ORIGINAL ARTICLES

The aggregation of the 5′ insulin gene polymorphism in insulin dependent (type I) diabetes mellitus families

L J Raffel, G A Hitman, H Toyoda, J H Karam, G I Bell, J I Rotter

Abstract

Population studies have suggested an increased frequency of small DNA insertions (class I alleles) 5′ to the insulin gene in insulin dependent (type I) diabetes mellitus (IDDM). The present study examined this relationship within families. Forty-one families with at least one diabetic offspring were studied. Analysis of the insulin gene polymorphism was performed by digestion of DNA with *BglI*, *SstI*, *RsaI*, or *PvuII* and hybridisation with an insulin gene probe or polymorphic region specific probes. An increased frequency of class I alleles was found among the parents of diabetics ($p = 0.02$), as well as a trend towards increased frequency of parents homozygous for class I alleles and matings of two homozygous subjects. This increased homozygosity for class I alleles was present in non-diabetic sibs as well ($p = 0.01$). These results show that ascertainment through an offspring with IDDM selects for families with high frequencies of homozygosity for the class I allele and thus suggests that the insulin gene polymorphism is indeed providing part of the genetic predisposition to IDDM. When the major portion of genetic predisposition is provided by other genes (estimates are that HLA accounts for 30 to 70% in IDDM), identification of additional susceptibility genes becomes difficult. Even when formal linkage analysis is uninformative, our studies indicate that analysis for aggregation of specific alleles within families is a useful approach to this problem.

Medical Genetics Birth Defects Center, Departments of Medicine and Pediatrics, Cedars-Sinai Medical Center, and UCLA School of Medicine, Los Angeles, USA.
L J Raffel
H Toyoda
J I Rotter

Department of Medicine, The London Hospital Medical College, London.
G A Hitman

Department of Medicine, Metabolic Research Unit, University of California, San Francisco, California, USA.
J H Karam

Howard Hughes Medical Institute, University of Chicago, Chicago, Illinois, USA.
G I Bell

Correspondence to Dr Raffel, Medical Genetics, Cedars-Sinai Medical Center, 8700 Beverly Boulevard, Los Angeles, California 90048-1869, USA.

Received 26 June 1991.
Revised version accepted 4 December 1991.

Insulin dependent (type I) diabetes mellitus is a disorder in which the genetic predisposition has been clearly established.[1 2] The association of IDDM with certain HLA types, particularly DR3 and DR4, has aided our understanding of the inheritance of this disorder.[1 2] It appears, however, that the HLA region does not account for all of the genetic predisposition to IDDM.[3–7] It has been estimated that loci in the HLA region account for from 30% to as much as 70% of the genetic predisposition.[4 6] This suggests that another locus or loci must provide the balance of genetic susceptibility to IDDM.

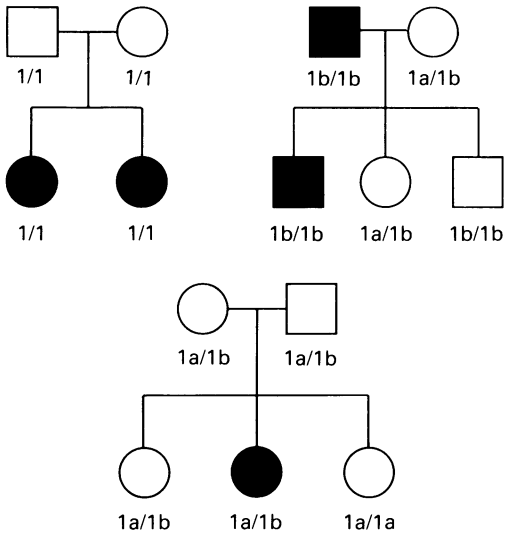
A number of candidate genes have been proposed as non-HLA susceptibility loci, but most have not been substantiated. There is some indication that the immunoglobulin heavy chain allotypes (Gm),[8 9] and the T cell receptor may be 'secondary loci'.[10–13] The most convincing data exist, however, for the polymorphic DNA region 5′ to the insulin gene. Although several studies in Caucasoid subjects have shown an increased frequency of class I alleles in IDDM populations as compared to controls,[14–16] attempts to show linkage of this locus with *IDDM* in family studies have been unsuccessful.[5 17–21]

The current study was also undertaken in an effort to determine if a relationship of the 5′ insulin gene polymorphism with IDDM existed within families. As described below, it was shown that family data support a role for the 5′ insulin gene polymorphism in IDDM predisposition, even though it is not possible to show classical linkage.

Material and methods

ASCERTAINMENT

Forty-one Caucasian families were ascertained through an offspring with IDDM as defined by National Diabetes Group criteria.[22] The sample was biased towards multiplex pedigrees to maximise the ability to test for linkage using affected sib pair and other linkage methodologies. Of the total, 11 were simplex families (the proband the only affected subject in the family), 20 were multiplex (two or more affected sibs), and 10 were multiplex/affected parent families (at least one offspring and one parent affected). All subjects consented to participate following explanation of the purpose of the study. Seventeen of the families were ascertained through the Bart's-Windsor Family Study, London, six at the University of California, San Francisco, and 18 at Harbor-UCLA Medical Center and Cedars-Sinai Medical Center, Los Angeles. In 37 of the 41 families, the genotype of both parents was determined, either by DNA analysis or by inference from DNA analysis of other family members. The British families were those previously studied.[16] In four families, it was not



Representative nuclear family pedigrees ascertained through IDDM probands. 5' insulin gene polymorphism genotypes are listed under each member of the pedigree. Those subjects 1/1 were truly homozygous for the class I allele within the confines of the methodology. Where different sizes of class I alleles could be detected these are described as 1a and 1b. The majority of the families were not informative for linkage because of a high frequency of class I alleles, even when different class I alleles could be distinguished within a family (for example, 1a, 1b).

possible to determine unequivocally both parental genotypes and these families were therefore excluded from analysis of the parental generation. All 41 families were used for analysis of offspring data.

DNA STUDIES

Blood was obtained by venepuncture for the isolation of DNA from all participants. The DNA samples were digested with the restriction endonucleases *Bgl*I, *Rsa*I, *Sst*I, or *Pvu*II according to manufacturers' instructions, electrophoresed in 0·85% to 1·0% agarose gels, and blotted onto nitrocellulose filters.[23 24] The digested DNA was then hybridised with a ^{32}P labelled probe. PHins 214, a 1650 base pair probe including the entire insulin gene sequence, was used for *Bgl*I, *Rsa*I, and *Sst*I digests.[25] PHins 310, an 879 base pair polymorphic region specific probe, was used for hybridisation to the *Pvu*II digests.

STATISTICAL ANALYSIS

Tests of proportions[26] were used for all comparisons of allele frequencies, genotype frequencies, and parental mating type frequencies.

Table 1 Frequency of the insulin gene polymorphism in Caucasian non-diabetic populations.

Population	No of cases	Insulin gene allelic frequency	
		Class I	Class III
San Francisco[14]	Control (n = 83)	0·67	0·33
St Louis[27]	Control (n = 33)	0·82	0·18
Copenhagen[15]	Control (n = 52)	0·73	0·27
London[16]	Control (n = 88)	0·67	0·33
Aggregate	Control (n = 256)	0·70	0·30

χ^2 for heterogeneity = 6·38, p > 0·09.

Results

Examples of representative pedigrees are shown in the figure. As can be seen, most families were uninformative for linkage owing to the common occurrence of homozygosity for class I alleles in all family members. In an effort to explain this unexpected finding, a comparison of the frequency of class I alleles in the non-diabetic parents was made with previously reported control population frequencies.[14–16 27] A control class I frequency of 0·70, the frequency in the aggregate data set of non-diabetic Caucasians, was chosen for comparison (table 1).

The frequency of class I alleles in the 65 non-diabetic parents, 0·78, was significantly greater than the control population (p = 0·02). When the nine parents with IDDM are also included, the frequency of class I alleles (0·79) is even more significantly increased (p < 0·01). Also increased in frequency among the parents is homozygosity for class I alleles and matings where both parents are homozygous for class I alleles, although these differences do not attain statistical significance when the parents with IDDM are removed from the sample (table 2).

The frequency of class I alleles was also determined in the diabetic and non-diabetic offspring (table 3). Consistent with the previously reported increase in class I alleles in IDDM subjects, the diabetic offspring had a higher incidence of homozygosity for class I alleles (0·72) than would be expected based on control population gene frequencies (p < 0·001). Although the incidence of homozygosity was slightly less in the non-diabetic offspring (0·62), it was still significantly greater than expected based on non-diabetic population frequencies (expected frequency 0·49, p = 0·01). The difference in homozygosity for class I alleles between the diabetic and non-diabetic offspring (0·72 *v* 0·62) was not statistically significant.

Discussion

The failure to show linkage of the insulin gene polymorphism with IDDM in family studies has led some investigators to question the validity of the population association of class I alleles with IDDM reported in several studies.[19] In actuality, this failure to show linkage is not surprising when it is assessed in the context that the insulin gene polymorphism is a 'secondary locus' rather than a primary (or major) locus. For major genes, a powerful way to confirm the association between a disease and a genetic marker is to show linkage in families. When the association is strong and the contribution of the locus is considerable, linkage can be shown using relatively small numbers of families, as in HLA and IDDM.[28] Unfortunately, the very large numbers of families needed to test adequately whether there is indeed linkage with a 'secondary locus' which accounts for a more minor proportion of genetic susceptibility makes it highly unlikely that studies of sufficient sample size will be available to test the hypothesis in this fashion. Less conventional analytic approaches to

Table 2 Homozygosity for class I alleles in parents of insulin dependent diabetics.

	Frequency of 1/1 homozygotes (expected = 0·49*)	Frequency of 1/1 × 1/1 matings (expected = 0·24)
Non-diabetic parents only	0·59 (n = 65), p = 0·06†	0·33 (n = 27), p = NS
All parents	0·61 (n = 74), p = 0·02	0·38 (n = 37), p = 0·02

* Expected calculated from population data (table 1).
† Statistical comparisons *v* expected.

Table 3 Homozygosity for class I alleles in sibs of insulin dependent diabetics.

	Type 1 diabetic probands (expected = 0·49*)	Non-diabetic sibs (expected = 0·49)
Proportion of 1/1 subjects	0·72 (n = 67), p < 0·001†	0·62 (n = 74), p = 0·01

* Expected calculated from population data (table 1).
† Statistical comparisons *v* expected.

family data are therefore necessary to assess the role of such 'secondary loci'.

The aggregation of class I alleles within families ascertained through an insulin dependent diabetic proband, as indicated in this study, is indeed a confirmation of the population association. The reason why it is not possible to show linkage is that the vast majority of pedigrees are not informative. The very reason that they are not informative is that most family members, regardless of diabetes status, are homozygous for the high risk class I alleles. The frequency of class I alleles in the non-diabetic parents and non-diabetic sibs is just as high as has been found in the diabetic population studies. Thus, by ascertaining these families through a diabetic proband, we have in actuality selected for families where the incidence of class I alleles will be high.

Aggregation of high risk alleles in unaffected family members would not be expected to occur with genes accounting for a large proportion of disease susceptibility. For a major locus, such as the HLA region in IDDM, linkage will be seen rather than familial aggregation. This is shown in table 4, where the proportion of subjects homozygous for diabetes associated HLA class II alleles (in this case HLA-DR3 and DR4) is significantly greater in IDDM cases than in their non-diabetic sibs (0·507 *v* 0·366, p < 0·001). This degree of difference contrasts with the analogous comparison reported here for the insulin gene polymorphism (table 3). Thus, the aggregation of class I insulin gene polymorphism alleles seen within the families in the current study, coupled with the failure to show linkage, suggests both that the population association previously reported is genuine and that this locus accounts for a lesser proportion of disease susceptibility than does the diabetes locus or loci in the HLA region.

*Table 4 Homozygosity for diabetes associated HLA class II alleles in insulin dependent diabetic subjects and their non-diabetic sibs.**

	Type 1 diabetic probands (n = 589)	Non-diabetic sibs (n = 153)	
Proportion homozygous for HLA class II alleles†	0·507	0·366	p < 0·001

* Data from continuing family studies of insulin dependent diabetes mellitus.[29 30]
† 'Homozygosity' for diabetes associated alleles defined as HLA DR3/3, 4/4, or 3/4.

The demonstration of probable segregation of diabetogenic 'disease' alleles as compared to control 'non-disease' alleles at the insulin gene polymorphism locus within IDDM families as reported by Thomson *et al*[31] using the haplotype relative risk method,[32] suggests that, with large sample sizes and the ability to distinguish all parental alleles, it can be possible to use family data to confirm the role of 'secondary loci' in disease predisposition. This method is fairly insensitive, however, and in the face of relatively small sample size or homozygosity for parental alleles will fail to identify accurately the importance of a minor allele. The problem of homozygosity at the insulin gene locus for parental alleles can be circumvented by the use of additional polymorphic markers in linkage with the insulin gene. Although Owerbach *et al*,[33] in a study of 27 families, constructed haplotypes based on DNA insertions at the insulin gene, polymorphism of the tyrosine hydroxylase locus, and a hypervariable region adjacent to the Harvey ras (HRAS1) oncogene and found an increase in a particular haplotype of the HRAS1 gene with IDDM, they still did not perform a formal linkage analysis. Testing for excess aggregation of the postulated high risk allele within families, as has been done in this study and the study of Owerbach *et al*,[33] is thus likely to be a more sensitive way to ascertain such secondary loci. This method will enable confirmation of a population association using family studies with sample sizes that can be more reasonably attained in most such investigations.

The primary role of the HLA region in genetic predisposition to IDDM is well established, although it is now increasingly being considered that other, non-HLA linked loci must also be involved.[4-6 8] Possibly the strongest data for involvement of non-major histocompatibility loci came from murine models of IDDM, the BB rat and NOD mouse.[34-36] These other loci, however, can each only account for a minor proportion of total genetic susceptibility of IDDM in man.[6] The aggregation of high risk alleles within families, such as has been shown here for the insulin gene polymorphism, is an indication of a high frequency of the polygene(s) which increase the likelihood of a subject developing IDDM. Within these high risk families, however, the determinant of which subject or subjects will actually develop IDDM is HLA, the major susceptibility locus. However, it is possible that if not for the occurrence of the 'secondary locus' high risk alleles within the family as well, the HLA related predisposition would be inadequate to produce disease.

We would like to thank Dr P Yen (Harbor-UCLA Medical Center, Los Angeles) for technical advice. We also thank Professor D J Galton (St Bartholomew's Hospital, London) and Professor G F Bottazzo (The Royal London Hospital, London) for their earlier involvement in the insulin gene work on the Bart's-Windsor Family Study. Part of this

work was supported by a grant from the Stuart Foundations, the William and Shyrlee Hamilton Fund, and the Cedars-Sinai Board of Governors' Chair in Medical Genetics (JIR).

1 Rotter JI. The modes of inheritance of insulin dependent diabetes. *Am J Hum Genet* 1981;**33**:835–51.
2 Thomson G, Robinson WP, Kuhner MK, *et al.* Genetic heterogeneity, modes of inheritance, and risk estimates for a joint study of Caucasians with insulin-dependent diabetes mellitus. *Am J Hum Genet* 1988;**43**:799–816.
3 Field LL. Insulin dependent diabetes mellitus: a model for the study of multifactorial disorders. *Am J Hum Genet* 1988;**43**:793–8.
4 Risch N. Assessing the role of HLA-linked and unlinked determinants of disease. *Am J Hum Genet* 1987;**40**:1–14.
5 Easton DF. Linkage analysis and genetic models for type 1 diabetes. *Genet Epidemiol* 1989;**6**:83–8.
6 Rotter JI, Landaw EM. Measuring the genetic contribution of a single locus to a multilocus disease. *Clin Genet* 1984;**26**:529–42.
7 Rich SS. Mapping genes in diabetes: genetic epidemiological perspective. *Diabetes* 1990;**39**:1315–9.
8 Rich SS, Weitkamp LR, Guttormsen G, Barbosa J. Gm, Km, and HLA in insulin-dependent type 1 diabetes mellitus: a log-linear analysis of association. *Diabetes* 1986;**35**:927–32.
9 Field LL, Dizier MH, Anderson CE, Spence MA, Rotter JI. HLA-dependent Gm effects in insulin-dependent diabetes: evidence from pairs of affected siblings. *Am J Hum Genet* 1986;**39**:640–7.
10 Hoover ML, Capra JD. HLA and T-cell receptor genes in insulin-dependent diabetes mellitus. *Diabetes Metab Rev* 1987;**3**:835–56.
11 Millward BA, Welsh KI, Leslie RDG, Pyke DA, Demaine AG. T-cell receptor beta chain polymorphisms are associated with insulin-dependent diabetes. *Clin Exp Immunol* 1987;**70**:152–7.
12 Ito M, Tanimoto M, Kamura H, *et al.* Association of HLA-DR phenotypes and T-lymphocyte-receptor beta-chain-region RFLP with type 1 diabetes in Japanese. *Diabetes* 1988;**37**:1633–6.
13 Niven MJ, Caffrey C, Moore RH, *et al.* T-cell receptor beta-subunit gene polymorphism and autoimmune disease. *Hum Immunol* 1990;**27**:360–7.
14 Bell GI, Horita S, Karam JH. A polymorphic locus near the human insulin gene is associated with insulin-dependent diabetes mellitus. *Diabetes* 1984;**31**:176–83.
15 Owerbach D, Nerup J. Restriction fragment length polymorphism of the insulin gene in diabetes mellitus. *Diabetes* 1982;**31**:275–7.
16 Hitman GA, Tarn AC, Winter RM, *et al.* Type 1 (insulin-dependent) diabetes and a highly variable locus close to the insulin gene on chromosome 11. *Diabetologia* 1985;**28**:218–22.
17 Cox NJ, Spielman RS. The insulin gene and susceptibility to type 1 diabetes. *Genet Epidemiol* 1989;**6**:65–9.
18 Dizier MH, Clerget-Darpoux F, Hochez J. Segregation analysis of two genetic markers in type 1 diabetes families under two locus models. *Genet Epidemiol* 1989;**6**:71–5.
19 Donald JA, Barendse W, Cooper DW. Linkage studies of HLA and insulin gene restriction fragment length polymorphisms in families with IDDM. *Genet Epidemiol* 1989;**6**:77–81.
20 Fimmers R, Neugebauer M, Dennert J, Wienker T, Baur MP. Association and sibpair analysis for the HLA, Gm, Km and insulin polymorphisms in multiplex IDDM families. *Genet Epidemiol* 1989;**6**:107–12.
21 Cox NJ, Baker L, Spielman RS. Insulin gene sharing in sib pairs with insulin-dependent diabetes mellitus: no evidence for linkage. *Am J Hum Genet* 1988;**42**:167–72.
22 National Diabetes Data Group International Workgroup. Classification of diabetes mellitus and other categories of glucose intolerance. *Diabetes* 1979;**28**:1039–57.
23 Southern EM. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 1975;**98**:503–17.
24 Wahl GM, Stern M, Stark GR. Efficient transfer of large fragments from agarose gels to diazobenzylomethyl-paper and rapid hybridization using dextran sulfate. *Proc Natl Acad Sci USA* 1979;**76**:615–9.
25 Bell GI, Karam JH, Rutter WJ. Polymorphic DNA region adjacent to the 5′ end of the human insulin gene. *Proc Natl Acad Sci USA* 1981;**78**:5758–66.
26 Dixon WJ, Massey FJ Jr. *Introduction to statistical analyses.* New York: McGraw-Hill, 1983.
27 Rotwein PS, Chirgwin J, Province M, *et al.* Polymorphism in the 5′ flanking region of the human insulin gene: a genetic marker for non-insulin dependent diabetes. *N Engl J Med* 1983;**308**:65–71.
28 Cudworth AG, Woodrow JC. Evidence for HLA linked genes in juvenile diabetes mellitus. *BMJ* 1975;**ii**:133–5.
29 Maclaren N, Riley W, Skordis N, *et al.* Inherited susceptibility to insulin-dependent diabetes is associated with HLA-DR1, while DR5 is protective. *Autoimmunity* 1988;**1**:197–205.
30 Vadheim CM, Rotter JI, Maclaren NK, Riley WJ, Anderson CE. Preferential transmission of diabetic alleles within the HLA gene complex. *N Engl J Med* 1986;**315**:1314–8.
31 Thomson G, Robinson WP, Kuhner MK, Joe S, Klitz W. HLA and insulin gene associations with type 1 diabetes. *Genet Epidemiol* 1989;**6**:155–60.
32 Falk CT, Rubinstein P. Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. *Ann Hum Genet* 1987;**51**:227–33.
33 Owerbach D, Gunn G, Gabbay KH. Multigenic basis for type 1 diabetes: association of HRAS1 polymorphism with HLA-DR3, DQw2, DR4, DQw8. *Diabetes* 1990; **39**:1504–9.
34 Colle E, Guttmann RD, Seemayer TA, Michel F. Spontaneous diabetes mellitus syndrome in the rat. IV. Immunogenetic interactions of MHC and non-MHC components of the syndrome. *Metabolism* 1983;**32**(suppl 1):54–61.
35 Prochazka M, Leiter E, Serreze DV, Coleman DL. Three recessive loci required for insulin dependent diabetes in non-obese mice. *Science* 1987;**237**:286–9.
36 Todd JA, Aitman JA, Cornall R, *et al.* Genetic analysis of autoimmune type 1 diabetes mellitus in mice. *Nature* 1991;**351**:542–7.