polymorphic and are inherited in standard Mendelian fashion. Thus, they are very useful for linkage analyses due to the high information content of the marker [Weber and May 1989].

Single strand conformation polymorphisms (SSCP's) are markers which identify sequence changes or deletions in DNA by electrophoretic shifts. DNA in a single-stranded form will fold into a conformation which is stabilized by intrastrand interactions. In cases of a base change, this conformation presumably changes which can be observed as an electrophoretic shift on non-denaturing polyacrylamide gels [Orita et al 1989] .

The primer sequences for amplifying markers to be used in this study were

obtained from published sources including: Lucassen et al 1993 - TH tetranucleotide repeat, Meloni et al 1992 - IGF1R trinucleotide repeat, Xiang et al 1991 - INSR dinucleotide repeat, Poduslo et al 1991 - IGF1R SSCP (2 bp deletion).

The IGF1R microsatellite is not very polymorphic, so a second marker was used to try to increase the information available for this locus. A single strand conformation polymorphism (SSCP) marker was used and haplotyped with the IGF1R microsatellite in order to increase informativity. SSCP markers can also be highly polymorphic and therefore very useful for genetic linkage studies.

1.5.2 Environmental Influences

The low concordance rate between monozygotic twins suggests that a large part of the susceptibility to IDDM is non-genetic. These non-genetic factors may include antigenic bovine serum albumin [Kostraba et al 1993, Robinson et al 1993], as well as viral infections from mumps, congenital rubella, Coxsackie B, and Cytomegalovirus [Banatvala et al 1985, Bruserud et al 1985, Pak et al 1988]. Antibodies to viral antigens have been reported to be increased in IDDM patients [Banatvala et al 1985]. These antigens may mimic self antigens such that, following infection in genetically susceptible individuals, an immune response mounted against the foreign antigens by T cells (primarily CD8+ that are involved in viral immunity) results in damage to islet β cells that express cross reactive antigens [Rubinstein 1991]. This immune reaction eventually leads to B lymphocyte involvement resulting in the production of self reactive antibodies leading to mass destruction of

the islet β cells. Alternatively, it has been suggested that some viruses can directly attack the pancreatic β cells through the inflammatory response [Yoon 1991].

1.5.2.1 Viral Influences

Viral infections are thought to possibly play a role in IDDM susceptibility due to the presence of antibodies to viral peptides in many newly diagnosed diabetics. In a study of children immunized with live mumps vaccine or who were infected with mumps virus, the presence of ICA's was investigated [Helmke et al 1986]. Of 127 children with severe mumps infection, 21 were ICA+ yet only one developed IDDM. Thus, while autoreactive antibodies developed, they were not indicative of susceptibility to IDDM.

In contrast, a recent study in NOD mice has shown that diabetogenicity may be due to amino acid position 776 of the polyprotein in the Encephalomyocarditis virus [Bae and Yoon 1993].

1.5.2.2 Bovine Serum Albumin

Recently, studies have focussed on the possible role of bovine serum albumin as a possible trigger in the pathogenesis of IDDM. One study found that in genetically susceptible individuals, early exposure to cow's milk (before 3 months of age) was associated with increased susceptibility to IDDM. Since concomitant exposure to solid foods was also observed in their sample population, they could only conclude that it was the early exposure to both cow's milk and solid foods which

accounted for this increased susceptibility [Kostraba et al 1993]. Another study by a Finnish group found that increased risk of IDDM occurred when dairy products were introduced prior to 2 months of age compared to after 2 months of age. They found that this increased risk was independent of year of birth, mother's age, education, child's birth order or birth weight. Also, they found that an inverse relation was found between overall duration of breast-feeding as well as the duration of exclusive breast-feeding and risk of IDDM [Virtanen et al 1993].

The primary antigenic sequence appears to be in the ABBOS peptide sequence which is a 17 amino acid sequence from aa 152-168 of the bovine serum albumin protein. This sequence is unique in the bovine form of the protein compared to homologous sequences in rat, mouse and humans. Anti-ABBOS antibodies appear to be present in many type I diabetics, much more than anti BSA antibodies. The ABBOS peptide appears to associate with a 69 kDa protein on the pancreatic β cell surface. This 69 kDa protein is induced by interferon (presumably following induction of the immune system by the BSA-ABBOS peptide) and slowly disappears following diagnosis of IDDM, most likely due to the auto-destruction of the pancreatic β cells which carry the 69 kDa protein [Karjalainen et al 1992]

1.6 Purpose

The purpose of this study is to clarify the role of the insulin gene region (IDDM2) in IDDM susceptibility using both random diabetics as well as multiplex diabetic families (families in which more two or more children are affected).

The specific aims of this study are:

1. to examine the role of the insulin gene region in IDDM susceptibility using the closely linked TH microsatellite as a marker. Despite strong evidence for an association between the insulin VNTR and IDDM, reports of linkage between the two have been inconsistent. Thus, given the increased informativity of the TH marker, we will attempt to show linkage between the insulin gene region and IDDM.
2. to determine whether any disease association can be observed between the chromosome 15 disease susceptibility gene (IDDM3) and the IGF1R microsatellite marker. IGF1R is in close proximity to D15S107 (the marker which showed significant evidence for linkage to IDDM3) and thus may be considered a candidate locus for IDDM3.
3. to determine whether any association exists between the insulin receptor gene and IDDM as previously reported using a microsatellite marker for the insulin receptor locus.
4. to determine whether different strengths of association can be observed in one marker given increased (or decreased) sharing in affected siblings at a second locus. For examples, using the D15S107 sharing information, associations between IDDM and the TH marker will be determined for families in which affected siblings share >50% of D15S107 alleles, and families in which affected siblings share ≤50% of D15S107 alleles. Marker sharing at other loci will be done for all markers including subdividing

families on the basis of HLA.

5. to examine the possibility of interactions between two marker loci of interest by analysis of joint sharing of marker alleles in affected sibling pairs. Since IDDM is clearly a multilocus disorder, interactive analyses of multiple loci would likely be more informative than single locus analyses.

6. to determine whether (as previously reported) there is a stronger association of IDDM with the insulin gene region markers in diabetics with paternally derived DR4.

7. to determine if microsatellite markers can be used to demonstrate linkage disequilibrium (ie. association), since some investigations have suggested their high mutation rate may make it impossible to use them for detecting associations with disease loci.

2. Materials and Methods

2.1 Random Patients and IDDM Multiplex Families

A total of 110 random IDDM diabetic patients and 250 multiplex IDDM diabetic families (1159 individuals) were studied. The random diabetics and multiplex diabetics were previously typed for HLA and the information obtained were then used in this study. (See Section 2.9 Typing Done by the Candidate). The multiplex families were also subdivided into two groups consisting of the British (BDA - 96 families) families and the North American families (HBDI and Canadian 155 families).

2.2 Blood Samples

Blood samples from the 110 random diabetics were obtained from newly diagnosed IDDM patients (10-15 mL) attending the Alberta Children's Hospital Diabetes Clinic after informed consent was given by the parents. Similarly, the Canadian multiplex families' (ie. families in which two or more children are affected - 25 families) blood samples were also obtained following informed consent. Blood was collected by venipuncture into Vacutainer ACD (acid citrate-dextrose) tubes (Becton Dickinson Ltd.) containing 1 mL of anticoagulant. The multiplex blood samples were shipped by courier to the lab in Calgary from various centres across Canada.

2.3 DNA Sources

DNA from multiplex family members were purchased from the British Diabetic Association - Warren Repository (96 families) and the Philadelphia Human Biological Disease Interchange (HBDI - 130 families). Other DNA's were extracted from blood samples collected across Canada (25 families).

2.4 DNA Extraction

DNA extraction was carried out in Dr Field's lab by other technical staff from blood samples using the following method which is modified from Miller et al 1988. Each ACD tube contains approximately 5 mL of whole blood plus 1 mL anticoagulant. The ACD tubes were first centrifuged at 2000 rpm in an IEC Centra-7R bench centrifuge at room temperature for 15 minutes (mins). The plasma layer was then transferred into 1.5 mL screw-top tubes and stored at -80°C for future studies. The buffy coat from one tube was then removed (approximately 1 mL) for HLA typing. The remaining cells were then decanted into a 50 mL conical tube and red blood cell lysis buffer (155 mM NH_4Cl, 10 mM $KHCO_3$, 1 mM Na_2EDTA, pH 7.4) was added to a total volume of 50 mL. The solution was then mixed and kept on ice until lysis was complete (approximately 15 mins when the solution went dark). The tubes were then centrifuged at 1500 rpm for 10 mins and the supernatant decanted. The pellet was then resuspended in 5-10 mL of red cell lysis buffer and recentrifuged at 1500 rpm for 10 mins. The supernatant was again removed and the white blood cell pellet was then resuspended in 3 mL of nuclei lysis buffer (10 mM Tris-HCl, 400 mM NaCl, 2 mM Na_2EDTA, pH 8.2). The cells were left to digest

overnight at 37°C with 0.2 mL 10% SDS and 0.5 mL proteinase K solution (50μL 1 mg/mL proteinase K solution, 50μL 10% SDS, 2 μL 0.5 M Na_2EDTA, 398 μL ddH_2O).

The next morning, 3 mL H_2O and 3 mL saturated NaCl (approximately 6M) was added. The solution was mixed and then centrifuged at 3000 rpm for 30 minutes. (If the proteins did not centrifuge down then an additional 3 mL H_2O and 3 mL of saturated NaCl was added and the solution respun at 3000 rpm for 30 mins). The supernatant was then transferred to a new 50 mL tube and 2 volumes of 95% ethanol was then added. The tubes were inverted to precipitate the DNA and the DNA then spooled out onto a bent glass pasteur pipette which had been heat sealed. The DNA was then washed with 70% ethanol and transferred to 1.5 mL eppendorf tubes. The DNA was then resuspended in approximately 500 μL of low TE (10 mM Tris-Cl, 1 mM Na_2EDTA, pH 8.0) and rotated overnight at 4°C.

2.5 DNA Quantification

DNA samples were diluted to 1:100 in 1 mL and then quantified by measuring the absorbance of light on an Amersham GeneQuant spectrophotometer at 260 nm in a 0.75 mL cuvette with a 1 cm light pathway. The ratio of the absorbance at 260/280 indicated the presence of any contaminating proteins in the solution with a ratio of 1.8-2.0 being desirable. An optical density (O.D.) of 1.0 is approximately equal to 50 μg/mL ds DNA.

2.6 HLA Typing

HLA typing for class I and class II antigens (A, B, C, and DR) were done in Dr. Field's lab by other technical staff for the Canadian multiplex families and the random diabetics by the following method. The 1 mL buffy coat that was isolated during DNA extraction was transferred to a 15 mL plastic test tube. 5-6 mL of phosphate-buffered saline - 0.6% sodium citrate (PBS-Na citrate) at 4°C was added and carefully mixed and centrifuged at 1200 rpm (no brake) for 10 mins. The supernatant was then removed and the cells washed with cold PBS and recentrifuged. The supernatant was again removed and the cells resuspended in PBS-Na citrate (cold) and transferred to a glass tube (13x100 mm disposable culture tube approx. 5-6 mL total volume). The resulting mix was then put on ice for 5 mins. 100 μL of dynabeads (HLA Prep II Dynabeads, Dynal Inc.) were added and mixed by gently rolling the tube. The tube was kept on ice for no more than 5 mins. and then placed in a rare earth magnet separator (Dynal MPC-6 magnetic particle concentrator - 6 tubes) to allow dynabead rosetted CD4+ cells to collect on the side of the tube (1-1.5 mins). The supernatant was then removed for class I HLA typing. The cells were washed 3 times with PBS-Na citrate (the first supernatant was kept for class II typing as well) then resupended in 200 μL of RPMI 1640 media (Gibco BRL) with glutamine and 15% fetal calf serum. The cells were then plated out on the HLA class II typing trays (GenTrak, Inc. and Canadian Red Cross Society National Laboratory) 1 μL per well with a Hamilton Syringe. The cells were then incubated at 37°C for 0.5 to 1 hour. 1 mL of rabbit complement was mixed with 20 μL acridine orange and

ethidium bromide stock solution (15 mg acridine orange and 50 mg ethidium bromide dissolved in 1 mL 95% ethanol mixed with 49 mL PBS) and 5 μL was then added to each well. The trays were then incubated in the dark for 1 hour and then read under a fluorescent light microscope. Simultaneous reading of viable and non-viable cells were done with cell death being indicative of the presence of the given HLA antigen.

The CD8+ cells were treated similarly to the non-CD8+ cells. The supernatant that was collected was washed with PBS, spun at 1200 rpm and the cells redissolved in 5 mL PBS. 3 mL of ficoll-paque was layered at the bottom of the test tube and was then centrifuged at 1200 rpm. The lymphocyte layer was then removed and washed three times with PBS and resuspended in 200μL RPMI (with glutamine and fetal calf serum). The cells were then treated as the non-CD8+ cells and plated out onto class I typing trays. All HLA typing was done using two different sets of trays for both class I and class II typing.

HLA typing results for the BDA and PH families were provided by the respective repositories.

2.7 Microsatellite and SSCP Polymerase Chain Reaction (PCR)

The polymerase chain reaction is a method to rapidly amplify a small segment of DNA (<2kb) between 2 primers located 3' and 5' to the region to be amplified. Primers are generally 15-25 nt in length to ensure specificity to one region in the genome. A thermostable DNA polymerase from the *Thermus aquaticus*

bacterium (Taq polymerase) is used so that the polymerase does not become inactive after DNA denaturing. The primers used were as follows:

Marker	Sequence
TH-1[1]	GGGTA TCTGG GCTCT GGGGT
TH-2	GGTCA CAGGG AACAC AGACT C
IGF1R SSCP-R[2]	GAGAC AGCTT CTCTG CAGTA
IGF1R SSCP-L	TCCGG ACACG AGGAT TCAGC
IGF1R microsatellite-F[3]	GCTGA GGGAG GAGGC GGC
IGF1R microsatellite-R	GGCGA GGGGC AGAAA CGC
INSR-R[4]	ATTGC TGCAT ATGCA GACAG
INSR-L	TGCAG CCGTG TGACT TACAG

The reaction mix for the PCR's varied according to the primers used. PCR's were done in 96 well microtitre plates on an Ericomp Twin-Block Machine with a temperature sensor. TH microsatellite PCR's were carried out in 15 μL reactions containing 25-50 ng dried DNA, 0.09 μL 100mM primers, 1.5 μL 10x GIBCO PCR buffer (200mM Tris-HCl pH8.4, 500mM KCl), 1.5μL 200 μM dNTP's, 0.3 μL 50

---

[1]Lucassen private communication based on PCR's in Lucassen et al 1993.

[2]Poduslo et al 1991

[3]Meloni et al 1992

[4]Xiang et al 1991

mM $MgCl_2$, 0.1μL P-32, 0.5 U Taq polymerase and ddH_2O. INSR and IGF1R microsatellite reactions were done in 15 μL reactions containing the same as the for TH save that 3 μL of 5x buffer (1M Tris-Cl pH 8.0, 1 M KCl, 170 μg/mL BSA, 1.5 mM $MgCl_2$, 12.5 μL Tween-20, 12.5 μL NP-40, ddH_2O to 5mL) was used. The IGF1R-SSCP PCR's were done in 10μL reactions containing 25-50ng DNA, 0.3μM primers, 20μM dNTP's, 0.1μL P-32, Gibco-BRL Taq Buffer, 0.5U Taq polymerase, 5% glycerol v/v, 3 mM $MgCl_2$ and ddH_2O. The reaction mixtures were covered with 15-30 μL of mineral oil. The different PCR conditions for the various markers are listed in Table 1.

Primers	Initial Denaturation	Denaturation	Annealing	Extension	Product Size (bp)
TH	94°C - 6 min	94°C - 1 min	55°C - 2 min	72°C - 1 min	105 - 125
INSR					130-148
IGF1R microsatellite	92°C - 7 min	92°C - 40 sec	63°C - 30 sec	72°C - 30 sec	90 - 99
IGF1R SSCP	94°C - 6 min	94°C - 1 min	58°C - 2 min	72°C - 1 min	2 bp deletion

Note: All PCR's were followed by a final extension at 72°C for 6 minutes.

Table 1. Primers and PCR Conditions.

2.8 Gel Electrophoresis

15 μL of loading dye (0.05% bromophenol blue, 0.05% xylene cyanol, 95% formamide) was added to the PCR product and 2 μL of the resulting mix was then loaded onto polyacrylamide gels and electrophoresed for size separation.

2.8.1 Denaturing Polyacrylamide Gels

TH, IGF1R microsatellite, and the INSR PCR products were separated on denaturing polyacrylamide gels. Each gel was made from 80mL 6% polyacrylamide, 7 M urea solution (150 mL 40% 19:1 acrylamide:bis-acrylamide solution - 380g acrylamide 20 g bis-acrylamide in 1L solution, 420 g urea, 200 mL 5 x TBE buffer, ddH_2O to 1L), 80 μL TEMED (N, N, N', N' Tetramethylethylenediamine) and 80 μL of 25% APS (ammonium persulfate) and the resulting solution was poured between glass plates of 33 x 40 cm and 33 x 42.5 cm dimension separated by 0.4 mm thick spacers and held together by 1" binder clips. A comb with approximately 56 wells was inserted into the gel and the gel allowed to solidify at room temperature. Upon polymerization, the comb and the bottom spacer were removed and the gel placed in an S2 electrophoresis apparatus by BRL Technologies Inc. 1X TBE (Tris-borate-EDTA) buffer was added to the buffer tanks, the samples were then loaded into the individual wells and an electrical current of 1000-1800 V was applied for 1-3 hours. Following electrophoresis, the plates were then separated and the gel transferred to 3M Whatman paper, covered with plastic wrap and dried at 80°C on a Model 583 GelDryer by BioRad. The gel was then placed into an autoradiograph cassette by

Fisher Scientific with Kodak X-OMAT XAR-5 X-ray film and left overnight to expose. The film was then developed in a Kodak M35A X-OMAT processor.

2.8.2 Non-denaturing Polyacrylamide Gels

The IGF1R-SSCP product was run out on native 5% polyacrylamide gels. Each gel was made of 10 mL 40% 19:1 acrylamide:bis-acrylamide solution, 16 mL 5 x TBE buffer, and 54 mL ddH_2O and polymerized by adding 80 μL of TEMED and 80 μL of 25% APS solution. The gels were poured as above for denaturing gels, and were run at 3 W for 16-24 hours in an S2 apparatus. Gel transfer and autoradiography were done as for denaturing gels.

2.9 INS VNTR Typing

The 5' insulin VNTR was typed in Dr. Field's lab by other technical staff for the random diabetics and the Canadian multiplex diabetic families by the method developed by Southern. Briefly, 3-5 μg of genomic DNA was digested by a restriction endonuclease (Taq or Sac I) and run out overnight on an agarose gel (0.6%). The DNA was then transferred by vacuum blot to a nitrocellulose membrane. The membrane was then hybridized with P-32 labelled probe (ie. for the INS VNTR). The membrane was then left to incubate overnight at 60°C. After washing excess probe of the membrane, it was then left to expose at -80°C with autoradiographic film [Elbein et al 1985].

2.10 Statistical Analysis

Statistical analysis of the data was done by two different methods; linkage analysis (both Lod score and affected sib pair methods) and association analysis. Lod score linkage analysis was carried out on a SUN-SPARC work station with the LINKAGE computer programs [Field et al 1994], while association analysis was done on an IBM compatible PC with the AFBAC (Affected Family BAsed Controls) computer program [Thomson 1995]. Significance of AFBAC analyses were calculated using the chi-square technique using Microsoft Excel v. 4.0 spreadsheet program. P values for the chi-square tests were determined using the spreadsheet program. Linkage analysis using the affected sibling pair method was performed manually (not using a computer program). Affected siblings were scored by determining whether maternally inherited and paternally inherited alleles were shared between the two siblings. The data was then tabulated and subjected to a standard chi-square test to determined if the sharing values obtained were significantly different from those expected from Mendelian segregation.

2.10.1 Linkage Analysis

Linkage analysis examines the actual inheritance of marker alleles from parents to offspring. Normally, there is a 25%, 50%, 25% probability of siblings sharing 2, 1, or 0 alleles according to Mendelian laws of segregation. However, given a gene involved in IDDM pathogenesis, affected siblings are likely to share alleles more often than the overall 50%. Affected sibling pair analysis looks at the

overall sharing of alleles to determine if more than 50% of alleles are shared for a given marker between two affected siblings; and if the difference is significant. The lod score method of linkage analysis is an additive, logarithmically based analysis which examines the probability of a marker being linked to a disease locus at a given θ (recombination frequency) compared to being unlinked (ie. θ=0.5) in families with multiple affected individuals. While linkage analysis often use a few large multigeneration, multi-affected families, the use of many small nuclear multiplex families can also be effective.

2.10.2 Association Analysis

Association analysis, in contrast to linkage analysis, does not examine sharing of alleles within affected families. Rather, it looks at the overall frequencies of marker alleles in affected versus normal populations. This method then identifies which allele(s) (if any) are in linkage disequilibrium with the gene contributing to the disease state. The AFBAC (Affected Family Based Controls) method is an association method in which control marker frequencies are derived from the affected families such that the affected and control populations are completely ethnically matched. This method divides the alleles into two subgroups; transmitted and non-transmitted. The transmitted group is composed of those alleles transmitted from a parent to an affected individual, while the non-transmitted group is comprised of those alleles not transmitted from parents to an affected individual for the family being studied. The non-transmitted group is then used for control frequencies for

comparison with the transmitted group [Thomson 1995, Field 1991].

The computer program runs three types of AFBAC analyses; simplex, multiplex, and double simplex analyses. The simplex analysis is the most basic which looks at the first affected child in all families and scores the paternal and maternal alleles as either transmitted or non-transmitted to that affected child, and tabulates the data over the entire data set. In this simplex analysis, the other affected children are ignored. (See Fig. 2). The scoring is separated into an overall analysis (ie transmitted and non-transmitted alleles) as well as by parental origin (ie. paternally derived transmitted and non-transmitted alleles and separately for maternally derived transmitted and non-transmitted alleles), and finally comparisons between parental transmitted alleles and between parental non-transmitted alleles. By comparing the two parental transmitted allele frequencies, possible imprinting effects can be detected. Thus, if disease expression occurred following allele transmission through the maternal lineage, the paternal transmitted frequencies would reflect population (control) frequencies while the maternal transmitted frequencies would show an increase in the disease associated allele frequency. Thus by comparing the paternal and maternal transmission frequencies, possible imprinting could be detected. Comparison of the non-transmitted allele frequencies between mothers and fathers should not reflect any differences generally, as they should both be a reflection of population frequencies. However, in the case of maternal-fetal interactions the non-transmitted allele frequencies between mothers and fathers may become significantly different from one another. In this case, the genotype of the

mother becomes important for disease in the offspring [Thomson 1995].

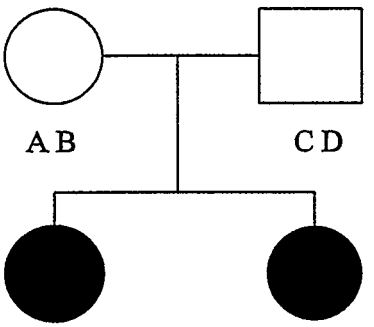
The multiplex analysis examines the first two affected children within a family and scores the alleles as transmitted or non-transmitted to those two children in a weighted manner. Thus, an allele present in both affected children would be scored 1 while an allele present in only one affected child would be scored 0.5. This method is the most accurate of the analysis methods in the sense that the non-transmitted allele frequencies accurately reflect general population frequencies. However, since it is a weighted analysis, the size of the control (non-transmitted alleles) group rapidly decreases such that a large sample size is needed to have meaningful results. (see Fig. 3).

The double simplex analysis is very similar to the simplex analysis except that both the first and the second affected children in the family are examined. In this case, alleles are scored transmitted or non-transmitted for the first affected child as in the simplex analysis and then the procedure is repeated for the second affected child. Also, the first and second affected child are treated completely independent of one another, effectively doubling the sample size.

While both the simplex and the double simplex analyses produce biased estimates of the control (non-transmitted) allele frequencies, they are biased in a conservative manner and as such are fully appropriate for testing transmitted vs. non-transmitted allele frequencies to detect marker associations with disease. This is because the non-transmitted frequencies are biased towards the transmitted allele frequencies and would tend to diminish any association between disease and the

marker. Since the simplex/double simplex analyses score the parental alleles as transmitted or non-transmitted irrespective of any other affected siblings in the family, it is more likely that possible disease associated alleles will be scored as non-transmitted even though they subsequently occur in affected individuals. Thus, in the multiplex analysis the non-transmitted allele frequencies accurately reflect general population frequencies while in the simplex/double simplex analyses the non-transmitted alleles frequencies will be conservatively biased.

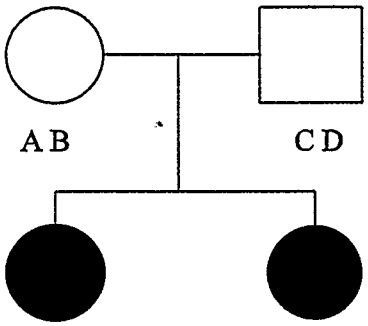
The AFBAC program allows rare alleles to be grouped together into an "other" category. This then reduces the likelihood of a false association due to minor differences in rare allele frequencies (ie increased importance of rare alleles) as well as making it possible to detect associations in the overall data set by reducing the degrees of freedom in the final chi-square analysis. In this study, rare alleles were lumped into the "other" category if the frequency of that allele was less than 0.05 in both the transmitted and non-transmitted group. Analyses were performed first to identify the rare alleles, and then re-run after appropriate lumping into the "other" category.



Case 1:	A C	A C
Case 2:	A C	A D
Case 3:	A C	B D

	Transmitted	Non-transmitted
Simplex Analysis (all cases)	A C	B D
Double Simplex Analysis		
Case 1	A C A C	B D B D
Case 2	A C A D	B D B C
Case 3	A C B D	B D A C

Fig. 2 AFBAC simplex and double simplex analysis in a fully informative family.



Case 1:	A C	A C
Case 2:	A C	A D
Case 3:	A C	B D

	Transmitted	Non-transmitted
Multiplex Analysis		
Case 1	A C	B D
Case 2	A 0.5 C 0.5 D	B
Case 3	0.5 A 0.5 B 0.5 C 0.5 D	-

Fig. 3 AFBAC multiplex analysis of a fully informative family.

In this study, AFBAC analyses were performed on the multiplex families using all three methods: simplex, multiplex and double simplex analyses. All analyses were further subdivided into maternally and paternally derived transmitted and non-transmitted frequencies. These parentally derived frequencies were analyzed to determine if an observed association was due to inheritance from a single sex parent or from both combined, as well as if any marked differences occurred between maternally compared to paternally inherited allele frequencies.

Families were then divided six different ways, including average age of onset of affected children in the family (greater than versus less than or equal to 11 years of age), by TH sharing (greater than versus less than or equal to 50% allele sharing between affected siblings), by D15S107 sharing (in lieu of IGF1R sharing), by INSR sharing, by HLA sharing, by DR4 positivity in both affected children, and by geographic location (British vs North American). Sharing criteria, age of onset and DR4 subgrouping was based on the average of all affected children in the family. For the IDDM3 region, the D15S107 marker was used in determining allele sharing instead of the IGF1R marker as it is much more informative than the IGF1R markers alone or combined in haplotypes. The D15S107 microsatellite is located approximately 3.5 cM from the IGF1R marker according to marker-marker linkage analysis in Dr. Field's lab.

2.11 Typing Done by the Candidate

Typing of the random diabetics for the PCR markers (TH, INSR, IGF1R, and

IGF1R SSCP) were done solely by the candidate. Other technical staff had previously completed the serological (HLA) typing and typing for the 5' insulin VNTR.

In the multiplex families, the CDN families were typed for HLA by other technical staff in Dr. Field's lab. Also, previously completed were the typings for the first 180 families for the IGF1R and the INSR microsatellite. Multiplex family typing done by the candidate include the TH microsatellite, the IGF1R SSCP, as well as completion of the IGF1R microsatellite and the INSR for the data set. Approximately 20 of the CDN families had also previously been typed for the INS VNTR by other technical staff within Dr. Field's lab.

3. Results

3.1 Random Diabetics

Random diabetics (89-109 individuals) were typed for the different PCR markers to determine allele frequencies in a diabetic population. Allele frequencies were determined by simple allele counting and the results are shown in Table 2.

Associations between IDDM and the genetic markers in the random diabetics were tested by using control frequencies determined from the multiplex families (using the AFBAC multiplex analysis; see Table 3) as well as published results where available. The multiplex family derived AFBAC control frequencies were used as control frequencies for the random diabetics as we had no random control group. Chi-square analysis of the markers shows no significant association between TH and IDDM in the random diabetics (χ^2=10.23, df=5. p=0.069) using the AFBAC control frequencies. An association is however, seen for the TH marker against published results [Puers et al 1993] with a χ^2=15.02, df=5, p=0.010. The remaining markers did not show any evidence of an association with IDDM using the AFBAC controls (IGF1R microsatellite χ^2=2.04, df=2, p=0.361, SSCP χ^2=0.17, df=1, p=0.680, INSR χ^2=0.09, df=4, p=0.999). Similarly, no significant difference was observed between the random diabetic allele frequencies and the diabetic allele frequencies determined by multiplex AFBAC analysis of the multiplex families for any of the markers (TH χ^2=9.47, df=5, p=0.092, IGF1R microsatellite χ^2=0.89, df=2, p=0.641, SSCP χ^2=0.46, df=1, p=0.498, INSR χ^2=2.11, df=4, p=0.629).

Since INS VNTR class I homozygotes are so common in IDDM, class I

homozygotes and non-homozygote diabetics were compared for IGF1R allele frequencies as a test for a possible TH interaction with IGF1R allele frequencies. However, no interaction was seen between the INS VNTR and IGF1R frequencies in random diabetics (χ^2=3.08, df=2, p=0.214). An analysis of the INSR also showed no interaction between the INS VNTR and INSR microsatellite frequencies (χ^2=1.53, df=4, p=0.821). Analysis of the TH allele frequencies in the INS VNTR class I homozygous diabetics compared to the non-class I homozygous diabetics revealed significant disequilibrium between the two loci (χ^2=39.48, df=5, p=0.000). Diabetic individuals who were homozygous for the class I allele had higher frequencies of allele 121 (0.339 vs 0.077 homozygous vs non-homozygous class I diabetics) and decreased frequencies of allele 124 (0.119 vs 0.538 homozygous vs non-homozygous class I diabetics). This agrees with the associations observed by Bennett et al 1995 when examining the class III allele. They found that there was striking disequilibrium with alleles 124, 117 and 109. In our data set, allele 109 was not significantly different between homozygous vs non-homozygous class I individuals (data not shown) although the allele frequency was slightly increased in class I homozygotes (0.220 vs 0.115 homozygous vs non-homozygous class I diabetics). Data from our Canadian multiplex families suggest that the class III allele only occurs in conjunction with TH alleles 117 and 124, however, our sample size is very small (n=13) and therefore is not entirely reliable.

TH				IGF1R microsatellite			IGF1R SSCP			INSR		
allele	n	freq	pub'd	allele	n	freq	allele	n	freq	allele	n	freq
109	38	0.183	84	93	77	0.360	1	117	0.657	132	20	0.092
113	39	0.188	59	96	136	0.635	2	61	0.343	138	146	0.670
117	19	0.091	41	other	1	0.005				140	23	0.106
121	52	0.250	53							142	18	0.082
124	57	0.274	129							other	11	0.050
other	3	0.014	6									
total	208	1.000	372		214	1.000		178	1.000		218	1.000

Table 2. Observed frequencies of PCR markers in the random diabetic population and published control frequencies for the TH marker [Puers et al, 1993]

3.2 Multiplex Diabetic Families

The multiplex families were typed as per the random diabetics. Overall allele frequencies (transmitted + non-transmitted) were determined from the parents for all markers and the maternal and paternal frequencies compared to determine if any marked difference exists between mothers and fathers. No significant differences were observed for any marker (ie. TH, IGF1R microsatellite, IGF1R SSCP, IGF1R haplotype, or INSR. Data not shown). The transmitted and non-transmitted allele frequencies were then determined by the AFBAC program and are summarized in Table 3. Results for individual markers will be discussed after considering whether there is any heterogeneity based on geographic origin of families (British versus North American).

TH					IGF1R microsatellite					INSR				
size (bp)	diabetic n	diabetic freq	control n	control freq	size (bp)	diabetic n	diabetic freq	control n	control freq	size (bp)	diabetic n	diabetic freq	control n	control freq
109	131	0.276	58.5	0.231	93	166	0.373	90.5	0.387	132	39	0.089	23.5	0.097
113	96	0.202	48.75	0.192	96	278.5	0.626	139.5	0.596	138	280.5	0.638	160	0.660
117	45.5	0.096	26.25	0.104	other	0.5	0.001	4	0.017	140	55.5	0.126	26	0.107
121	89.5	0.188	34.5	0.136						142	31	0.070	19.75	0.081
124	106	0.223	81.5	0.321						other	34	0.077	13.25	0.055
other	7	0.015	4	0.016										
total	475	1.000	253.5	1.000		445	1.000	234	1.000		440	1.000	242.5	1.000

Table 3. Diabetic and control frequencies in multiplex families of microsatellite markers. (AFBAC multiplex analysis)

IGF1R SSCP					IGF1R Haplotype				
Allele	diabetic		control		Allele	diabetic		control	
	n	freq	n	freq		n	freq	n	freq
1	258	0.686	122	0.637	3	105.5	0.260	57.5	0.282
2	118	0.314	69.5	0.363	4	44.5	0.110	17.5	0.086
					5	170	0.419	83	0.407
					6	85.5	0.211	42	0.206
					other	0.5	0.001	4	0.019
Total	376	1.000	191.5	1.000		406	1.000	204	1.000

Table 3. (cont.) Diabetic and control frequencies in multiplex families of microsatellite markers. (AFBAC multiplex analysis)

3.2.1 Analysis of Markers Based on Geographic Location

Analysis of the data by geographical origin was also done to determine if the two data groups were significantly different from one another. The data was subdivided into British (BDA) versus North American (NA, ie. PH and CDN) families. The Canadian families were not separated into their own subgroup due to the small number of families available (n=20). Overall data for all markers by geographical location is summarized in Table 4. An association is seen between the TH marker and IDDM for the NA families in the simplex and multiplex analyses as well as in the paternally inherited group for the NA families for the simplex and double simplex analysis. Comparable analyses in the British diabetic group showed no significant association. The IGF1R also shows significant results with the overall NA family data in the double simplex analysis but not in the British for the same analyses. No other markers divided by geographic location show any significant differences between transmitted and non-transmitted allele frequencies.

Comparisons between the two overall geographic locations show no significant difference between the BDA and the NA families for the IGF1R microsatellite and the INSR markers for BDA transmitted vs NA transmitted allele frequencies or for BDA non-transmitted vs NA non-transmitted control allele frequencies. Significant differences exist between the two sets for the TH marker in the non-transmitted control allele frequencies for the multiplex and double simplex analyses (χ^2=13.01, df=5, p=0.023; χ^2=13.50, df=5, p=0.019, data not shown). The differences between the non-transmitted frequencies in the two groups appears to be

due to alleles 109 and 117. Allele 109 has a higher frequency in the BDA population compared to the NA families (0.309 vs 0.173 multiplex analysis; 0.287 vs 0.209 double simplex), while allele 117 is decreased in the BDA group (0.049 vs 0.144 multiplex; 0.082 vs 0.124 double simplex). Similarly, significant differences exist between the BDA and the NA IGF1R SSCP marker for the transmitted allele frequencies (simplex χ^2=3.85, df=1, p=0.050; multiplex χ^2=6.51, df=1, p=0.011; double simplex χ^2=8.96, df=1, p=0.003). Also, significant differences exists between the non-transmitted frequencies for the SSCP double simplex analysis (χ^2=4.88, df=1, p=0.027). Significant differences also exists between the BDA and the NA families transmitted frequencies for the haplotyped markers using the AFBAC double simplex analysis (χ^2=12.78, df=4, p=0.012).

Marker	Location	Analysis	χ^2	p	paternal χ^2 transmitted vs controls	p	maternal χ^2 transmitted vs controls	p
TH	Overall	simplex	12.45	***0.029**	8.31	0.140	7.42	0.191
	(n=247)	multiplex	10.35	0.066	5.63	0.344	7.04	0.218
		double	14.51	***0.013**	13.26	***0.021**	5.98	0.308
	BDA	simplex	3.39	0.640	0.41	0.995	8.08	0.152
	(n=99)	multiplex	5.26	0.386	0.69	0.984	7.88	0.163
		double	5.22	0.389	0.98	0.964	6.79	0.237
	NA	simplex	13.91	***0.016**	11.05	***0.050**	5.38	0.371
	(n=148)	multiplex	15.03	***0.010**	9.71	0.084	6.38	0.271
		double	19.60	****0.002**	17.33	****0.004**	4.52	0.478
INSR	Overall	simplex	2.36	0.798	1.54	0.909	1.18	0.947
	(n=247)	multiplex	2.29	0.808	4.01	0.548	1.63	0.898
		double	1.04	0.959	3.31	0.652	2.90	0.716

Table 4. Analysis of markers by geographic location and by parental type *p<0.05 **p<0.01 n=# of families

Marker	Location	Analysis	χ^2	p	paternal χ^2 transmitted vs controls	p	maternal χ^2 transmitted vs controls	p
INSR	BDA	simplex	4.16	0.527	1.01	0.962	6.17	0.29
	(n=99)	multiplex	2.04	0.844	1.94	0.857	1.61	0.9
		double	3.79	0.58	2.22	0.818	2.88	0.718
	NA	simplex	7.26	0.202	4.4	0.494	8.08	0.152
	(n=147)	multiplex	3.9	0.565	5.3	0.38	6.63	0.25
		double	5.51	0.357	5.67	0.34	9.86	0.079
IGF1R	Overall	simplex	6.29	***0.043**	4.81	0.09	1.5	0.472
	(n=246)	multiplex	6.2	***0.045**	4.12	0.127	2.24	0.326
		double	7.59	***0.022**	4.91	0.086	2.83	0.244
	BDA	simplex	2.04	0.36	2.02	0.364	0.09	0.958
	(n=99)	multiplex	2.03	0.362	1.1	0.576	1.05	0.592
		double	3.8	0.149	2.25	0.324	1.58	0.454

Table 4. (cont.) Analysis of markers by geographic location and by parental type *p<0.05 **p<0.01 n=# of families

Marker	Location	Analysis	χ^2	p	paternal χ^2 transmitted vs controls	p	maternal χ^2 transmitted vs controls	p
IGF1R	NA	simplex	5.62	0.06	3.29	0.193	2.33	0.312
	(n=147)	multiplex	5.76	0.056	3.85	0.146	1.92	0.383
		double	6.03	***0.049**	4.05	0.132	2.03	0.363
SSCP	Overall	simplex	2.53	0.112	1.71	0.191	0.88	0.349
	(n=217)	multiplex	1.38	0.24	0.88	0.349	0.53	0.468
		double	0.67	0.411	0.51	0.475	0.2	0.658
	BDA	simplex	0.90	0.343	0.02	0.883	1.47	0.226
	(n=96)	multiplex	0.01	0.907	0.00	0.979	0.05	0.825
		double	0.02	0.881	0.00	1.000	0.05	0.830
	NA	simplex	1.70	0.193	2.80	0.094	0.02	0.885
	(n=121)	multiplex	2.41	0.120	1.96	0.162	0.63	0.429
		double	0.99	0.319	0.98	0.322	0.17	0.682

Table 4. (cont.) Analysis of markers by geographic location and by parental type *p<0.05 **p<0.01 n=# of families

Marker	Location	Analysis	χ^2	p	paternal χ^2 transmitted vs controls	p	maternal χ^2 transmitted vs controls	p
Haplo	Overall	simplex	6.54	0.163	6.45	0.168	4.51	0.342
	(n=225)	multiplex	7.28	0.122	6.68	0.154	2.25	0.691
		double	8.18	0.085	7.6	0.107	4.22	0.377
	BDA	simplex	3.23	0.52	3.46	0.484	4.67	0.323
	(n=87)	multiplex	3.81	0.432	6.36	0.174	0.46	0.977
		double	5.5	0.24	6.72	0.152	2.34	0.673
	NA	simplex	6.39	0.172	5.7	0.223	2.3	0.68
	(n=138)	multiplex	6.29	0.178	4.86	0.302	2.01	0.734
		double	8.85	0.065	6.19	0.185	3.81	0.432

Table 4. (cont.) Analysis of markers by geographic location and by parental type *p<0.05 **p<0.01 n=# of families

3.2.2 TH Data

Overall association analysis for the TH marker shows an association between TH and IDDM (χ^2=12.45, df=5, p=0.029 simplex analysis, see Table 4). Further analysis of the TH marker shows no association in the multiplex analysis (χ^2=10.35, df=5, p=0.066) whereas an association is again seen in the double simplex analysis (χ^2=14.50 df=5, p=0.013). The TH association appears to be due primarily to the decreased frequency of allele 124 in the diabetic compared to control groups (χ^2=8.84, df=1, p=0.003; simplex data used, data not shown; multiplex data shown in Table 3).

Results for association analysis between TH and IDDM by subgrouped data are shown in Table 5. Further subdivision of the data indicates an association between TH and IDDM in families with increased HLA sharing in affected siblings in simplex and double simplex analyses (χ^2=16.05, df=5, p=0.007; χ^2=12.99, df=5, p=0.024) as well as in families with $\leq$ 50% HLA sharing in affected siblings in double simplex analyses (χ^2=11.12, df=5, p=0.049). An association is also seen for the double simplex analysis of the D15S107 decreased sharing group (χ^2=18.01, df=5, p=0.003). Likewise an association is seen for all three analyses for families where affected siblings share $\leq$ 50% of INSR alleles (χ^2=13.98, df=5, p=0.016, simplex; χ^2=11.94, df=5, p=0.036, multiplex; χ^2=19.88, df=5, p=0.001, double simplex), while no association is seen when subdividing families by age of onset, or by DR4 positivity. Further analysis of the subgroups to one another (eg. D15S107 decreased compared to increased sharing in the transmitted allele frequency groups)

shows that no differences exist between transmitted allele frequencies or non-transmitted frequencies for all subgroups except for the double simplex analysis of non-transmitted allele frequencies in the HLA increased vs decreased sharing comparison group (χ^2=12.04, df=5, p=0.034), double simplex analysis of transmitted allele frequencies for the INSR sharing comparison subgroups (χ^2=11.92, df=5, p=0.036), as well as the double simplex analysis non-transmitted frequencies for the age of onset comparison subgroups (χ^2=13.84, df=5, p=0.017).

Group	Analysis Type	χ^2	p	Group	Analysis Type	χ^2	p
D15S107 sharing >50%	simplex	10.54	0.391	D15S107 sharing ≤50%	simplex	10.56	0.061
(n=97)	multiplex	4.83	0.437	(n=143)	multiplex	10.80	0.055
	double	7.01	0.220		double	18.01	**0.003
HLA sharing >50%	simplex	16.05	**0.007	HLA sharing ≤50%	simplex	5.89	0.317
(n=133)	multiplex	7.75	0.170	(n=85)	multiplex	10.04	0.074
	double	12.99	*0.024		double	11.12	*0.049
INSR sharing >50%	simplex	5.77	0.330	INSR sharing ≤50%	simplex	13.98	*0.016
(n=78)	multiplex	3.34	0.647	(n=153)	multiplex	11.94	*0.036
	double	5.51	0.357		double	19.88	**0.001
age of onset >11 yrs	simplex	5.93	0.313	age of onset ≤11 yrs	simplex	4.18	0.524
(n=128)	multiplex	4.77	0.444	(n=94)	multiplex	3.02	0.700
	double	6.16	0.291		double	8.38	0.136

Table 5. AFBAC analysis results of the TH marker split into various data groups. df=5 * $p<0.05$ ** $p<0.01$ n=# of families

Group	Analysis Type	χ^2	p	Group	Analysis Type	χ^2	p
DR4 positive	simplex	5.97	0.309	DR4 negative	simplex	7.71	0.173
(n=168)	multiplex	5.57	0.35	(n=64)	multiplex	8.3	0.14
	double	7.67	0.175		double	10.26	0.068

Table 5. (cont) AFBAC analysis results of the TH marker split into various data groups. df=5 * $p<0.05$ ** $p<0.01$ n=# of families

The TH data was then analyzed by parental origin of the alleles (results summarized in Table 6). A paternal association is seen for the TH marker with IDDM in the overall data set as well as for the ≤ 50% affected sibling sharing group of D15S107, decreased INSR sharing group, as well as in non-DR4 families, while a maternal association is seen in families with decreased HLA sharing in affected siblings. Comparisons between paternal and maternal data (ie. paternal versus maternal transmitted allele frequencies within a data group), show no difference between the two groups save for the non-DR4 families double simplex transmitted data by sex analysis (χ^2=19.52, df=5, p=0.002). Analysis between sharing groups by parental origin (ie. paternal transmitted allele frequencies for D15S107 increased and decreased affected sibling sharing of alleles) shows that significant differences exist between paternally non-transmitted alleles for D15S107 increased compared to decreased sharing subgroups (χ^2=11.83, df=5, p=0.037, double simplex), paternally transmitted and non-transmitted alleles for INSR sharing comparison groups (χ^2=14.97, df=5, p=0.010; χ^2=14.09, df=5, p=0.015, double simplex), transmitted allele frequencies for age of onset subgroups (χ^2=13.43, df=5, p=0.020, double simplex), and transmitted allele frequencies for DR4 positive vs negative families (χ^2=20.00, df=5, p=0.001, double simplex). Maternally, significant differences exists only between the transmitted allele frequencies for the age of onset comparison groups (χ^2=17.14, df=5, p=0.004, double simplex).

Paternal Data Group	Analysis Type	χ^2	p	Maternal Data Group	Analysis Type	χ^2	p
overall data	simplex	8.31	0.140	overall data	simplex	7.42	0.191
(n=247)	multiplex	5.63	0.344	(n=247)	multiplex	7.04	0.218
	double	13.26	***0.021**		double	5.98	0.308
D15S107 sharing >50%	simplex	3.75	0.586	D15S107 sharing >50%	simplex	9.92	0.077
(n=97)	multiplex	1.96	0.855	(n=97)	multiplex	7.11	0.212
	double	7.12	0.212		double	3.35	0.646
D15S107 sharing ≤50%	simplex	10.40	0.065	D15S107 sharing ≤50%	simplex	4.78	0.443
(n=143)	multiplex	8.58	0.127	(n=143)	multiplex	4.74	0.448
	double	19.30	****0.002**		double	5.11	0.403
HLA sharing >50%	simplex	9.69	0.084	HLA sharing >50%	simplex	8.31	0.140
(n=133)	multiplex	5.47	0.362	(n=133)	multiplex	4.47	0.484
	double	10.98	0.052		double	3.68	0.597

Table 6. Paternal and Maternal data analysis for TH marker. df=5 * $p<0.05$ ** $p<0.01$ n=# of families

Paternal Data Group	Analysis Type	χ^2	p	Maternal Data Group	Analysis Type	χ^2	p
HLA sharing ≤50%	simplex	3.87	0.568	HLA sharing ≤50%	simplex	4.83	0.437
(n=85)	multiplex	4.18	0.524	(n=85)	multiplex	9.19	0.102
	double	3.01	0.699		double	12.07	***0.034**
INSR sharing >50%	simplex	6.47	0.264	INSR sharing >50%	simplex	4.14	0.529
(n=78)	multiplex	4.16	0.527	(n=78)	multiplex	3.74	0.587
	double	10.05	0.074		double	2.28	0.810
INSR sharing ≤50%	simplex	11.28	***0.046**	INSR sharing ≤50%	simplex	6.08	0.298
(n=153)	multiplex	7.42	0.191	(n=153)	multiplex	5.35	0.374
	double	15.53	****0.008**		double	7.08	0.215
age of onset >11 yrs	simplex	4.30	0.507	age of onset >11 yrs	simplex	10.00	0.075
(n=128)	multiplex	4.05	0.542	(n=128)	multiplex	2.82	0.728
	double	9.56	0.089		double	2.54	0.771

Table 6. (cont.) Paternal and Maternal data analysis for TH marker. df=5 * $p<0.05$ ** $p<0.01$ n=# of families

Paternal Data Group	Analysis Type	χ^2	p	Maternal Data Group	Analysis Type	χ^2	p
age of onset ≤11 yrs	simplex	3.18	0.672	age of onset ≤11 yrs	simplex	1.73	0.885
(n=94)	multiplex	2.64	0.755	(n=94)	multiplex	2.48	0.780
	double	9.90	0.078		double	5.78	0.328
DR4 positive	simplex	4.52	0.477	DR4 positive	simplex	5.08	0.406
(n=168)	multiplex	3.78	0.581	(n=168)	multiplex	5.00	0.416
	double	6.87	0.230		double	6.00	0.307
DR4 negative	simplex	6.88	0.230	DR4 negative	simplex	8.33	0.139
(n=64)	multiplex	9.70	0.084	(n=64)	multiplex	3.17	0.674
	double	15.30	**0.009		double	2.18	0.824

Table 6. (cont.) Paternal and Maternal data analysis for TH marker. df=5 * $p<0.05$ ** $p<0.01$ n=# of families

3.2.3 IGF1R Data

AFBAC analysis of the IGF1R marker shows that overall, an association exists between IGF1R and IDDM (simplex χ^2=6.29, df=2, p=0.043; multiplex χ^2=6.20, df=2, p=0.045; double simplex χ^2=7.59, df=2, p=0.022, see Table 4). Overall marker information was also examined for the IGF1R SSCP and haplotypes using the AFBAC method. No significant results were observed between the diabetic and control frequencies for these markers. Marker information from the microsatellite was used exclusively for the IGF1R marker analysis as the microsatellite marker, while not as informative as the haplotypes, had a larger sample size, and is more informative than the SSCP. Subgrouped data for the IGF1R marker is shown in Table 7. An association is observed between IGF1R and IDDM in families where affected siblings share more than 50% of TH alleles (multiplex analysis χ^2=7.74, df=2, p=0.021, double simplex χ^2=8.89, df=2, p=0.012) as well as with DR4 positive families (multiplex χ^2=6.44, df=2, p=0.040, double simplex χ^2=8.30, df=2, p=0.016). Also an association is seen in the double simplex analysis between IGF1R and IDDM in families where affected siblings share more than 50% of HLA haplotypes (χ^2=7.20, df=2, p=0.027). AFBAC analysis of the D15S107 marker interestingly showed no evidence of an association with IDDM in our data set (data not shown).

Differences in the transmitted and non-transmitted allele frequencies between the data subgroups was then analyzed. Differences were found between the non-transmitted frequencies of the IGF1R microsatellite when the data set was divided

by TH sharing (ie. TH sharing greater than 50% versus sharing ≤ 50%) for multiplex and double simplex analyses (χ^2=7.61, df=2, p=0.022, multiplex; χ^2=13.29, df=2, p=0.001, double simplex) as well as the non-transmitted group for the age of onset subgroups in the double simplex analysis (χ^2=6.46, df=2, p=0.040).

Analysis of the IGF1R marker data by parental origin shows that there is a paternal association seen in the double simplex analysis of the increased INSR sharing group (χ^2=6.39, df=2, p=0.041) and a paternal association seen in the increased TH sharing group double simplex analysis(χ^2=6.72, df=2, p=0.035) (data not shown). No other significant associations are seen in the data subgroups by parental origin for the IGF1R microsatellite marker.

Comparison of the transmitted and non-transmitted frequencies between data subgroups (ie. TH greater than versus ≤ 50% sharing) show that significant differences exists between the TH subgroups non-transmitted allele frequencies for the multiplex and double-simplex analyses (χ^2=7.61, df=2, p=0.022 and χ^2=13.29, df=2, p=0.001 respectively), as well as in the non-transmitted allele frequencies for the age of onset subgroups in the double simplex analysis (χ2=6.46, df=2, p=0.040).

Group	Analysis Type	χ^2	p	Group	Analysis Type	χ^2	p
TH sharing >50%	simplex	4.60	0.100	TH sharing ≤50%	simplex	1.05	0.591
(n=82)	multiplex	7.74	***0.021**	(n=141)	multiplex	1.66	0.435
	double	8.89	***0.012**		double	3.84	0.147
HLA sharing >50%	simplex	3.16	0.206	HLA sharing ≤50%	simplex	5.31	0.070
(n=132)	multiplex	4.76	0.093	(n=86)	multiplex	3.64	0.162
	double	7.20	***0.027**		double	3.45	0.178
INSR sharing >50%	simplex	2.02	0.364	INSR sharing ≤50%	simplex	4.00	0.136
(n=77)	multiplex	5.31	0.070	(n=152)	multiplex	2.37	0.306
	double	5.69	0.058		double	2.88	0.237
age of onset >11 yrs	simplex	1.00	0.605	age of onset ≤11 yrs	simplex	5.86	0.054
(n=130)	multiplex	2.23	0.327	(n=95)	multiplex	4.26	0.119
	double	2.66	0.264		double	4.65	0.098

Table 7. AFBAC analysis of IGF1R microsatellite data split into different analysis groups. df=2 *p<0.05 n=# of families

Group	Analysis Type	χ^2	p	Group	Analysis Type	χ^2	p
DR4 positive	simplex	4.1	0.129	DR4 negative	simplex	3.17	0.075
(n=168)	multiplex	6.44	***0.040**	(n=65)	multiplex	3.44	0.064
	double	8.3	***0.016**		double	2.07	0.15

Table 7. (cont) AFBAC analysis of IGF1R microsatellite data split into different analysis groups. df=2 *p<0.05 n=# of families

3.2.4 INSR Marker

Analysis of the overall data for the INSR marker shows that no significant differences exist between the diabetic and control frequencies for this marker (data shown in Tables 4 and 8). Further analysis of this marker by parental origin of the alleles shows that no significant differences exists between the diabetic and the control frequencies for either parent (data not shown). However, analysis of the transmitted and non-transmitted allele frequencies by sex shows significant differences exist between paternal compared to maternal allele frequency transmission (data shown in Table 9).

From the table we see that significant differences exists between paternally and maternally inherited alleles of the INSR for the D15S107 decreased sharing data group, as well as the TH decreased sharing group and HLA increased sharing group.

Group	Analysis Type	χ^2	p	Group	Analysis Type	χ^2	p
TH sharing >50%	simplex	5.05	0.410	TH sharing ≤50%	simplex	7.36	0.195
(n=83)	multiplex	6.43	0.267	(n=145)	multiplex	1.28	0.936
	double	5.78	0.329		double	2.21	0.820
HLA sharing >50%	simplex	6.72	0.243	HLA sharing ≤50%	simplex	4.83	0.437
(n=135)	multiplex	0.58	0.989	(n=87)	multiplex	3.55	0.615
	double	1.76	0.882		double	3.56	0.614
D15S107 sharing >50%	simplex	1.52	0.910	D15S107 sharing ≤50%	simplex	3.34	0.648
(n=95)	multiplex	2.26	0.811	(n=142)	multiplex	1.17	0.948
	double	1.74	0.884		double	0.84	0.974
age of onset >11 yrs	simplex	3.00	0.699	age of onset ≤11 yrs	simplex	2.91	0.714
(n=124)	multiplex	1.72	0.887	(n=97)	multiplex	4.40	0.493
	double	2.23	0.816		double	4.78	0.443
DR4 positive	simplex	6.42	0.268	DR4 negative	simplex	6.38	0.271
(n=166)	multiplex	3.79	0.580	(n=66)	multiplex	4.82	0.438
	double	6.56	0.255		double	9.30	0.098

Table 8. AFBAC analysis of INSR data df=5. n=# of families

Data Group	Analysis	Paternal vs Maternal transmitted χ^2	p	Paternal vs Maternal non-transmitted χ^2	p
overall data	simplex	6.05	0.301	4.91	0.428
(n=247)	multiplex	7.56	0.182	4.20	0.521
	double	16.71	**0.005	10.28	0.069
D15S107 sharing >50%	simplex	4.44	0.488	5.21	0.391
(n=95)	multiplex	5.51	0.357	4.82	0.439
	double	13.23	*0.021	7.16	0.209
D15S107 sharing ≤50%	simplex	15.10	*0.010	3.14	0.690
(n=142)	multiplex	13.19	*0.022	1.86	0.869
	double	29.10	**0.000	7.33	0.197
TH sharing >50%	simplex	4.42	0.490	5.31	0.379
(n=83)	multiplex	4.22	0.518	6.42	0.267
	double	9.37	0.095	11.24	*0.047

Table 9. Transmitted and non-transmitted by sex analysis of the INSR marker, df=5. *p<0.05 **p<0.01

Data Group	Analysis	Paternal vs Maternal transmitted χ^2	p	Paternal vs Maternal non-transmitted χ^2	p
TH sharing ≤50%	simplex	11.92	***0.036**	6.62	0.251
(n=145)	multiplex	10.22	0.069	5.79	0.327
	double	19.83	****0.001**	16.65	****0.005**
HLA sharing >50%	simplex	12.67	***0.027**	8.09	0.151
(n=135)	multiplex	8.61	0.126	2.65	0.754
	double	19.63	****0.001**	14.97	***0.010**
HLA sharing ≤50%	simplex	2.25	0.814	2.14	0.829
(n=87)	multiplex	2.94	0.709	6.61	0.252
	double	7.14	0.210	5.88	0.318
DR4 positive	simplex	6.17	0.290	4.55	0.473
(n=166)	multiplex	5.51	0.357	5.90	0.316
	double	13.80	***0.017**	15.72	****0.008**

Table 9. (cont). Transmitted and non-transmitted by sex analysis of the INSR marker, df=5. *p<0.05 **p<0.01

Data Group	Analysis	Paternal vs Maternal transmitted χ^2	p	Paternal vs Maternal non-transmitted χ^2	p
DR4 negative	simplex	1.79	0.877	4.09	0.536
(n=66)	multiplex	1.38	0.927	2.45	0.784
	double	2.75	0.738	6.17	0.29
age of onset >11yrs	simplex	2.95	0.708	9.31	0.097
(n=124)	multiplex	3.56	0.615	5.1	0.404
	double	7.26	0.202	16.13	**0.006
age of onset ≤11yrs	simplex	7.02	0.219	5.53	0.355
(n=97)	multiplex	7.87	0.164	4.14	0.529
	double	16.4	**0.006	9.52	0.09

Table 9. (cont). Transmitted and non-transmitted by sex analysis of the INSR marker, df=5. *p<0.05 **p<0.01

3.2.5 Paternally inherited DR4

A previously published article suggested that an association existed between paternally inherited DR4 and the TH/INS gene with IDDM. In our data group, no evidence exists for any paternal DR4 association between IDDM and TH (data shown in Table 10). There is some evidence of weak association between non-paternally inherited DR4 and IGF1R, however, this is not seen in the haplotype or the SSCP.

Paternal DR4 Subgroup	Analysis Type	χ^2	p	Non-pat. DR4 Subgroup	Analysis Type	χ^2	p
TH	simplex	8.49	0.131	TH	simplex	4.61	0.466
(n=97)	multiplex	5.73	0.333	(n=71)	multiplex	5.44	0.364
	double	8.57	0.128		double	6.43	0.267
IGF1R	simplex	1.03	0.597	IGF1R	simplex	3.09	0.213
(n=97)	multiplex	2.20	0.332	(n=71)	multiplex	6.19	***0.045**
	double	3.72	0.155		double	6.58	***0.037**
INSR	simplex	9.10	0.105	INSR	simplex	4.44	0.488
(n=96)	multiplex	8.22	0.144	(n=70)	multiplex	1.24	0.941
	double	14.36	***0.013**		double	3.39	0.640
IGF1R SSCP	simplex	2.89	0.089	IGF1R SSCP	simplex	4.60	***0.032**
(n=89)	multiplex	1.51	0.219	(n=57)	multiplex	2.18	0.139
	double	2.11	0.147		double	1.49	0.222

Table 10. AFBAC analysis of paternally inherited DR4 families (TH, INSR df=5; haplo df=4, IGF1R df=2, SSCP df=1) *$p<0.05$

Paternal DR4 Subgroup	Analysis Type	χ^2	p	Non-pat. DR4 Subgroup	Analysis Type	χ^2	p
IGF1R Haplo	simplex	3.58	0.466	IGF1R Haplo	simplex	4.52	0.34
(n=94)	multiplex	2.14	0.709	(n=63)	multiplex	6.23	0.182
	double	6.38	0.173		double	6.19	0.185

Table 10. (cont) AFBAC analysis of paternally inherited DR4 families (TH, INSR df=5; haplo df=4, IGF1R df=2, SSCP df=1) *p<0.05

3.3 Linkage Analysis

3.3.1 LOD Scores

Linkage analysis of the markers was done by the LOD score method. The model used included a disease allele frequency of 0.25, and penetrance of 0, 0.4, 0.6 for genotypes with 0, 1 or 2 disease alleles present. Strong evidence against linkage is seen for TH at θ=0.00 (LOD=-4.19) and the maximum LOD score is 0.39 at θ=0.3 (ie. at a distance of 30 cM). Splitting of the data into increased HLA and decreased HLA sharing shows no evidence for linkage at θ=0.00 in either subgroup (-2.71 with increased HLA sharing and -1.50 with decreased HLA sharing).

LOD score analysis of the INSR shows little evidence for linkage at θ=0.00 (LOD score = -1.46). (Max score obtained was LOD=0.57 at θ=0.20). Analysis of the IGF1R microsatellite shows greatest evidence for linkage at θ=0.20 with a score of LOD=0.61. In families with increased HLA sharing, a maximum LOD score of 0.59 is obtained at θ=0.10.

LOD score analysis of the markers by geographic location shows maximum scores of LOD=0.25 at θ=0.30 in the NA families (n=151) for the TH microsatellite, while the BDA families (n=100) show a maximum LOD score of LOD=0.14 at θ=0.30. Maximum LOD scores in the NA families (n=111) for the INSR marker is LOD=0.34 at θ=0.20, while in the BDA families (n=80) the maximum LOD score obtained was LOD=0.24 at θ=0.20. The maximum LOD scores obtained for the IGF1R marker in the NA families (n=104) was at θ=0.05, LOD=1.48, while in the BDA families (n=83) the maximum LOD score was obtained at θ=0.40, LOD=-0.03.

3.3.2 Affected Sibling Pair Sharing

Linkage was also determined using average sibling sharing of alleles at a given marker. Expected sharing of genes between siblings purely by chance (ie. Mendelian segregation) is 0.50. Average sharing for the TH locus was 0.53 (1-sided χ^2=1.25, df=1, p=0.13), D15S107 had an average sharing of 0.56 (χ^2=6.11, df=1, p=0.005) and 0.53 for INSR (χ^2=0.73, df=1, p=0.195). D15S107 was used instead of the IGF1R to score affected sibling sharing as D15S107 is much more informative than the IGF1R microsatellite alone or combined in haplotypes.

3.4 Interactions Between Loci

Marker-marker interaction was determined by doing a 3x3 contingency χ^2 table of the sharing observed in the two markers. Sharing groups were arbitrarily defined as 0, $0<x\leq1$, and $1<x\leq2$. Interaction between markers was determined for HLA and TH (see Table 4), HLA and D15S107, D15S107 and TH, INSR and HLA, INSR and TH, and INSR and D15S107. No significant results were observed between any of the groups examined (D15S107-TH χ^2=3.75, df=4, p=0.44; HLA-TH χ^2=3.71, df=4, p=0.45; HLA-D15S107 χ^2=1.85, df=4, p=0.76; INSR-TH χ^2=5.89, df=4, p=0.207; INSR-D15S107 χ^2=8.95, df=4, p=0.062; INSR-HLA χ^2=9.06, df=4, p=0.060).

TH D15S107	2	1	0	SUM
2	34	45	6	85
(Expected)	(35.63)	(43.43)	(5.94)	
1	49	61	6	116
(Expected)	(48.63)	(59.27)	(8.10)	
0	13	11	4	28
(Expected)	(11.74)	(14.31)	(1.96)	
SUM	96	117	16	229

Table 11. 3x3 contingency χ^2 table of interaction between TH and D15S107 allele sharing in affected sibling pairs. χ^2=3.75, df=4, p=0.44

4. Discussion

4.1 TH and the INS gene region

Microsatellite PCR is a rapid and precise method of typing markers for both association and linkage analysis. The INS gene region has been extensively studied for almost 10 years in the hopes of further defining the observed association between IDDM and the INS 5' VNTR. However, attempts at demonstrating linkage in this region have been inconsistent. Owerbach and Gabbay 1994 showed by affected sib pairs that the sharing ratios for 0, 1, or 2 alleles were in fact significantly different from expected values. However, this result has not held in every data set. In this study, using the more informative tetranucleotide repeat in the TH gene which is tightly linked to the INS VNTR, it was hoped that actual linkage could be demonstrated. However, this was not the case. LOD score analysis of the TH marker shows significant evidence against close linkage with IDDM (LOD = -4.19 at θ=0.00) and no significant evidence for linkage (average affected sibling sharing of alleles was 0.53, p=0.13). While the average sharing of TH alleles between affected siblings is higher than the expected 0.5 from Mendelian segregation, this difference was not significant. Despite the lack of significant evidence for linkage at this locus, there is evidence of an association between IDDM and TH (p=0.029, simplex analysis; p=0.013 double simplex analysis, see Table 4). In our data set, the TH association is due primarily to two alleles; allele 124 which is decreased in the diabetic population, and allele 121 which is increased in diabetics compared to controls (see Table 3).

Analysis of the TH allele frequencies in our random diabetic population showed significant differences between the class I homozygous IDDM diabetics compared to non-class I homozygous diabetics. This indicates that linkage disequilibrium can be detected between the INS VNTR and TH, and that the TH microsatellite is therefore a valid substitute for study of this region. In our random diabetic population, an association was seen between TH and IDDM only when using published frequencies and not with the AFBAC derived control frequencies (published controls: χ^2=15.02, df=5, p=0.010; AFBAC controls: χ^2=10.23, df=5, p=0.069 multiplex analysis). Strongest evidence for a TH association was observed in the NA families, while no association was observed from the BDA families alone. This may suggest that there is a geographic difference in frequencies of the TH/INS-associated diabetes susceptibility gene between the two groups, as the non-transmitted allele frequencies between the NA and BDA families were significantly different. Since the TH association within the multiplex diabetic families was seen in the NA families and not seen with the BDA families, it is possible that this reflects differences in susceptibility to IDDM due to ethnic background. Thus, AFBAC control frequencies derived from the NA multiplex families would be more closely matched with the random Canadian diabetics in our study than would AFBAC control frequencies derived from BDA families. When the random diabetics were re-analyzed using the NA TH control frequencies, an association between IDDM and TH was observed with a χ^2=12.04, df=5, p=0.034 (data not shown). Comparisons between the NA multiplex families transmitted allele frequencies with the random

diabetic allele frequencies shows no significant differences between the two groups (χ^2=5.34, df=5, p=0.378). These results emphasize the extreme importance of using ethnically matched controls for association studies and the usefulness of a method which derives ethnically matched controls from within the families (ie. the AFBAC method). Also, the results show that a microsatellite marker can indeed be used to detect associations due to linkage disequilibrium and that they do not mutate so rapidly as to destroy disequilibrium effects.

Julier et al 1991 reported that the 5' VNTR association with IDDM occurred primarily in individuals with paternally derived DR4. However, this was not supported in our data set (p=0.131) and likewise has yet to be replicated in an independent study [Cornall 1993]. This may reflect differences between the two populations being studied (ie. BDA and NA vs French). Instead, an association between TH and IDDM was observed in the subgroup of families whose affected siblings shared $\leq$ 50% of alleles at INSR (see Table 5). The associations observed between IDDM and TH in the $\leq$ 50% sharing of D15S107 subgroup and the $\leq$ 50% sharing of HLA subgroup may have occurred by chance as they are only observed in one analysis group. Also, while the TH association in the increased HLA sharing group is observed in two analysis groups (see Table 5), the interactive analysis between TH and HLA joint sharing shows no evidence for any interaction between the two groups (data not shown). Thus, it is unlikely that the observed association is indicative of a true interaction between TH and HLA for IDDM.

When doing multiple analyses on some data, the p-values should be corrected

for the number of different analyses run such that for the TH data set divided by INSR sharing, the p-values would be corrected by multiplying by 2 (ie. the number of separate analyses). Thus the association between TH and IDDM in the ≤50% HLA sharing group p-value would be corrected to 0.098 (see Table 5). The remaining observed associations even when corrected for data subgrouping maintain p-values less than 0.05.

Parental sex-dependent effects were observed to a degree, especially paternal-dependent associations between IDDM and TH (see Tables 4 and 6). However, while the allele frequencies were significantly different between the transmitted and control groups, the p-values were consistently lower in the overall (ie. maternal and paternal data combined) data compared to the individual parental data. Thus, while most of the susceptibility may be derived through one parent, it cannot be ruled out that susceptibility is in fact due to both. In our dataset, while a paternal association is observed especially in the NA families, a similar trend (though non-significant) is observed in the BDA mothers. Thus this trend does not follow typical imprinting influences. Also, while the TH gene is close to an imprinted gene there is no evidence that the TH gene is imprinted.

Genetic imprinting is a mechanism of differential expression of a gene depending on the sex of the parent from whom the allele is inherited. Thus a disease allele which is maternally imprinted may not be expressed when inherited maternally and the child would develop normally. However, when the disease allele is then transmitted paternally, the child would then be affected. There is evidence showing

that the IGF2 gene is imprinted. Expression studies have shown that one of the IGF2 promoters is imprinted in the fetal liver [Vu and Hoffman 1994]. Also, there are reports that the human insulin gene's mouse homologue is imprinted in the yolk sac [Giddings et al 1994]. Should genetic imprinting be involved in IDDM susceptibility in this region, a difference in paternal vs maternal transmission frequencies of "disease" alleles would be observed. While imprinting is not likely to be a factor in this study, it is important to keep in mind the potential effects it may have as the IGF2 gene may be imprinted. Also, there is evidence of skewed transmission rates between mothers and fathers of affected children for the HLA DR4 allele.

Recent studies indicate that the actual susceptibility locus for IDDM in the INS gene region is in fact the 5' VNTR and that the VNTR may act through altering transcription rates. The means of transcription alteration is not entirely understood and reports are conflicting regarding the direction of change. However, if the VNTR is the actual locus of susceptibility, it is important to note that the VNTR may alter the transcription rate of any three of the genes in the area (ie. TH, INS, or IGF2). A possible role for the VNTR may be to alter susceptibility to IDDM depending on transcription levels of either INS or IGF2 possibly through induction or loss of immunologic tolerance to INS. Likewise, it may be that higher/lower (depending on the actual direction of change of transcription rate) levels of either INS or IGF2 may act through a secondary receptor eg. INSR or IGF1R resulting in susceptibility to IDDM. This would require looking at specific INS VNTR alleles and re-examining

the data according to the sharing criteria. Also, it remains to be determined whether altered transcription rates observed results in changes in INS levels in the blood.

The difference seen in TH/INS linkage and association results may be indicative of a difference in sensitivity of the two analysis methods. Since association analysis examines specific allele frequencies, it may be more sensitive for detecting susceptibility loci which may or may not be "necessary" for pathogenesis (ie. when susceptibility alleles only modify risk upwards or downwards). Also, association analysis is a model independent method of analyzing data, such that in cases where the mode of inheritance is poorly understood, meaningful results may still be obtained. Finally, association analysis reflects linkage disequilibrium between two loci over many generations of a population, whereas linkage analysis is performed only on 2-3 generations of specific families.

Linkage analysis is especially useful when large multi-generation families with multiple affected individuals are being analyzed. However, with small multiplex families, large numbers of families are required to obtain significant results. In the case of susceptibility loci, where only a small group of families may have the susceptibility allele, linkage may not detect the effect of the susceptibility allele as the remaining background families may obscure the significance of the effect. When an incorrect model is specified for linkage analysis, the results tend to be less significant or have increased θ (recombination) values [Durner and Greenberg, 1992]. Also, LOD score analyses tend to exaggerate recombination distances when genetic heterogeneity is involved (affected individuals who do not

share similar genes are thought to be recombinants). Thus results from LOD score analyses may indicate that the disease locus may be fairly far away (and therefore uninteresting for association analyses) despite being fairly nearby. The affected sibling pair analysis is a model-independent method of linkage analysis however, linkage by the LOD score method under the correct model is more informative for linkage. While association analysis (by the AFBAC method) also requires large numbers of small nuclear families, it is also possible to do association studies using random affected individuals. However, with random affected individuals, control individuals must then be obtained which are as closely matched for ethnicity as possible to the study individuals.

4.2 IGF1R

Association analysis of the data also shows an association between the IGF1R microsatellite and IDDM in multiplex families ($p=0.043$ simplex analysis; $p=0.045$ multiplex analysis; $p=0.022$ double simplex analysis; see Table 4). The IGF1R gene is close to D15S107 (3.5 cM) which Field et al 1994 reported as being linked to IDDM in an HLA-independent manner. While IGF1R LOD score analysis of the data suggests that the actual locus of interest may be up to 20 cM away it is well known that distance (θ) estimates are usually exaggerated for complex multigenic disorders [Durner and Greenberg, 1992]. Also, an observable association in families between IDDM and the IGF1R locus suggests that the locus of interest is fairly close in order to detect linkage disequilibrium between IDDM and IGF1R. Affected

sibling sharing analysis indicates that there is significantly increased allele sharing in the D15S107 marker (average sharing 0.56, p=0.005). The increased allele sharing may reflect the D15S107 linkage to IDDM due to close linkage between IGF1R and D15S107. Although association analysis of the IGF1R with the random diabetics showed no association between IGF1R and IDDM (p=0.361), no association was observed between the D15S107 marker and IDDM (data not shown). This suggests that the susceptibility locus may in fact be closer to the IGF1R as a weak association is observable with this marker despite lack of significant linkage.

AFBAC analysis of the IGF1R data, split by marker sharing, shows that the association between IGF1R and IDDM may be stronger in families where affected siblings share more TH alleles. Significant differences are observed in the non-transmitted allele frequencies between TH increased and decreased sharing groups for both multiplex and double simplex analyses. Since the control frequencies are significantly different from one another, this suggests that the two groups are genetically distinct from one another. Yet, the transmitted allele frequencies are not significantly different from each other. This may suggest that with increased TH sharing, there is an increased predisposition to IDDM. However, no evidence for an interaction between TH and IGF1R is supported from the joint sharing analysis between the two markers (data not shown). Given the lack of evidence for an interaction between these two loci, it may instead be that increased TH sharing alone influences the IGF1R association and that there is no reciprocity in this interaction since there is no evidence for an association between TH and IDDM in a D15S107

dependent manner (see Table 5). Also, there is evidence of increased linkage in the NA families for the IGF1R marker (LOD=1.48, θ=0.05, disease allele frequency 0.25; penetrances 0, 0.4, 0.6) such that the observed association may reflect this increase in linkage depending on the average sharing of alleles between the two geographic locations. However, the average allele sharing by geographic location was not determined.

The observed parental differences in the IGF1R associations (data not shown), much like those observed in TH, are likely due to chance. The overall probabilities are lower than the individual parental probabilities, but this would be expected as the sample size being examined becomes smaller and therefore more difficult to show significant association between the marker and IDDM. The results may suggest that although there is some influence from both parents, the greatest influence may be seen from one parent. However, if one corrects the p-values for the number of data subsets created, then the p-values lose their significance.

While most IDDM diabetics (61/87 - 70% of our random diabetics) are homozygous for the INS class I VNTR, no difference was observed in the IGF1R marker for homozygous class I diabetics compared to non-homozygous class I diabetics in our random diabetics samples. However, since class I non-homozygous individuals are relatively rare (30% of our random diabetic population), the sample size for this analysis was fairly small and therefore the results may not be accurate.

Despite no evidence of an HLA-D15S107 joint sharing interaction (see Section 3.4 Interaction Between Loci), there is evidence of an association between

IGF1R and IDDM which is modified by DR4 (ie. the IGF1R association was strongest in DR4+ families). This association may not have been reflected in the HLA-D15S107 interaction as the interaction looks at the allele sharing distribution in these markers whereas the DR4 dependent association is dependent only on DR4 positivity in the two affected children (ie. the other DR marker did not matter).

4.3 INSR

A previously reported INSR association was not supported in our data set. Raffel et al 1990 reported an association between INSR and IDDM using PCR-RFLP markers, however, using our PCR microsatellite marker, the same was not observed. This may be due to the increased information in a microsatellite marker compared to PCR-RFLP markers. LOD score analysis likewise did not show any evidence of the involvement of INSR in IDDM susceptibility (LOD=-1.46 at θ=0.00, average sibling sharing 0.53, p=0.195). However, in our data set (see Table 9), a significant difference was observed between the maternal and paternal transmitted allele frequencies of the INSR alleles especially in the D15S107 decreased sharing group (p=0.010 simplex, p=0.022 multiplex, p=0.000 double simplex; corrected p-values would be 2p for splitting the data into 2 subsets and would read p=0.020 simplex, p=0.044 multiplex, and p=0.000 double simplex), TH decreased sharing group (uncorrected p=0.036 simplex, 0.001 double simplex) and HLA increased sharing group (uncorrected p=0.027 simplex, 0.001 double simplex). (While other significant differences are observed, they occur only in one analysis group, eg.

double simplex analysis, and likely arose due to chance). Especially compelling is the observed difference in the D15S107 ≤ 50% sharing group as a significant difference is observed in all three analyses. While there is no significant evidence for a joint sharing interaction between INSR and D15S107, there is a low p-value (p=0.062, see Section 3.4) which may suggest that with more families an interaction may be observed.

It is interesting that D15S107 is so near the IGF1R gene, which encodes a receptor as does the INSR gene and that both are able to bind insulin albeit with different affinities. One is tempted to speculate that some susceptibility to IDDM may be due to different binding affinities of insulin itself to one of its receptors and one (or more) of its related receptors. Also, despite evidence that the 5' INS VNTR is the actual susceptibility locus in the insulin gene region, it is interesting to note that the IGF2 (insulin-like growth factor 2) gene is just downstream of the insulin locus. In addition, it has been suggested that the 5' INS VNTR may alter the transcription rate of any of the three genes in the region (ie. TH, INS, or IGF2).

4.4 Interactions between loci

Possible interactions between loci was examined for TH - HLA, TH-D15S107, D15S107-HLA, INSR-TH, INSR-D15S107, and INSR-HLA. No evidence indicative of any interaction between any two loci was seen in the data set despite evidence of marker sharing dependent associations with IDDM (eg. INSR sex transmitted effect in the D15S107 ≤ 50% sharing group). Since IDDM is a

multilocus disorder, analysis of single loci while informative to a degree, would not be as informative as multi-locus analyses. Also, if heterogeneity is involved, interactive models may be able to define sets of disease susceptibility genes rather than one set of many genes all involved in increasing susceptibility.

5.0 Conclusions

1. While TH is observed to be associated with IDDM (p=0.029 simplex, p=0.013 double simplex), actual linkage cannot be demonstrated despite using a much more informative marker than the INS VNTR. Thus, TH and the INS VNTR both show the paradoxical presence of an association with IDDM without showing any linkage.
2. There is a weak association between IGF1R microsatellite and IDDM (p=0.043 simplex, p=0.045 multiplex, p=0.022 double simplex). Affected sibling pair linkage analysis shows significantly increased D15S107 sharing in affected siblings (0.56, p=0.005), although no D15S107 association is detectable. The presence of an IGF1R association would indicate that the locus of interest is nearby (or even that IGF1R is the susceptibility locus).
3. There is no evidence for either an association between INSR and IDDM (contrary to a previously published report) nor for IDDM-INSR linkage.
4. There is weak evidence for an increased TH sharing dependent IGF1R association with IDDM, and conversely weak evidence for a decreased D15S107 sharing dependent TH association with IDDM, but no evidence for interaction between TH and D15S107 using a joint sharing analysis.
5. There is no significant evidence for joint sharing interaction between any two of the four markers examined ie. TH, HLA, D15S107, and INSR.
6. There is no evidence for a paternally dependent DR4 association between TH and IDDM within this data set (contrary to a previously published report).

7. Microsatellite markers can be used for association analyses as seen by the observed association between TH and IDDM. These markers do not mutate so fast as to destroy linkage disequilibrium between the microsatellite and disease locus.

8. No compelling evidence exists for a purely paternal or maternal dependent association for any of the markers examined.

9. Evidence for genetic heterogeneity exists. TH non-transmitted (control) allele frequencies for the two geographic locations are seen as significantly different from one another, suggesting that "Caucasian" allele frequencies can vary across small geographic regions of Europe, such that the British allele frequencies are different from the mixed European allele frequencies found in NA.

References

A'Hern O R, Heaton D A, Millward B A, Risley D, Pyke D A and Leslie R D G. (1988) The significance of the concordance rate for Type 1 (insulin-dependent) diabetes in identical twins. Diabetologia 31: 747-750.

Bach J-F. (1991) Insulin-dependent diabetes mellitus. Current Opinion in Immunology 3: 902-905.

Bae J-S, Yoon J-W. (1993) Determination of diabetogenicity attributable to a single amino acid, ala^{776}, on the polyprotein of Encephalomyocarditis virus. Diabetes 42: 435-43.

Ball E J, Dombrausky L, Hoover M, Capra J D, and Stastny P. (1987) Restriction fragment length polymorphism of the human T cell receptor alpha gene. Immunogenetics 26: 48-55.

Banatvala J E, Bryant J, Schernthaner G, Borkenstein M, Schober E, Brown D, De Silva L M, et al. (1985) Coxsackie B, Mumps, Rubella and Cytomegalovirus specific IgM responses in patients with Juvenile-onset insulin-dependent diabetes mellitus in Britain, Austria, and Australia. The Lancet i: 1409-1412.

Barnett A H, Eff C, Leslie R D G, and Pyke D A. (1981) Diabetes in Identical Twins. Diabetologia 20: 87-93.

Bell G I, Horita S, Karam J H. (1984) A polymorphic locus near the human insulin gene is associated with insulin-dependent diabetes mellitus. Diabetes 33: 176-183.

Bell G I, Haram J, Rutts W. (1981) Polymorphic DNA region adjacent to the 5' end

of the human insulin gene. Proceedings of the National Academy Science 78: 5759-5763.

Bennett S T, Lucassen A M, Gough S C L, Powell E E, Undlien D E, Pritchard L E, Merriman M E, et al. (1995) Susceptibility to human type 1 diabetes at IDDM2 is determined by tandem repeat variation at the insulin gene minisatellite locus. Nature Genetics 9: 284-292.

Bliss R. (1982) The Discovery of Insulin. The University of Chicago Press. McClelland & Stewart Limited. Toronto.

Bruserud O, Jervell J, Thorsby E. (1985) HLA-DR3 and -DR4 control T-lymphocyte responses to mumps and Coxsackie B4 virus: studies on patients with Type I (insulin-dependent) diabetes and healthy subjects. Diabetologia 28: 420-426.

Clerget-Darpoux F, Babron M C, Deschamps I, and Hors J. (1991) Complementation and maternal effect in insulin-dependent diabetes. American Journal of Human Genetics 49: 42-48.

Copeman J B, Cucca F, Hearne C M, Cornall R J, Reed P W, Ronningen K S, Undlien D E, et al. (1995) Linkage disequilibrium mapping of a type 1 diabetes susceptibility gene (IDDM7) to chromosome 2q31-q33. Nature Genetics 9: 80-85.

Cornall R J. (1993) Genetics of a multifactorial disease: autoimmune type I diabetes mellitus. Clinical Science 84: 257-262.

Cox N J, Baker L, and Spielman R S. (1988) Insulin-gene sharing in sib pairs with

insulin-dependent diabetes mellitus: No evidence for linkage. American Journal of Human Genetics 42: 167-172.

Davey M P, Meyer M M, Bakke A C. (1994) T cell receptor Vβ gene expression in monozygotic twins Discordance in CD8 subset and in disease states. Journal of Immunology 152: 315-321.

Davies J L, Kawaguchi Y, Bennet S T, Copeman J B, Cordell H J, Pritchard L E, Reed P E, et al. (1994) A genome-wide search for human type 1 diabetes susceptibility genes. Nature 371: 130-136.

Deschamps I, Hors J, Clerget-Darpoux F, Gardais E, Robert J, Marcelli-Barge A, Lestradet H, and Dausset J. (1990) Excess of maternal HLA-DR3 antigens in HLA DR3, 4 positive type 1 (insulin-dependent) diabetic patients. Diabetologia 33: 425-430.

Doberson M J, Scharff J E, Ginsberg-Fellner F, Notkins A L. (1980) Cytotoxic autoantibodies to beta cells in the serum of patients with insulin-dependent diabetes mellitus. New England Journal of Medicine 303: 1493-8.

Docherty K. (1992) 1992 R.D. Lawrence Lecture. The regulation of insulin gene expression. Diabetic Medicine. 9: 792-797.

Durner M and Greenberg D A (1992) Effects of heterogeneity and assumed mode of inheritance on Lod scores. American Journal of Human Genetics 42: 271-275.

Elbein S, Rotwein P, Permutt A M., Bell G I, Senz N and Karam J H. (1985) Lack of association of the polymorphic locus in the 5'-flanking region of the

human insulin gene and diabetes in American Blacks. Diabetes 34: 433-439.

Erlich H A, Gelfand D, Sninsky J J. (1991) Recent advances in the polymerase chain reaction. Science 252: 1643-1651.

Erlich H A, Bugawan T L, Scharf S, Nepom G T, Tait B, and Griffith R L. (1990) HLA-DQβ sequence polymorphism and genetic susceptibility to IDDM. Diabetes 39: 96-103.

Field L L. (1988) Invited Editorial: Insulin-dependent diabetes mellitus: a model for the study of multifactorial disorders. American Journal of Human Genetics 43: 793-798.

Field L L. (1991) Non-HLA region genes in insulin-dependent diabetes mellitus. Balliere's Clinical Endocrinology and Metabolism 5(3): 413-438.

Field L, Tobias R, Magnus T. (1994) A locus on chromosome 15q26 (IDDM3) produces susceptibility to insulin-dependent diabetes mellitus. Nature Genetics 8: 189-194.

Foster D W. (1988) Diabetes Mellitus. In: Scriver C R (ed) The Metabolic Basis of Inherited Disease, 6 ed. McGraw-Hill, New York, pp. 375-397.

Gelber C, Paborsky L, Singer S, McAteer D, Tisch R, Jolicoeur C, Buelow R, et al. (1994) Isolation of nonobese diabetic mouse T-cells that recognize novel autoantigens involved in the early events of diabetes. Diabetes. 43: 33-39.

Ghosh S, Palmer S M, Rodrigues N R, Cordell H J, Hearne C M, Cornall R J, Prins J-B, et al. (1993) Polygenic control of autoimmune diabetes in nonobese diabetic mice. Nature Genetics 4: 404-409.

Giddings S J, King C D, Harman K W, Flood J F, and Carnaghi L R. (1994) Allele specific inactivation of insulin 1 and 2, in the mouse yolk sac, indicates imprinting. Nature Genetics 6: 310-313.

Gill R G, Haskins K. (1993) Molecular mechanisms underlying diabetes and other autoimmune diseases. Immunology Today 14(2): 49-51.

Greenberg D A. (1993) Linkage analysis of "necessary" disease loci versus "susceptibility" loci. Am J Hum Genet 52: 135-143.

Hagopian W A, Michelsen B, Karlsen A E, Larsen F, Moody A, Grubin C E, Rowe R, et al. (1993) Autoantibodies in IDDM primarily recognize the 65,000-M_r rather than the 67,000-M_r isoform of glutamic acid decarboxylase. Diabetes 42: 631-636.

Hammond-Kosack M C U, Dobrinkski B, Lurz R, Docherty K, Kilpatrick M W. (1992) The human insulin gene linked polymorphic region exhibits an altered DNA structure. Nucleic Acids Research 20(2): 231-236.

Hashimoto L, Habita C, Beressi JP, Delepine M, Besse C, Cambon-Thomsen A, Deschamps I, et al. (1994) Genetic mapping of a susceptibility locus for insulin-dependent diabetes on chromosome 11q. Nature 371: 161-164.

Helmke K, Otten A, Willems W R, Brockhaus R, Mueller-Eckjardt G, Stief T, Bertrams J, et al. (1986) Islet cell antibodies and the development of diabetes mellitus in relation to mumps infection and mumps vaccination. Diabetologia 29: 30-33.

Honeyman M C, Cram D S, Harrison L C. (1993) Transcription Factor jun-B is

Target of Autoreactive T-cells in IDDM. Diabetes 42: 626-630.

Jenkins D, Penny M A, Uchigata Y, Jacobs K H, Mijovic C H, Hirata Y, Otani T, et al. (1992) Investigation of the mode of inheritance of Insulin-dependent diabetes mellitus in Japanese subjects. American Journal of Human Genetics 50: 1018-1021.

Julier C, Hyer R N, Davies J, Merlin F, Soularue P, Briant L, Cathelineau G, et. al. (1991) Insulin-IGF2 region on chromosome 11p encodes a gene implicated in HLA-DR4-dependent diabetes susceptibility. Nature 354: 155-159.

Karjalainen J, Marktin J M, Knip M, Ilonen J, Robinson B H., Savilahti E, Åkerblom H K, et al. (1992). A bovine albumin peptide as a possible trigger of insulin-dependent diabetes mellitus. The New England Journal of Medicine 327: 302-307.

Kennedy G C, German M S, and Rutter W J. (1995) The minisatellite in the diabetes susceptibility locus IDDM2 regulates insulin transcription. Nature Genetics 9: 293-298

Kostraba J N, Cruikshanks K J, Lawler-Heavner J, Jobim J F, Rewers M J, Gay E C, Chase P, et al. (1993) Early exposure to cow's milk and solid foods in infancy, genetic predisposition, and risk of IDDM. Diabetes 42: 288-95.

Kumar D, Gemayel N S, Deapen D, Kapadia D, Yamashita P H, Lee M, Dwyer J H, et al. (1993) North American twins with IDDM: genetic, etiological, and clinical significance of disease concordance according to age, zygosity, and interval after diagnosis in first twin. Diabetes 42: 1351-63.

Laakso M, and Pyörälä K. (1985) Age of Onset and Type of Diabetes. Diabetes Care 8(2): 114-117.

Lernmark Å, Bärmeier H, Dube S, Hagopian W, Karlsen A, and Wassmuth R. (1991) Autoimmunity of diabetes. Endocrinology and Metabolism Clinics of North America. Vol. 20(3). pp. 589-617.

Lucassen A M, Julier C, Beressie J-P, Boitard C, Froguel P, Lathrop M, Bell J I. (1993) Susceptibility to insulin dependent diabetes mellitus maps to a 4.1 kb segment of DNA spanning the insulin gene and associated VNTR. Nature Genetics 4: 305-310.

Luo D-F, Bui M M., Muir A, Malaren N K., Thomson G, She J-X. (1995) Affected-sib-pair mapping of a novel susceptibility gene to insulin-dependent diabetes mellitus (IDDM8) on chromosome 6q25-27. American Journal of Human Genetics 57: 911-919.

Macdonald M J, Gottschall J, Hunter J B, Winter K L. (1986) HLA-DR4 in insulin-dependent diabetic parents and their diabetic offspring: a clue to dominant inheritance. Proceedings of the National Academy of Science USA. 83: 7094-7053.

Maclaren N. (1992) Immunology of diabetes mellitus. Annals of Allergy 68: 5-9.

McCarthy B J, Dorman J S, Aston C E. (1991) Investigating genomic imprinting and susceptibility to insulin-dependent diabetes mellitus: an epidemiological approach. Genetic Epidemiology 8: 177-186.

Meloni R, Fougerousse F, Roudaut C, and Beckmann J S. (1992) Trinucleotide

repeat polymorphism at the human insulin-like growth factor I receptor gene (IGF1R). Nucleic Acids Research. 20(6): 1427.

Miller S A, Dykes D D and Polesky H F (1988) A simples salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Research 16: 1215.

Morel P A, Dorman J S, Todd J A, McDevitt H O, Trucco M. (1988) Aspartic acid at position 57 of the HLA-DQβ chain protects against type I diabetes: a family study. Proceedings of the National Academy of Science USA 85: 8111-8115.

Nepom G T. (1989) Determinants of genetic susceptibility in HLA-associated autoimmune disease. Clinical Immunology and Immunopathology 53 Suppl.: S52-S62.

Nepom G T. (1990) A unified hypothesis for the complex genetics of HLA associations with IDDM. Diabetes 39: 1153-1157.

Nerup J, Platz P, Anderson O, Ortved C M, Lyngsoe J, Poulsen J E, Ryder L P, et al. (1974) HL-A antigens and diabetes mellitus. The Lancet. ii: 864-868.

Orita M, Suzuki Y, Sekiya T, and Hayashi K. (1989) Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. Genomics 5: 874-879.

Owerbach D and Gabbay K H. (1995) The HOXD8 Locus (2q31) is Linked to Type I Diabetes. Diabetes 44: 132-136.

Owerbach D and Gabbay K H (1994) Linkage of the VNTR/insulin-gene and type

I diabetes mellitus: increased gene sharing in affected sibling pairs. American Journal of Human Genetics 54: 909-912.

Pak C Y, Eun H-M, McArthur R G, Yoon J-W. (1988) Association of Cytomegalovirus infection with autoimmune type I diabetes. The Lancet ii: 1-4

Palmer J P, McCulloch D K. (1991) Prediction and prevention of IDDM-1991. Diabetes 40: 943-947.

Permutt M A, Elbein S C. (1990) Insulin gene in diabetes. Analysis through RFLP. Diabetes Care. 13(3): 364-374.

Poduslo S E, Dean M, Kolch U, and O'Brien S J. (1991) Detecting high-resolution polymorphisms in human coding loci by combining PCR and single-strand conformation polymorphisms (SSCP) analysis. American Journal of Human Genetics. 49: 106-111.

Puers C, Hammond H A, Jin L, Caskey C T, Schumm J W. (1993) Identification of repeat sequence heterogeneity at the polymorphic short tandem repeat locus HUMTH01[AATG]$_n$ and reassignment of alleles in population analysis by using a locus specific allele ladder. American Journal of Human Genetics 53: 953-958.

Pujol-Borrell R, Solfevila G, Buscema M, Vives M, Badenas J, Marakian R, Bottazzo G F. (1989) Inappropriate expression of HLA class II molecules in endocrine epithelial cells: the phenomenon, the new experimental data and comparison with animal models. Journal of Autoimmunity 2(Suppl): 163-

169.

Raffel L J, Collins R M Jr, Xiang K, Klein R, Rotter J I. (1990) RFLP analysis of the insulin receptor gene: evidence for a role in IDDM susceptibility. American Journal of Human Genetics 47 (Suppl): A143.

Risch N. (1987) Assessing the role of HLA-linked and unlinked determinant of disease. American Journal of Human Genetics 40: 1-14.

Rjasanowski I, Vogt L, Michaelis D, Keilacker H, and Kohnert K. (1993) The frequency of diabetes in children of type I diabetic parents. Diabete et Metabolisme 19: 173-177.

Robinson B H, Dosch H-M, Martin J M, Akerblom H K, Savilahti E, Knip M, Ilonen J. (1993) A model for the involvement of MHC class II proteins in the development of Type I (insulin-dependent) diabetes-mellitus in response to bovine serum albumin peptides. Diabetologia 36: 364-368.

Roep B O, Arden S D, de Vries R R P, Hutton J C. (1990) T-cell clones from a type-1 diabetes patient respond to insulin secretory granule proteins. Nature 345: 632-634.

Rotter J I, Anderson C E, Rubin R, Congleton J E, Terasaki P I, Rimoin D L. (1983) HLA genotypic study of Insulin-dependent diabetes. The excess of DR3/DR4 heterozygotes allows rejection of the recessive hypothesis. Diabetes 32: 169-174.

Rubinstein Pablo, Suciu-Foca Nicole, Nicholson John F. (1977) Genetics of juvenile diabetes mellitus A recessive gene closely linked to HLA D and with 50 per

cent penetrance. New England Journal of Medicine 297: 1036-1040.

Rubinstein P. (1991) HLA and IDDM: facts and speculations on the disease gene and its mode of inheritance. Human Immunology 30: 270-277.

Sheehy M J. (1992) HLA and insulin-dependent diabetes. a protective perspective. Diabetes 41: 123-129.

Sheehy M J, Scharf S J, Rowe J R, Neme de Gimenez M H, Meske L M, Erlich H A, and Nepom B S. (1989). A disease-susceptible HLA haplotype is vest defined by a combination of HLA-DR and -DQ alleles. Journal of Clinical Investigation 83: 830-835.

Thomson G. (1995). Mapping disease genes: family-based association studies. American Journal of Human Genetics 57: 487-498.

Todd J A, et al. (1991) Genetic analysis of a complex, multifactorial disease, autoimmune type 1 (insulin-dependent) diabetes. Res. Immunol. 142: 483.

Toguchi Y, Ginsberg-Fellner F, Rubinstein P. (1985) Cytotoxic islet cell surface antibodies (ICSA) inpatients with Type I diabetes and their first-degree relatives. Diabetes 34: 855-860.

Tun R Y M, Smith M D, Lo S S S, Rook G A W, Lydyard P, Leslie R D G. (1994) Antibodies to heat shock protein 65 kD in type I diabetes mellitus. Diabetic Medicine 11: 66-70.

Virtanen S M, Räsänen L, Ylönen K, Aro A, Clayton D, Langholz B, Pitkäniemi J, et al. (1993) Early introduction of dairy products associated with increased risk of IDDM in Finnish children. Diabetes 42: 1786-90.

Vu T H, and Hoffman A R. (1994) Promoter-specific imprinting of the human insulin-like growth factor II gene. Nature: 371: 714-717.

Warram J H, Krolewski A S, Gottlieb M S, Kahn C R. (1984) Differences in risk of insulin-dependent diabetes in offspring of diabetic mothers and diabetic fathers. New England Journal of Medicine 311(3): 149-152.

Weber J L, May P E. (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. American Journal of Human Genetics 44: 388-396.

Wilkin T, Armitage M, Casey C, Pyke D A, Hoskins P J, Rodier M, Diaz J-L, et al. (1985) Value of insulin autoantibodies as serum markers for insulin-dependent diabetes mellitus. Lancet i: 480-482

Xiang K, Granqvist M, Seino S, and Bell G I. (1991) Microsatellite polymorphism in human insulin receptor gene (INSR) on chromosome 19. Nucleic Acids Research 19(18): 5094.

Yagi H, Matsumoto M, Kunimoto K, Kawaguchi J, Makino S, Harada M. (1992) Analysis of the roles of CD4+ and CD8+ T cells in autoimmune diabetes of NOD mice using transfer to NOD athymic nude mice. European Journal of Immunology 22: 2387-2393.

Yoon, J-W. (1991) Role of viruses in the pathogenesis of IDDM. Annals of Medicine. 23: 437-445.