Journal of Medical Genetics 1986, **23**, 210–216

DNA polymorphic haplotypes on the short arm of chromosome 11 and the inheritance of type I diabetes mellitus

G A A FERNS, G A HITMAN*, R TREMBATH, L WILLIAMS, ANNE TARN, E A GALE, AND D J GALTON

From the Departments of Diabetes and Lipid Research and Diabetes and Immunogenetics, St Bartholomew's Hospital, London EC1A 7BE.

SUMMARY The linked polymorphic loci 5′ to the insulin gene and 3′ to the c-Harvey-ras-1 (c-Ha-ras) gene, both localised to the short arm of chromosome 11, have been studied in 14 type I diabetic pedigrees. The use of a cloned gene probe corresponding to the polymorphic locus adjacent to the insulin gene, in combination with the restriction endonuclease *Pvu*II, has permitted an improvement in the resolution of sizes of insert at this locus. An *Msp*I restriction fragment length polymorphism at the c-Ha-ras proto-oncogene locus (4 cM upstream from the insulin gene) was used to identify parental insulin gene related alleles unambiguously, and subsequently a pedigree analysis was performed to determine whether subclasses of inserts at this locus track with insulin dependent diabetes. Segregation analysis demonstrated no linkage between the polymorphic loci 5′ to the insulin gene, nor 3′ to the c-Ha-ras, and type I diabetes. However, a similar analysis confirmed an association between the HLA locus on chromosome 6 and insulin dependent diabetes.

Reiterated DNA sequences are dispersed throughout the human genome and have been found to be individually specific in their distribution.[1] One such sequence is the polymorphic locus 5′ to the insulin gene on chromosome 11.[2] The variability of this locus is due to an insertion, 363 base pairs (bp) from the start codon,[3] of a 14 bp oligonucleotide consensus sequence, repeated in tandem, to yield three main allelic classes: a long (class 3) insert of 1600 to 2200 bp, a short (class 1) insert of 0 to 600 bp, and an intermediate (class 2) insert rarely found in Caucasians. Bell *et al*[4] and Hitman *et al*[5] have recently shown an association between the class 1 insertion and insulin dependent diabetes in Caucasians. However, in the latter study, pedigree analysis of 17 type I diabetic families did not demonstrate linkage between the class 1 insert and diabetes. Hitman *et al*[5] therefore excluded the possibility of close linkage between the polymorphic locus and diabetes, using a single locus recessive model. Four centiMorgans 5′ to the insulin gene is yet another hypervariable region, located 1400 bp 3′ to the c-Ha-ras-1 proto-oncogene.[6 7] It is composed of a tandemly repeated 28 bp oligonucleotide. One difficulty encountered in previous studies of the insulin gene hypervariable locus has been the problem of differentiating alleles of the same insert class. We have used two approaches to solve this problem. The restriction endonuclease *Pvu*II excises the insulin gene associated polymorphic locus on a small restriction fragment (approximately 800 bp for a class 1 insert and 2200 bp for a class 3 insert) (fig 1b). This permits an improved resolution of insert length measurement. Because in some pedigrees it may still be difficult to differentiate parental alleles which are of the same subclass, we decided to use the c-Ha-ras associated polymorphic locus as an additional allelic marker. Although crossover events may be expected to occur between the two polymorphic loci (four out of 100 meioses), such recombinants would be apparent from the non-Mendelian inheritance of alleles within pedigrees. We have applied these techniques to the study of 14 insulin dependent diabetic pedigrees, in an attempt to discover if subclasses of insulin gene polymorphic inserts track with diabetes. Previous workers[5] have suggested that an inability to differentiate insert subclasses may explain the paradox of

*Present address: Medical Unit, The London Hospital, London E1.
Received for publication 9 September 1985.
Accepted for publication 17 October 1985.

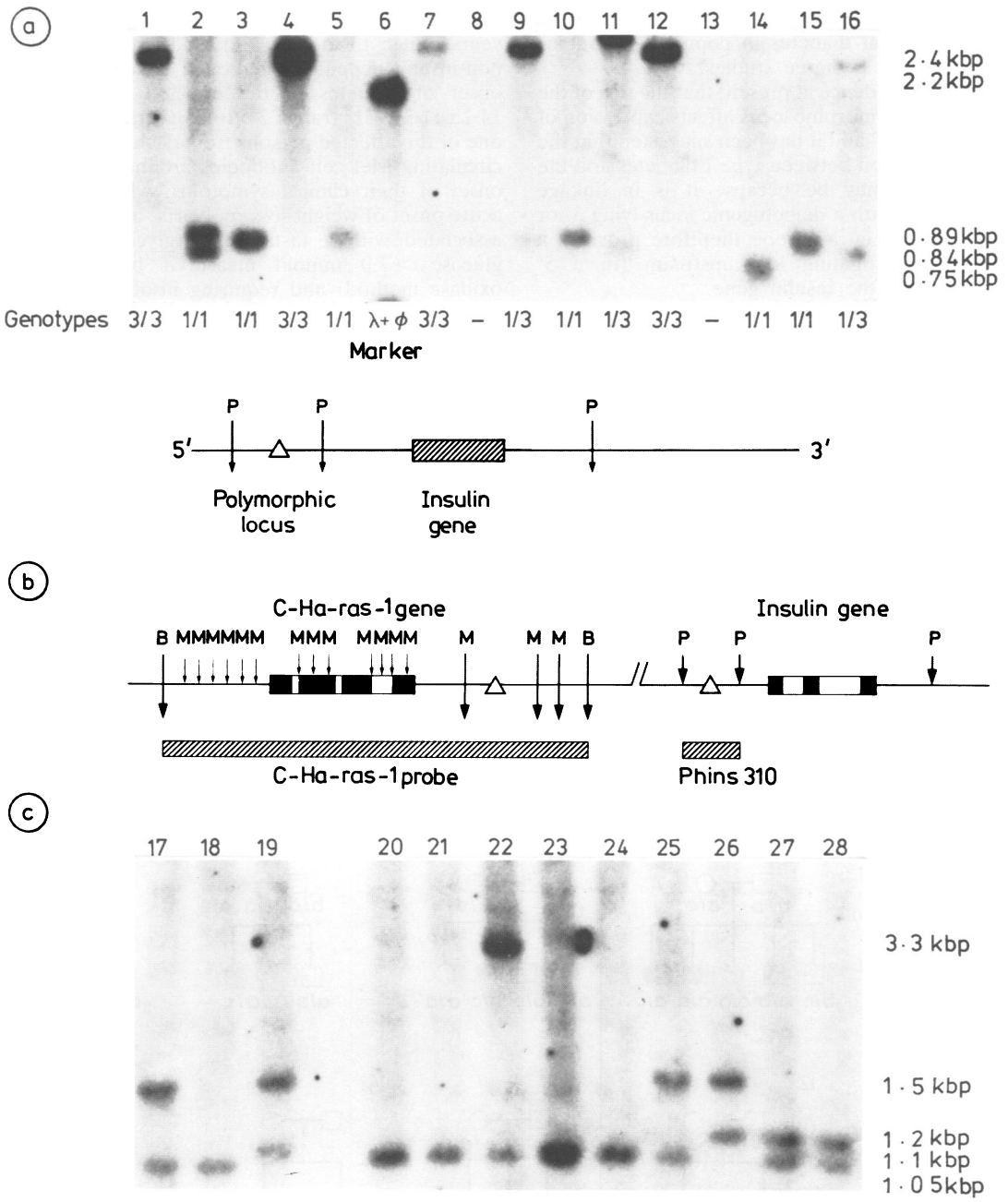

FIG 1 *(a) A representative autoradiogram of a Southern blot containing DNA digested with PvuII and hybridised to a probe complementary to the hypervariable region 5′ to the insulin gene. Fragment sizes are indicated in kilobase pairs (kbp). Genotypes are also indicated below each lane. (b) A simplified gene map of the c-Ha-ras-1 and insulin gene loci on 11p. B, M, and P indicate the cleavage sites for the enzymes BamHI, MspI, and PvuII respectively. The hatched regions represent the areas corresponding to the gene probes used. The solid blocks are the coding regions. (c) A representative autoradiogram of a Southern blot containing DNA digested with MspI and hybridised to a probe complementary to the c-Ha-ras gene. Fragment sizes are indicated in kbp.*

disease association between the class 1 insertion and insulin dependent diabetes in populations, but no close linkage in pedigree studies.

There is no evidence at present that the size of the insert at the polymorphic locus affects expression of the insulin gene,[8] and it has been suggested that the disease association between type I diabetes and the class 1 insert may be because it is in linkage disequilibrium with a diabetogenic locus lying 3′ or 5′ to it. The c-Ha-ras probe therefore provides a means of investigating loci upstream (in a 5′ direction) from the insulin gene.

Methods

PEDIGREE STUDY

Fourteen families (fig 2) were assembled, all from the Bart's-Windsor study, and comprised a total of 27 type I diabetics (14 female, 13 male, mean age 26·3±13·0 years, range 13 to 50 years), 38 non-diabetics (18 female, 20 male, mean age 35·5±16·7 years, range 13 to 58 years), and one person with non-insulin dependent diabetes. The mean age of onset of diabetes in the affected subjects was 14·1±11·3 years (range 2·5 to 53 years) and at least one of the affected persons from each pedigree had circulating islet cell antibodies around the time of onset of their clinical symptoms, which included acute onset of weight loss, polyuria, and polydipsia associated with a fasting hyperglycaemia (blood glucose >7·0 mmol/l measured by a glucose oxidase method) and requiring insulin therapy.

PREVALENCE STUDY

Forty-four unrelated subjects (21 female, 23 male, mean age 28·7±17·0 years) with type I diabetes mellitus (mean age of onset 12·8±9·9 years) were obtained from the outpatients department of St Bartholomew's hospital or from the Bart's-Windsor study. Controls (20 male, 17 female, mean age

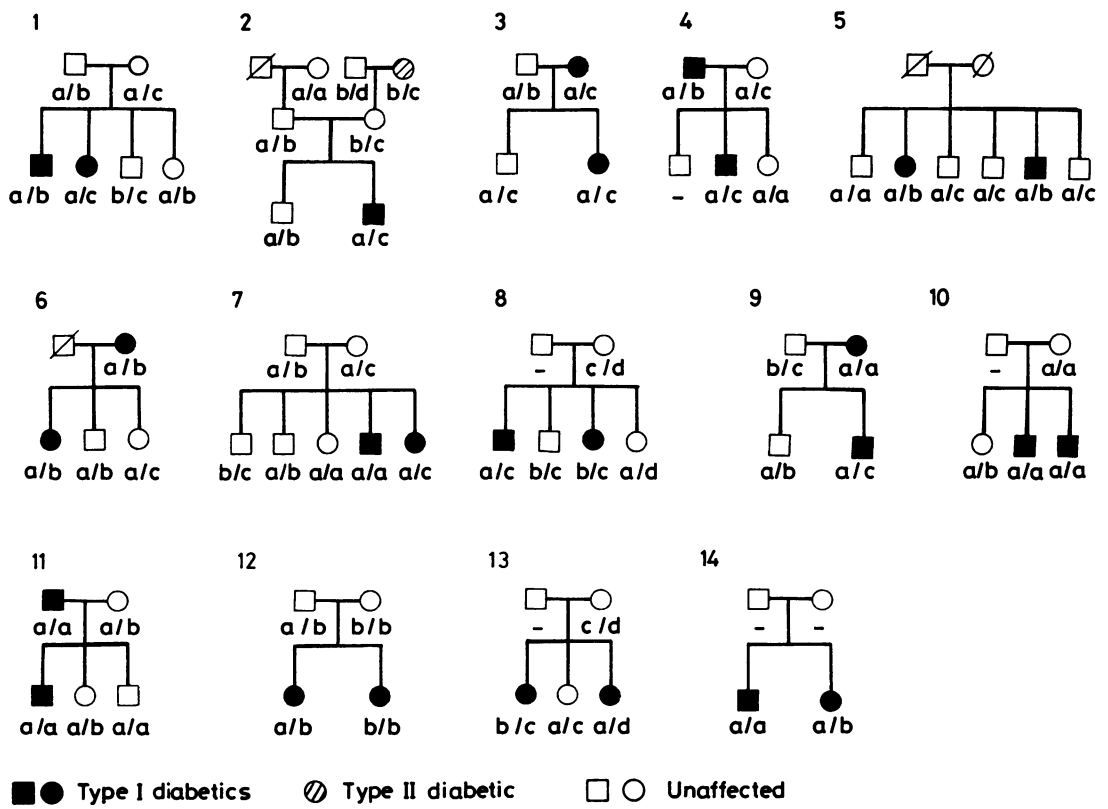

FIG 2 *Fourteen insulin dependent diabetic pedigrees. Letters (a, b, c, d) indicate different insulin/c-Ha-ras haplotypes and have been separately defined for each family.*

44·2±9·6 years) were obtained from a health screening centre; none had a family history of diabetes mellitus and overt diabetes was excluded by measuring a fasting blood glucose (fasting blood glucose <6·0 mmol/l).

DNA ANALYSIS

DNA was obtained from 10 ml of whole blood and anticoagulated with 1 mg/ml disodium EDTA, essentially using the method of Kunkel *et al.*[9] Red cells were lysed with a sucrose buffer (0·32 mmol/l sucrose, 5 mmol/l $MgCl_2$, 1% Triton-X-100, 10 mmol/l Tris, pH 7·5). Leucocyte nuclei were pelleted by centrifuging at 10 000 *g* for 10 minutes at 4°C, resuspended in 4·5 ml saline-EDTA (75 mmol/l NaCl, 24 mmol/l EDTA, pH 8·0), lysed with sodium dodecyl sulphate (SDS) (17 mmol/l), and treated with proteinase K (200 μg/ml) (Sigma Chemicals, St Louis, USA). After incubation for 16 hours at 37°C, the DNA containing solution was extracted once with 5 ml Tris buffered phenol (pH 8·0) and twice with 10 ml chloroform-iso-octanol (24:1 v/v). The phases were separated by centrifuging at 10 000 *g* for 10 minutes at 10°C, the aqueous phase being recovered at each step. DNA was then precipitated by adding 0·5 ml sodium acetate (3 mol/l, pH 5·0) and 11 ml absolute ethanol at room temperature. The precipitated DNA was recovered and dissolved in Tris-EDTA buffer (10 mmol/l Tris, 1 mmol/l EDTA, pH 7·5). DNA 8 μg was digested with the restriction enzymes *Msp*I or *Pvu*II in accordance with the manufacturer's (BRL Ltd, Cambridge, UK) instructions. The restricted fragments were electrophoresed on a 1·2% agarose gel, transferred to a nitrocellulose sheet (Schleicher and Schull) using the technique of Southern,[10] and then baked for three hours at 80°C. The filters were then placed into a hybridisation buffer (0·75 mol/l NaCl, 75 mmol/l sodium citrate, 50 μg/ml sheared herring sperm DNA, Denhardt's solution, 3·5 mmol/l SDS, 5% dextran sulphate (Pharmacia)) with a ^{32}P labelled plasmid, containing either the entire c-Ha-ras-1 proto-oncogene sequence (including 3′ and 5′ flanking regions)[6] or the polymorphic locus 5′ to the insulin gene (Phins 310)[11] at 65°C for 48 hours. The filters were then washed once for 30 minutes in each of the following solutions at 65°C: (1) 0·54 mol/l NaCl, 45 mmol/l sodium citrate (3 × SSC), pH 7·5, Denhardt's, 3·5 mmol/l SDS; (2) 1 × SSC, 3·5 mmol/l SDS; and (3) 0·1 × SSC, 3·5 mmol/l SDS. Hybridisation bands were then visualised by autoradiography at −70°C using Kodak CURIX MR4 *x*-ray film with X-ograph speed X intensifying screens. Band sizes were determined by running *Hind*III fragments of lambda phage and *Hae*II fragments of PhiX phage (BRL) with each batch of DNA digests.

TABLE 1 *Comparison of type I diabetics with unaffected subjects: clinical and genotype data of pedigrees and populations.*

Study group of type I diabetics	*Subgroup*	*No*	*Male*	*Female*	*Mean age (y)*	*Age of onset (y)*	*Insulin genotypes (%)*			*Allelic frequencies*		*Mean size of class I allele (bp)*	*Mean size of c-Ha-ras allele (bp)*
							1/1	*1/3*	*3/3*	*Class 1*	*Class 3*		
Pedigrees (14 families)	Diabetics	27	13	14	26·3± 13·0	14·1± 11·3	22 (82)	5 (18)	0* (0)	0·908	0·092	810± 62‡	1320± 510‡
	Unaffected relatives	38	20	18	35·5± 16·7	—	26 (68)	11 (29)	1 (3)	0·829	0·171	790± 50	1360± 520
Population	Unrelated diabetics	44	23	21	28·7± 17·0	12·8± 9·9	33 (75)	11 (25)	0‡ (0)	0·875	0·125	790± 65§	—
	Unrelated Caucasian controls	37	20	17	44·2± 9·6	—	22 (62)	11 (28)	4 (10)	0·743	0·257	790± 78	—

Constructing a 2 × 2 contingency table containing absolute allelic frequencies and applying a χ^2 test.
*p>0·05 compared to unaffected relatives.
†p<0·05 compared to unrelated Caucasian controls.
‡p>0·05 compared to unaffected family members using an unpaired *t* test.
§p>0·05 compared to Caucasian controls using an unpaired *t* test.

HLA TYPING

This was performed as previously described[5] using lymphocytes isolated from 20 ml citrated blood and a two colour fluorescent technique.

STATISTICAL ANALYSIS

An unpaired *t* test was used to compare mean allelic sizes of the hypervariable loci of insulin and c-Ha-ras genes. Allelic frequencies for the insulin locus were compared by calculating the absolute frequency of class 1 and class 3 alleles within each group, constructing a 2 × 2 contingency table, and applying a χ^2 test. The distribution of haplotypes within the pedigrees was compared by combining relevant cells into a 2 × 2 contingency table and applying Fisher's exact test.

Results

Autoradiography of a Southern blot containing DNA digested with the restriction enzyme *Pvu*II and hybridised to Phins 310, labelled with ^{32}P, resulted in hybridisation fragments of about 800 bp for a class 1 insertion and about 2200 bp for a class 3 insertion (fig 1a). About 50% of the parents of the pedigrees were unambiguously heterozygous at this locus. However, it was not possible to differentiate class 1 inserts that differed in size by less than about 50 bp.

Autoradiography of Southern blots containing DNA digested with the restriction enzyme *Msp*I and hybridisation to a genomic DNA probe for c-Ha-ras-1, labelled with ^{32}P, resulted in hybridisation fragments of very variable size. In the pedigrees we investigated, fragments between 900 and 3300 bp were obtained (fig 1c). *Msp*I cleaves the c-Ha-ras locus at a number of sites (fig 1b), resulting in the generation of many small fragments. Most of these are not detectable, but the fragment containing the polymorphic locus is sufficiently large to give a hybridisation signal.

Haplotypes, comprising the alleles at the hypervariable insulin and c-Ha-ras loci, could thus be defined for each subject. These haplotypes were assigned letters (a, b, c, d) for simplicity (fig 2), but represent different combinations of alleles for each family.

Segregation analysis of the pedigrees revealed that, within a particular family, subjects haplo-identical to the diabetic proband on both their homologous chromosomes had a 4/13 (31%) chance of also having diabetes, whereas if they were haplo-identical to the proband on one chromosome, or non-haplo-identical, their chance of diabetes was 9/38 (33%). This difference, however, failed to reach statistical significance. Similarly, subjects who were haplo-identical to the diabetic proband on one or both chromosomes had a 12/48 (25%) chance of diabetes, whereas 1/3 (33%) of subjects who shared no common haplotype with the proband had diabetes (table 1). These data suggest that the inheritance of type I diabetes mellitus is not strongly related to the insulin-c-Ha-ras locus.

A similar analysis was performed for the HLA locus on chromosome 6 (table 2). It was found, in contrast, that a significantly greater proportion of first degree relatives who were HLA identical to the diabetic proband were also diabetic (6/8), compared with those who were HLA non-identical (7/48) ($p<0.001$).

The inheritance of the insulin-c-Ha-ras haplotypes were found to be Mendelian in all cases and no recombinational events were detected in the pedigrees examined in this study, confirming the close linkage between the insulin gene and c-Ha-ras loci.

The mean size of the class 1 insulin gene related polymorphic insert was 810±62 bp for the affected family members and 790±55 bp in the unaffected subjects ($p>0.05$) (fig 3). Similarly, there was no significant difference in the mean size of the insert at the c-Ha-ras related polymorphic locus (affected 1320±510 bp, unaffected 1360±520 bp).

TABLE 2 *Number of affected and unaffected relatives who possess shared haplotypes compared with probands of 14 type I diabetic pedigrees.*

Number of homologous chromosomes at which haplo-identical to proband	*Ins/c-Ha-ras locus (chromosome 11)*		*HLA (DR) locus (chromosome 6)*	
	Diabetic	*Non-diabetic*	*Diabetic*	*Non-diabetic*
	No	*No*	*No*	*No*
Neither	1	2	0	7
One chromosome	8	27	7	34
Both chromosomes	4	9	6	2

The c-Ha-ras alleles were used to identify insulin gene related alleles unambiguously. Subjects HLA haplo-identical on both homologous chromosomes (chromosomes 6) were defined as HLA identical, other subjects as HLA non-identical. Similarly, subjects ins/c-Ha-ras identical on both homologous chromosomes (chromosomes 11) were defined as ins-C-Ha-ras identical and others as c-Ha-ras non-identical. 2 × 2 contingency tables were constructed, containing numbers of family members who were diabetic or unaffected, either HLA identical or non-identical, or ins/c-Ha-ras identical or non-identical. A Fisher's exact test was then applied, $p<0.001$ and $p>0.05$ respectively.

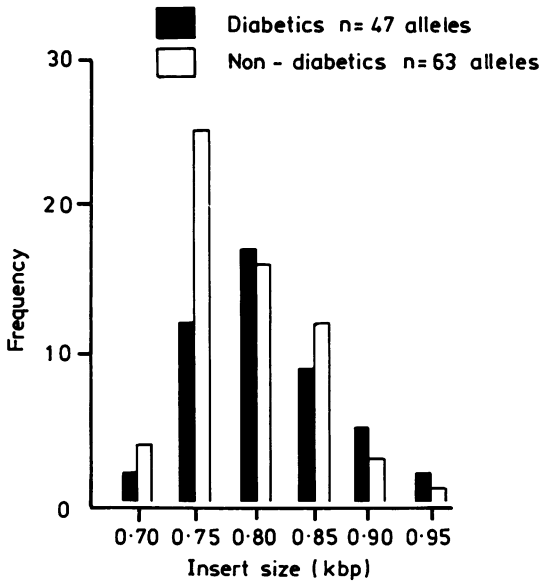

FIG 3 *A histogram showing the distribution of sizes of class 1 insert 5′ to the insulin gene in diabetic and non-diabetic subjects of 14 diabetic pedigrees.*

Analysis of the population data showed no significant difference between the size of the class 1 insert in type I diabetics (mean 789±67 bp) compared to normal controls (mean 791±78 bp, $p > 0{\cdot}05$). However, we confirmed the finding of previous studies that the class 1 insert is more prevalent in type I diabetics compared with Caucasian controls, comparing allelic frequencies in a 2 × 2 contingency table and applying a χ^2 test ($p < 0{\cdot}05$).

Discussion

One difficulty encountered by Hitman *et al*[5] has been overcome by using a combination of gene probes in order to track the inheritance of hypervariable insulin alleles through pedigrees. The ability to differentiate these alleles, even though they may be of the same subclass, has permitted the inclusion in our analysis of families that were initially uninformative in Hitman's study. The inability to distinguish between subclasses of insulin allele has previously been cited as a reason for the discrepancy between the results of population and pedigree studies. However, using a more refined technique, we have found no difference between the distribution of insert sizes in type I diabetics compared to non-diabetics, either within pedigrees or populations. Analysis of the c-Ha-ras locus in the pedigrees suggests that if a diabetogenic locus exists on the short arm of chromosome 11, it is probably not 5′ to the insulin gene.

Segregation analysis of the pedigrees showed that irrespective of whether a person shared the same insulin-c-Ha-ras haplotype as an affected family member, the chance of him also being diabetic was still about 1/4, suggesting that if the polymorphic locus of the insulin gene is involved in the genesis of type I diabetes, it is probably a minor component. There is strong evidence that type I diabetes is either a bigenic or polygenic disease. Chromosome 6 markers have been strongly implicated in the aetiology of the disease[11] and the HLA locus should also be considered. A segregation analysis considering this locus showed that persons who were HLA identical to the diabetic proband were more likely to be diabetic than if they were HLA non-identical. Our data therefore suggest that the HLA locus is of greater importance to the inheritance of type I diabetes mellitus than the chromosome 11 short arm markers.

The paradox of disease association between the class 1 insert and type I diabetes in prevalence studies, but no close linkage between this locus and inheritance of diabetes mellitus in families, still requires explanation. A polygenic model of inheritance may account for these findings. If such a model is correct, it may be possible that in the pedigrees we have analysed the genetic contribution from loci other than the polymorphic insulin locus is of more importance. A class 1 insertion may be sufficient to cause diabetes in combination with other genetic and environmental factors, but it may not be a necessary factor in all cases.

The use of linked hypervariable loci provides a powerful tool for identifying and tracking the inheritance of certain segments of chromosomes in pedigrees. The present study suggests that the insulin gene polymorphic locus is not itself of major importance in the aetiology of type I diabetes. However, a similar approach may be adopted to study the inheritance of other disease associated loci.

We are grateful to the Wellcome Trust for a Pathology Training Fellowship to GAAF, The Fritz-Thyssen Foundation for a grant to DJG, Dr Alan Hall for his gift of the c-Ha-ras gene probe, and Dr Graham Bell for the Phins 310 probe. We are also indebted to Dr John Wyke and Val Fincham of the Imperial Cancer Research Fund for cloning facilities and technical assistance, and the Departments of Medical Illustration and Chemical Pathology, St Bartholomew's Hospital.

References

[1] Jeffreys AJ, Wilson V, Thein SL. Individual specific 'fingerprints' of human DNA. *Nature* 1985;**316**:76–7.

[2] 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.

[3] Bell GI, Selby KN, Rutter WJ. The highly polymorphic region near the insulin gene is composed of simple tandemly repeating sequences. *Nature* 1982;**295**:31–5.

[4] Bell GI, Horita S, Karam JH. A highly polymorphic locus near the human insulin gene is associated with insulin dependent diabetes. *Diabetes* 1984;**33**:176–83.

[5] Hitman GA, Tarn AC, Winter RM, *et al.* Type I (insulin dependent) diabetes and a highly variable locus close to the insulin gene on chromosome 11. *Diabetologia* 1985;**28**:218–22.

[6] Capon DJ, Chen EY, Levison AD, Seeberg Ph, Goeddel DV. Complete nucleotide sequences of the T24 bladder carcinoma oncogene and its normal homologue. *Nature* 1983;**302**:33–7.

[7] Fearon ER, Antonarakis SE, Meyers DA, Levine MA. c Ha ras 1 oncogene lies between beta globin and insulin loci on chromosome 11p. *Am J Hum Genet* 1984;**36**:329–37.

[8] Walker MD, Edmund T, Boulet AM, Rutter WJ. Cell specific expression controlled by the 5' flanking region of the insulin and chymotrypsin genes. *Nature* 1983;**306**:557–61.

[9] Kunkel LM, Smith KD, Boyer SH, *et al.* Analysis of human Y chromosome specific reiterated DNA in chromosomal variants. *Proc Natl Acad Sci USA* 1977;**74**:1245–9.

[10] Southern EM. Gel electrophoresis of restriction fragments. *Methods Enzymol* 1979;**68**:27–32.

[11] Thomson G. A two locus model of juvenile diabetes. *Ann Hum Genet* 1980;**43**:383–98.

Correspondence and requests for reprints to Dr G A A Ferns, Department of Diabetes and Lipid Research, St Bartholomew's Hospital, West Smithfield, London EC1A 7BE.