

## Type 1 (insulin-dependent) diabetes and a highly variable locus close to the insulin gene on chromosome 11

G. A. Hitman<sup>1</sup>, A. C. Tarn<sup>1</sup>, R. M. Winter<sup>2</sup>, V. Drummond<sup>1</sup>, L. G. Williams<sup>1</sup>, N. I. Jowett<sup>1</sup>, G. F. Bottazzo<sup>3</sup> and D. J. Galton<sup>1</sup>

<sup>1</sup>Departments of Diabetes, Lipids and Immunogenetics, St. Bartholomew's Hospital, London,

<sup>2</sup>Division of Inherited Metabolic Diseases, Northwick Park Hospital, Harrow and

<sup>3</sup>Department of Immunology, The Middlesex Hospital, London, UK

**Summary.** A polymorphic DNA sequence in the 5'-flanking region of the human insulin gene was studied in relation to Type 1 (insulin-dependent) diabetes. In 141 Caucasoid subjects analysed by Southern blot hybridisation techniques, two major DNA insertions were observed: a Class 1 allele or a Class 3 allele. The Class 2 allele was not observed in this group of subjects. Genotype frequencies in a control population ( $n = 88$ ) were: homozygous 1/1, 42%; heterozygous 1/3, 50%; and homozygous 3/3, 8%. Corresponding genotype frequencies in 53 Type 1 diabetic patients were 79%, 21% and 0%, respectively ( $p < 0.0005$  from  $\chi^2$  test). This confirms prev-

alence data reported by Bell et al. [16]. There appeared to be no coinheritance with HLA-DR3/DR4 related antigens, nor with autoimmune features. Analysis of 17 Type 1 diabetic pedigrees including 34 diabetic and 69 non-diabetic subjects did not demonstrate genetic linkage of these DNA inserts with diabetes, using an autosomal recessive, single locus model of inheritance.

**Key words:** DNA inserts, insulin gene, Type 1 diabetes, HLA antigens, prevalence, pedigree studies.

Although associations of Type 1 diabetes with HLA DR3 and DR4 antigens exist [1, 2], these are not alone sufficient to account for all of the genetic component of this disease. Investigators have therefore proposed a two locus model of inheritance for Type 1 diabetes [3–5]. Recently, Hodge et al. reviewed 27 possible genetic markers for Type 1 diabetes in 100 families [6]. They confirmed genetic linkage of diabetes with the HLA antigens, properdin factor B, and glyoxalase, all of which are coded for by genes on the short arm of chromosome 6. In addition, they investigated the possibility of non-chromosome 6 markers (on chromosomes 1, 2, 4, 9, 13 and 16) but failed to find any evidence for such associations, except weakly with the Kidd blood group (chromosome 2). However, when their data were pooled with those from the Minnesota families [7], linkage of this blood group with Type 1 diabetes was not confirmed.

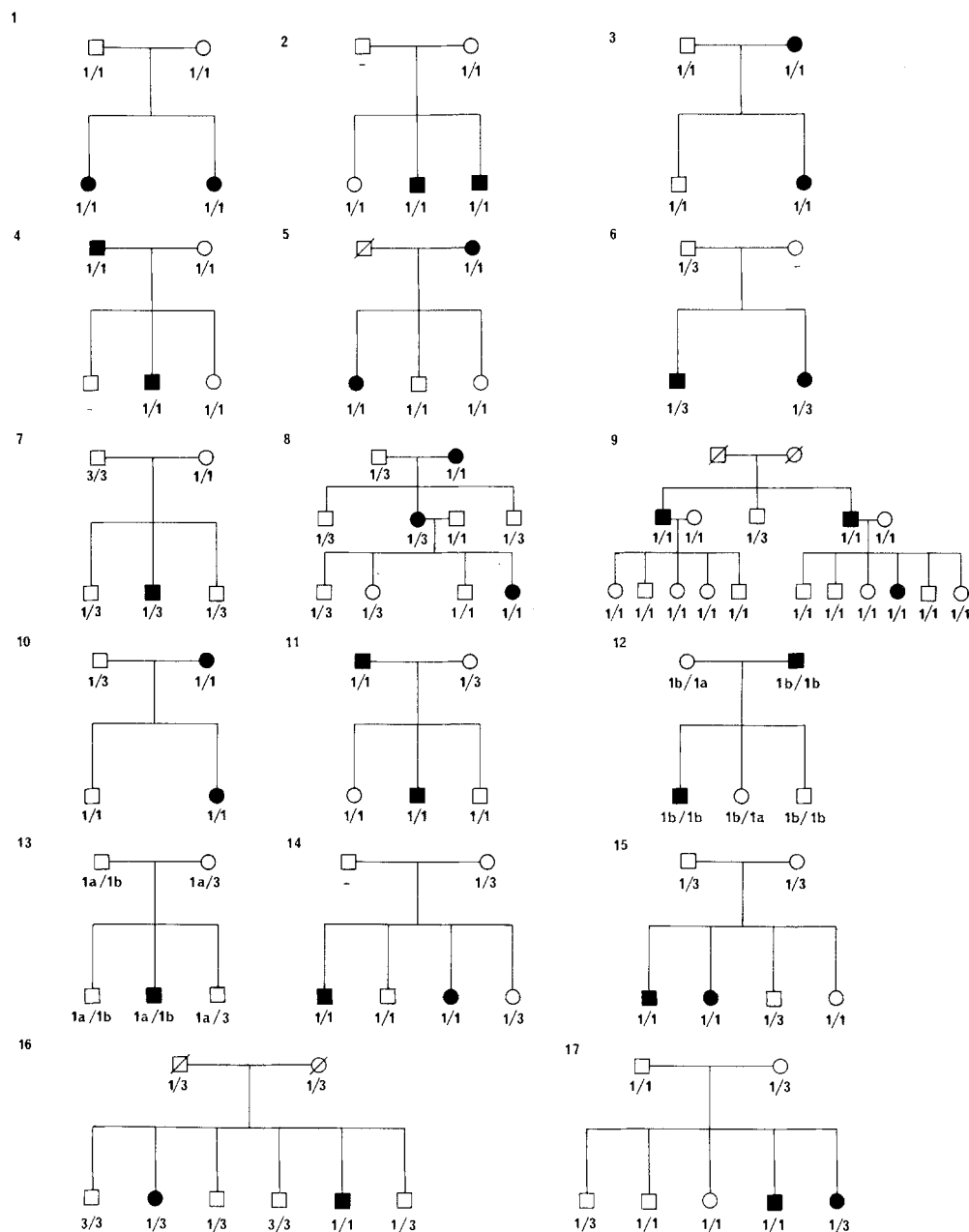
The human insulin gene locus is situated on the short arm of chromosome 11 [8] and near this locus (within 363 base pairs from the start of insulin gene transcription) a highly polymorphic region has been identified [9]. In Caucasoid populations this region can be divided into two main DNA-insertion classes (small, Class 1; or large, Class 3) distinguishable by the number of tandem repeats of a 14-base-pair nucleotide sequence [10]. As approximately 50% of individuals are heterozygous with respect to the length of the DNA insertion at this point, it is possible to use this as a genetic marker on chromosome 11 in proximity to the insulin gene. Associations of this region with the Class 3 DNA

insertion have been observed for Type 2 (non-insulin-dependent) diabetes [11–13], diabetic hypertriglyceridaemia [14] and atherosclerosis [15]. Recently, Bell et al. [16] have demonstrated an association of Type 1 diabetes with the Class 1 insertion. We have attempted to verify the latter association of Type 1 diabetes and adopted an alternative approach of looking for genetic linkage of this polymorphism adjacent to the insulin gene using pedigree analysis. In addition, we have investigated the possibility of certain HLA-DR types or autoimmune features segregating with the Class 1 allele.

### Subjects and methods

#### *Diabetic subjects*

Fifty-three unrelated Caucasoid Type 1 diabetic subjects were examined. Type 1 diabetes was defined by acute onset of symptoms (maximum several weeks) with weight loss, polyuria and polydipsia with an age of onset before 35 years, or an episode of diabetic ketoacidosis. Thirty-four were recruited from the Barts-Middlesex-Windsor family study [1] (the eldest sibling in each pedigree studied) and the rest were sequentially selected from the diabetic clinic at St. Bartholomew's Hospital. Clinical details of this group (mean  $\pm$  SD) age  $26.6 \pm 16.7$  years; age at onset of diabetes  $15.2 \pm 11.1$  years, range 9 months to 46 years; 36 were male and four subjects were above the accepted weight range for height (from the recommendations of the Fogarty Centre Conference on Obesity [17] and based on the original Metropolitan Life Insurance Tables 1959). All patients were on insulin treatment and seven had diabetic complications (including macroangiopathy, proliferative retinopathy, cataracts, neuropathy and necrobiosis lipoidica).



**Fig. 1.** Seventeen pedigrees analysed for linkage of Type 1 diabetes with the Class 1 allele. Numbers under the family symbol represent insulin genotypes (1 = Class 1 allele, 3 = Class 3 allele). Lod score analysis of all the families gave a value of  $-1.79$  at a recombination fraction of 0 using an autosomal recessive model of inheritance with 50% penetrance.  $\square$  = unaffected male;  $\circ$  = unaffected female;  $\emptyset$   $\emptyset$  = member deceased;  $\bullet$   $\blacksquare$  = member diabetic

### Control subjects

Eighty-eight Caucosoid subjects were studied as described previously [11]. None had a family history of diabetes and all had fasting blood glucose levels  $< 6$  mmol/l (mean  $\pm$  SD,  $4.2 \pm 0.67$  mmol/l). Clinical details were (mean  $\pm$  SD) age  $46.6 \pm 14.4$  years; 55 were male and 32 were above accepted weight range for height.

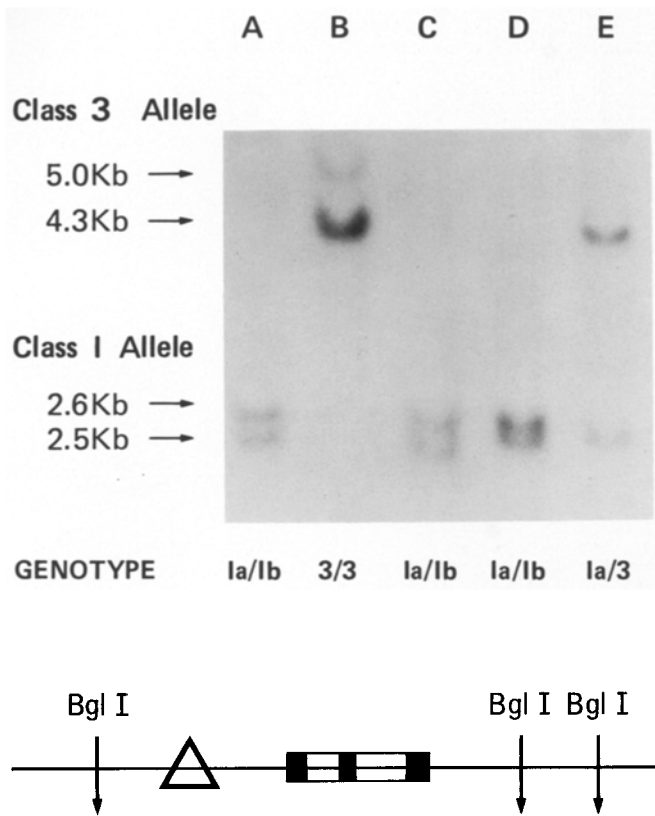
### Diabetic pedigrees

Seventeen families were serially selected from the Barts-Middlesex-Windsor family study with emphasis on pedigrees with at least two members with Type 1 diabetes ( $n = 15$ ). The family tree of each family is presented in Figure 1. In 34 diabetic subjects the mean ( $\pm$  SD) age was  $30 \pm 17$  years, 17 were male and mean ( $\pm$  SD) age at onset was  $17 \pm 15$  years. In 69 non-diabetic subjects, the mean ( $\pm$  SD) was  $34 \pm 16$  years and 41 were male. At least one Type 1 diabetic in each family has had circulating islet cell antibodies during their illness, and in twelve families there is an association of diabetes with a high-risk HLA genotype (i.e. HLA-DR3/HLA-DR3, HLA-DR3/HLA-DR4,

HLA-DR4/HLA-DR4; family Nos. 1, 3, 5, 7, 9, 10, 11, 12, 14, 15, 16, 17). In eight families there is a first degree history of Type 2 diabetes (family nos. 1, 3, 5, 8, 9, 10, 15, 17).

### DNA analysis

DNA was prepared by the method of Kunkel et al. [18] from thawed 10 ml blood samples anticoagulated with EDTA, which was then digested for 16 h (according to manufacturers specifications) with either restriction enzymes Sst I, Bgl I (Bethesda Research Laboratories, Cambridge, UK) or Rsa I (New England Bio Laboratories, Beverly, Mass., USA). The digested DNA was electrophoresed on a 1% agarose gel, transferred to nitrocellulose filters by the method of Southern [19] and hybridised to a  $^{32}$ P-labelled insulin gene probe (kindly donated by Dr. G. I. Bell, San Francisco, California, USA) containing the entire coding sequence and close flanking regions of the human insulin gene. Hybridisation bands were visualised by autoradiography and sized by comparing with Hind III digested  $\lambda$  phage (Bethesda Research Laboratories).



**Fig. 2.** Autoradiograph of insulin gene fragments (Bgl I digestion). Tracks A-E are DNA samples from the blood of five patients after digestion with restriction enzyme Bgl I (for 108 assays) studied by Southern-blot techniques [19] and hybridised to a <sup>32</sup>P-labelled DNA sequence containing the human insulin gene and closely flanking regions (-59 - 1600 base pairs). Sizes (in kilobases) of hybridisation bands are indicated to the left of the figure. Below the autoradiogram is a restriction enzyme map showing the cutting sites of Bgl I (indicated by arrows); the box represents the insulin gene and the triangle the polymorphic region. In other experiments the restriction enzyme Rsa I (for 158 assays) or Sst I (for 71 assays) was used

**Table 1.** Genotype distribution and allele frequencies of the polymorphic locus adjacent to the insulin gene in Type 1 diabetic patients and control subjects

	Insulin genotype			Allele frequency	
	1/1	1/3	3/3	Class 1	Class 3
Type 1 diabetic patients (n = 53)	42 (79%)	11 (21%)	0 (0%)	0.90	0.10
Control subjects (n = 88)	35 (42%)	44 (50%)	7 (8%)	0.67	0.33

A 3x2  $\chi^2$  test for differences in DNA insert frequencies between diabetic and non-diabetic subjects is 19.6 with 2 degrees of freedom ( $p < 0.0005$ ). **Terminology:** Following Bell et al. [16]; Class 1 DNA inserts = S (small) (11) = L (lower) (13) = - (12); Class 3 DNA inserts = L (large) (11) = U (upper) (13) = + (12)

**HLA typing**

HLA-DR typing was performed by the method of Van Rood et al. [20] in 44 of the Type 1 diabetic patients and in all family members using peripheral blood lymphocytes obtained from 20-ml citrated blood

samples defibrinated with thrombin (two drops, 50  $\mu$ /l) and calcium gluconate (0.6 ml, 10%). All ten officially recognized DR specifications were defined with a minimum of 60 antisera, using a two colour fluorescent technique.

**Statistical methods**

Allele frequencies were enumerated from the genotype frequencies. Initially the overall prevalence data were analysed by  $\chi^2$  using a 3x2 contingency table. Relative incidence of concordance of diabetes with pairs of insulin genotypes was calculated from the formula  $x = \frac{hK}{Hk}$  where h = number of Type 1 diabetic patients with marker  $\alpha$ ; H = number of controls with marker  $\alpha$ ; k = number of Type 1 diabetic patients with marker  $\beta$ ; K = number of control subjects with marker  $\beta$ . The 95% confidence limits of x (the relative risk for Type 1 diabetes with marker  $\alpha$ ) were calculated by the method of Woolf [21]. Pooled data from Bell et al. [16] and the present study were analysed similarly.

Preliminary analysis of the pedigrees was performed by comparing the genotype distribution found in the diabetic with that in the non-diabetic siblings. Linkage of DNA inserts with diabetes was analysed by lod scores using the computer program LIPED [22]. A lod score for a family is the logarithm for the ratio of the probability of that family given a recombination fraction < 0.5 to the probability of the same family given a recombinant fraction of 0.5 (i.e. no linkage). Autosomal dominant, intermediate and recessive models were analyzed with various values of penetrance with a linear age of onset correction, using identical parameters to those of Hodge et al. [6].

**Results**

Two main insulin gene-related fragments were observed depending on the presence of the Class 1 or Class 3 insertions. The sizes of the DNA fragments detected, using restriction enzymes Rsa I (cutting sites within the third coding region of the insulin gene and 5' to the polymorphic locus), Bgl I or Sst I, were for the Class 1 insertion 2.1  $\pm$  0.1 Kb, 2.7  $\pm$  0.25 Kb, 6.2  $\pm$  0.3 Kb, respectively, and for the Class 3 insertion 3.7  $\pm$  0.3 Kb, 4.5  $\pm$  0.4 Kb, 7.5  $\pm$  0.2 Kb, respectively. Two separate sized alleles could often be distinguished in patients homozygous for the Class 1 insertion (1a/1b Fig. 2). However, because of the inaccuracies of sizing for the prevalence data, these alleles were pooled for statistical analysis. Thus, the genotypes 1/1, 1/3 and 3/3 were distinguished as illustrated on a representative autoradiogram (Fig. 2). Results from the prevalence data are summarized in Table 1. In Type 1 diabetic patients a preponderance of subjects were homozygous for the Class 1 insertion, i.e. 79% compared with 42% in the control subjects (3x2 contingency table for differences in genotype distribution between Type 1 diabetic and control subjects  $\chi^2 = 19.6$ ; 2 degrees of freedom;  $p < 0.005$ ).

Comparison of genotypes (1/1 versus 1/3 + 3/3) in diabetic and control subjects revealed a relative incidence of 5.3 (95% confidence limits being 2.4-11.6) for concordance of Type 1 diabetes with genotype 1/1 ( $p < 0.005$  by the method of Woolf [21]). Distribution of the HLA-DR types within the unrelated Type 1 diabetic patient group was similar to that in other studies [1, 2] with 95% possessing HLA-DR3 or HLA-DR4 or both antigens. The percentage of diabetic patients with HLA-

**Table 2.** LOD score values of individual insulin-dependent pedigrees presented in Figure 1

Family	Models of inheritance <sup>a</sup>			
	Autosomal recessive		Intermediate LOD score <sup>b</sup>	Autosomal dominant 20% penetrance LOD score <sup>b</sup>
	50% penetrance LOD score <sup>b</sup>	20% penetrance LOD score <sup>b</sup>		
8	-0.34	-0.09	+0.21	+0.32
9	+0.38	+0.31	+0.31	+0.21
10	-0.05	-0.02	-0.01	0.00
11	-0.15	-0.05	-0.02	0.00
12	+0.02	+0.01	0.00	0.00
13	-0.01	0.00	0.00	0.00
14	-0.08	-0.03	-0.02	-0.03
15	+0.50	+0.48	+0.34	+0.28
16	-1.00	-0.45	-0.04	+0.03
17	-1.07	-0.66	-0.40	-0.29
Total LOD scores for all families <sup>c</sup>				
Maximum	+0.04 (0.3) <sup>d</sup>	+0.10 (0.2)	+0.4 (0.0)	+0.52 (0.0)
Minimum	-1.79 (0.0) <sup>d</sup>	-0.52 (0.0)	0.0 (0.5)	-0.0 (0.5)

Families 1-7 were uninformative with respect to linkage analysis.

All lod scores were computed using LIPED with age-of-onset correction. (a) Models of inheritance have been used as suggested by Hodge et al. [6]; (b) all lod scores are for a combination fraction of 0.0; (c) a lod score of  $< -2.0$  or  $> +3.0$  suggests no linkage or linkage respectively;  $-1.79$  means the odds against linkage are 60:1; (d) recombination fraction (0) is given in parentheses by corresponding total lod score

DR3 and DR4 individually was 73% and 68% respectively, with 45% possessing both antigens. There was independent segregation of HLA-DR types and insulin-related genotypes (1/1 versus 1/3) in the Type 1 diabetes group. The antigen frequencies of HLA-DR3 and HLA-DR4 in patients with the insulin-related genotype 1/1 were 0.46 and 0.29, respectively, in this group. Among the diabetics with the genotype 1/1 or 1/3, 44% and 40% respectively had autoimmune features (i. e. the presence of circulating islet cell antibodies for longer than 5 years or presence of organ-specific antibodies).

Distribution of the insulin-related genotypes in the 17 pedigrees are presented in Figure 2. Among diabetic siblings, 21, 6, and 0 respectively, were homozygous for the Class 1 allele, heterozygous and homozygous for the Class 3 allele and numbers of non-diabetic siblings were 28, 11 and 2. Lod score analysis (Table 2) was only significant using the autosomal recessive model of inheritance with 50% penetrance and recombination fraction 0.0, the total score for all informative families being  $-1.79$ , indicating the odds against linkage under this model to be 60:1. However, linkage under other models (in particular autosomal dominant) is not excluded.

## Discussion

We have found an association between patients genotyped homozygous 1/1 and Type 1 diabetes. This confirms the findings of Bell et al. [16] in California, who have also shown an association between Type 1 diabetes and this genotype in Caucasoid subjects, reporting that 76% of 113 Type 1 diabetic patients were homozygous for the Class 1 insertion. The allele frequencies of the Class 1 and Class 3 DNA inserts were: In Type 1 diabe-

tes 0.88 and 0.12, and in control subjects 0.67 and 0.33, the latter being similar as reported here. Both clinical studies from London and California are comparable, as shown by a non-significant heterogeneity  $\chi^2$  test (0.33, 2 degrees of freedom). Additionally, we have not found co-segregation of HLA-DR types or autoantibodies with the Class 1 allele. These data therefore confirm an additional association of Type 1 diabetes other than those with markers on the short arm of chromosome 6.

In the pedigree studies, a similar distribution of insulin-related genotypes was found in the non-diabetic and diabetic siblings. Formal statistical analysis by lod scores excludes close linkage between the insulin gene locus and diabetes under one single locus recessive model, but linkage under a dominant model cannot be excluded.

The question therefore arises why there is an association in the prevalence study, but no linkage between Type 1 diabetes and subjects homozygous for the Class 1 allele? Firstly, a possibility exists that the controls are unsuitable as they are older and more obese than the diabetic subjects. Looking at the subgroups of our control population matched for age (mean  $\pm$  SD  $30.3 \pm 5.6$  years;  $n = 32$ ), and for acceptable weight range for height ( $n = 56$ ), the respective genotype distributions (homozygous Class 1, heterozygous and homozygous for Class 3) were 38%, 50% and 12%; and 47%, 43% and 10%, which are still significantly different from those found in Type 1 diabetes. Secondly, the Class 1 allele is one of many different DNA insertions whose size varies from 0-600 base pairs in length. Unfortunately, using these methods, differences of less than 100 pairs within this insertion cannot be detected. There may be thus a specific small allele which asso-

ciates with Type 1 diabetes, but not with healthy control subjects. This possibility is currently being investigated by using a different genomic probe with another restriction endonuclease. Thirdly, the diabetic pedigrees are likely to be heterogeneous. Evidence for this is that in five families the high risk HLA genotypes are not found and in eight pedigrees there is a family history of Type 2 diabetes. Furthermore, the pedigrees study may be biased by the inclusion of so many families with two or more affected individuals. Lastly, Type 1 diabetes is likely to be a polygenic disease requiring the inheritance of genes at two or more loci, which in addition to environment factors and immunological predisposition leads to the development of diabetes. It is therefore not surprising that using a single insulin genomic probe, the linkage analysis on a small number of families is negative; linkage has already been proven with HLA-DR locus on chromosome 6, and thus penetrance of any insulin-linked gene must be low. For this reason, we analysed both our patient group and the non-diabetic siblings for coinheritance of HLA-DR antigens with that of the polymorphic locus adjacent to the insulin gene. Interestingly, a trend was observed for homozygous 1/1 to associate with HLA-DR3 antigens in the diabetic group (20 out of 42 compared with 8 out of 56 in the non-diabetic subjects), whereas HLA-DR4 antigen distributed equally between the two groups (13 out of 42 versus 14 out of 56, respectively). This result should be interpreted with caution as it is a comparison within related individuals.

In conclusion, we have demonstrated an additional association of Type 1 diabetes with patients who are homozygous for the Class 1 allele in the 5'-flanking region of the insulin gene on chromosome 11, but have so far failed to confirm close linkage between the insulin locus and susceptibility to Type 1 diabetes using single-locus models of inheritance.

*Acknowledgement.* We gratefully acknowledge financial support from the following institutions: British Diabetic Association, Joint Research Board of St. Bartholomew's Hospital, Medical Research Council, Fritz-Thyssen Foundation, Wellcome Trust Foundation, Juvenile Diabetes Federation and Novo Laboratories. We also acknowledge assistance from Dr. S. Humphries and Professor R. Williamson (Department of Biochemistry, St. Mary's Hospital, London) and J. London and A. Toups-Neumann (Düsseldorf) for secretarial help.

## References

1. Wolf E, Spencer KM, Cudworth AG (1984) The genetic susceptibility to Type 1 (insulin dependent) diabetes: analysis of the HLA-DR association. *Diabetologia* 24: 224-230
2. Svejgaard A, Platz P, Ryder LP (1980) Joint Report: Insulin dependent diabetes mellitus. In: Terasaki PI (ed) *Histocompatibility testing*. UCLA Tissue Typing Laboratory, Los Angeles, Cal, pp 638-656
3. Hodge SE, Anderson CE, Neiswanger K, Field LL, Spence MA, Sparkes RS, Sparkes MC, Crist M, Terasaki PI, Rimon DL, Roter JI (1981) Close genetic analysis between diabetes mellitus and Kidd blood group. *Lancet* 2: 893-895
4. Thomson G (1980) A two locus model for juvenile diabetes. *Ann Hum Genet* 43: 383-398
5. Clerget-Darpoux F, Bonaiti-Pellie C, Deschamps I, Hors J, Feingold N (1981) Juvenile insulin dependent diabetes: a possible susceptibility gene in interaction with HLA. *Ann Hum Genet* 45: 199-206
6. Hodge SE, Anderson CE, Neiswanger K, Sparkes RS, Rimon DL (1983) The search for heterogeneity in insulin-dependent diabetes (IDDM): linkage studies, two locus models and genetic heterogeneity. *Am J Hum Genet* 35: 1139-1155
7. Dunsforth TS, Rich SS, Swanson J, Barbosa J (1982) No evidence for linkage between diabetes and the Kidd marker. *Diabetes* 31: 990-993
8. Owerbach D, Bell GI, Rutter WJ, Brown JA, Shows TB (1981) The insulin gene is located on the short arm of chromosome 11 in humans. *Diabetes* 30: 267-270
9. Bell GI, Karam JH, Rutter WJ (1981) Polymorphic DNA region adjacent to the 5' end of the human insulin gene. *Proc Natl Acad Sci (USA)* 78: 5758-5766
10. Bell GI, Selby KN, Rutter WJ (1982) The highly polymorphic region near the human insulin gene is composed of simple tandemly repeating sequences. *Nature* 295: 31-35
11. Hitman GA, Jowett NI, Williams LG, Humphries S, Winter RM, Galton DJ (1984) Polymorphism in the 5' flanking region of the insulin gene and non-insulin dependent diabetes. *Clin Sci* 66: 383-388
12. Rotwein PS, Chirgwin J, Province M, Knowler WC, Pettitt DJ, Cordel B, Goodman HM, Permutt MA (1983) Polymorphism in the 5' flanking region of the human insulin gene: a genetic marker for non-insulin dependent diabetes. *N Engl J Med* 308: 65-71
13. Owerbach D, Nerup J (1982) Restriction fragment length polymorphism of the insulin gene in diabetes mellitus. *Diabetes* 31: 275-277
14. Jowett NI, Williams LG, Hitman GA, Galton DJ (1984) Diabetic hypertriglyceridaemia and related 5' flanking polymorphism of the human insulin gene. *Br Med J* 288: 96-99
15. Mandrup-Poulsen T, Owerbach D, Mortensen SA, Johansen K, Meinertz-Sørensen H, Nerup J (1984) DNA sequences flanking the insulin gene on chromosome 11 confer risk of atherosclerosis. *Lancet* 1: 253-255
16. Bell GI, Horita S, Karam JH (1984) A highly polymorphic locus near the human insulin gene is associated with insulin-dependent diabetes mellitus. *Diabetes* 33: 176-183
17. Bray GA (1973) Obesity in perspective. DHEW publication (National Institute of Health, Bethesda, USA). Fogerty International Centre for Advanced Study, pp 75-708
18. Kunkel LM, Smith KD, Boyer SH, Borgeanker DS, Wachtel SS, Miller OJ, Breg WR, Jones HW Jr, Rary JM (1977) Analysis of human Y chromosome specified reiterated DNA in chromosome variants. *Proc Natl Acad Sci (USA)* 74: 1245-1249
19. Southern EM (1975) Detection of specified sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 98: 503-517
20. Van Rood JJ, Van Leeuwen A, Ploem JS (1976) Simultaneous detection of two cell populations by two colour fluorescence and application on the recognition of B cell determinants. *Nature* 262: 795-797
21. Woolf B (1955) On estimating the relation between blood group and disease. *Ann Hum Genet (London)* 19: 251-253
22. Editorial (1979) Age of onset correction available for linkage analysis (Liped). *Am J Hum Genet* 31: 761-762

Received: 25 May 1984

and in revised form: 13 March 1985

Dr. G. A. Hitman  
The Medical Unit  
The London Hospital  
Whitechapel  
London E1  
UK