

Parental imprinting effect at the *INS-IGF2* diabetes susceptibility locus

C. Polychronakos, A. Kukuvtis, N. Giannoukakis, E. Colle

Montreal Children's Hospital Research Institute, Department of Pediatrics, Division of Endocrinology, McGill University, Montreal, Québec, Canada

Summary Although association of insulin-dependent diabetes mellitus with a haplotype at a locus encompassing the genes for insulin and the insulin-like growth factor II has been well established, two major studies disagree as to whether linkage to this locus is confined to paternally inherited alleles, or is present in alleles transmitted from either parental sex. Towards resolving this discrepancy, we examined parent-of-origin specific association rather than linkage, using the haplotype relative risk method in a mixed Caucasian population. We find that the haplotype relative risk (HRR) conferred by paternal chromosomes was much higher (5.1, $p < 0.01$) than the corre-

sponding maternal value (2.3, $p = 0.07$), which narrowly failed to reach statistical significance. Thus, although we cannot exclude an effect of the maternal allele, such an effect appears to be considerably weaker. We review evidence that parental imprinting is genotype-dependent, which may explain the different degrees to which the paternal effect is seen in different populations. [Diabetologia (1995) 38: 715–719]

Key words Genetic association, imprinting, insulin, insulin-like growth factor II.

Susceptibility to insulin-dependent diabetes mellitus (IDDM) appears to depend on more than one genetic locus [1]. In addition to the major role of the HLA complex, substantial contribution from a second locus, at or near the insulin gene (*INS*), is known to exist: in a French population, both linkage to and association with IDDM was demonstrated over a 4.1 kb area, from the *INS* 5' flanking region to within 1 kb of the first promoter of the insulin-like growth factor II gene *IGF2* [2–3]. Alleles in non-coding sequence polymorphisms in that region form a highly conserv-

ed haplotype common in the general population but even more common in diabetic patients, conferring a four-fold relative risk (RR). Linkage of this locus to IDDM was demonstrated in the original report [2] by finding that pairs of diabetic siblings shared one haplotype from heterozygous parents with a frequency of 0.67, significantly higher than the expected 0.5.

This transmission distortion was exclusively due to non-random sharing of haplotypes from heterozygous fathers (0.74 paternal, vs 0.54 maternal), suggesting that the susceptibility is transmitted by the paternal allele only. A more recent study in a British population questioned this parental effect, as pairs of diabetic siblings were found to share haplotypes at a frequency higher than 0.5, regardless of the sex of the heterozygous parent [4].

In an effort to resolve this controversy we re-examined the question using association rather than linkage data: we evaluated the IDDM association with the *INS-IGF2* locus by the haplotype relative risk method (HRR), an approach which eliminates the problem of population stratification in the study

Received: 12 July 1994 and in revised form: 16 December 1994

Corresponding author: Dr. C. Polychronakos, Montreal Children's Hospital Research Institute, Department of Pediatrics, McGill University, 2300 Tupper, Montréal, Québec, Canada, H3H 1P3

Abbreviations: IDDM, Insulin-dependent diabetes mellitus; RFLP, restriction fragment length polymorphism; *INS*, insulin gene; *IGF2*, gene for the insulin-like growth factor II; HRR, haplotype relative risk; OR, odds ratio; RR, relative risk.

of genetic association. More importantly, we describe a modification of the HRR method that permits examination of the effect of parental allele origin on disease-locus association to see whether the paternal effect on linkage, found in one of the studies, also applied to association.

Subjects, materials and methods

Subjects. Sixty-two IDDM diabetic subjects and their parents were recruited through the Diabetic Clinic of the Montréal Children's Hospital. The population was Caucasian and mostly of European descent, with French Canadians being the largest single group (approximately one-third). Diagnosis of IDDM was based on hyperglycaemia and ketonuria before age 18 years. Informed consent was obtained, and the study was approved by the hospital review board.

Polymorphism. The IDDM-associated region at 11 p 15.5 extends over 4.1 kb and consists of several polymorphisms at tight linkage disequilibrium. For the purpose of this study, we focused on a cluster of three RFLPs (Fig. 1) at the 3' end of *INS*, where association seems to be maximal [2, 3], and which are always concordant for the presence of IDDM-associated alleles [3, 4, and our own observations]. Our data were generated using the *PstI* site, but the other two sites gave identical information in our patients, as they did in previous studies [3, 4]. Approximately 55% of unselected Caucasian subjects are homozygous for the absence of the site, compared to 80–85% of diabetic patients [3, 4].

The RFLP was tested by polymerase chain reaction (PCR) as described by Bain et al. [4]. The cycling protocol was 94°C for 30 s, 60°C at 45 s, and 72°C at 90 s, in the presence of 2 mmol/l Mg⁺⁺. The PCR product was digested with *PstI* and resolved on an 8% polyacrylamide gel.

In the families where the diabetic child and both parents were heterozygous, assignment of parental origin of each allele was made based on haplotypes using two RFLPs at the *IGF2* gene (*ApaI* at the 3' untranslated region [5] and *AluI* on exon 3 [3]) which are within 30 kb from the *PstI* site but not at linkage disequilibrium with it. The short physical distance between the three markers, corresponding to a small fraction of one cM, allowed us to assume absence of recombination in the four families in which this method was successfully used. In three families no assignment of parental origin could be made. We were, therefore, able to examine 59 families.

Statistical analysis

The HRR method evaluates genetic association with no need for control subjects from the general population. This is accomplished by comparing the two parental alleles that are transmitted to the affected offspring from unaffected parents with the two that are not [6–8]. The latter can be thought of as equivalent to the genotype of an unselected control individual. The advantage of the HRR method is that it eliminates the possibility that the inevitable differences between recruiting procedures for patients and control subjects result in genetic differences between the two groups (population stratification).

As previously proposed [6, 7], the method compares the two transmitted and the two non-transmitted haplotypes together. It can be shown to be equally valid when haplotypes from the parent of each sex are separately examined. Thus, if

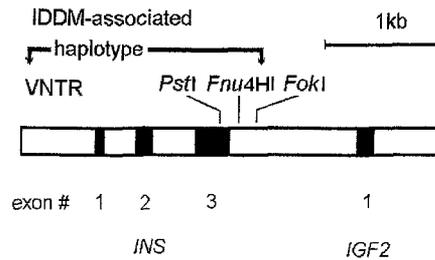


Fig. 1. The IDDM associated haplotype extends from the insulin 5' flanking VNTR to the *FokI* RFLP, less than 1 kb upstream of the first *IGF2* exon

$P(D|+)$ and $P(D|-)$ are the conditional probabilities of having the disease (D) when the paternally transmitted haplotype is the high-risk one (+) or its alternative(s) (-) then, regardless of the nature of the maternal allele, the relative risk conferred by the paternal allele is by definition:

$$RR = \frac{P(D|+)}{P(D|-)} \quad (1)$$

Although this formula requires prospective data, it can be transformed, as discussed by Fleiss [9], for application to case-control studies:

$$RR = \frac{P(+|D)}{1-P(+|D)} \times \frac{1-P(+)}{P(+)} \quad (2)$$

where $P(+|D)$ is the probability of having a paternally transmitted high-risk haplotype if affected with the disease, and $P(+)$ the frequency of that haplotype in the general population. If $P(D)$, the prevalence of the disease, is reasonably low, then $P(+)$ can be reliably estimated by the frequency of (+) among paternal haplotypes not transmitted to the affected offspring.

Thus, an estimate of the RR, based on N affected offspring can be obtained by the odds ratio (OR):

$$OR = \frac{T(+)}{T(-)} \times \frac{N-T(-)}{N-T(+)}$$

where $T(+)$ and $T(-)$ is the number of, respectively, high and low risk paternal haplotypes transmitted to the child. The identical reasoning can, of course, be applied to maternal alleles in separate calculations. These computations require no assumption about mode of inheritance or the presence of Hardy-Weinberg equilibrium.

Sample size. Preliminary data in our population, very similar to those found in case-control studies [3, 4], showed a frequency of the protective (-) allele of approximately 0.1 in transmitted, and 0.3 in non-transmitted chromosomes, conferring an HRR of approximately 4. In order to detect this difference in the parent-specific studies at the 0.05 level with a power of 95%, 111 data points are required (Table A3 in [9]). Our sample slightly exceeded this requirement ($59 \times 2 = 118$ paternal and 118 maternal alleles).

Results

The percentage of (+/+) homozygotes among our IDDM patients was very close to that reported in the French and British studies [2–4], as were our derived control genotype frequencies compared to the

Table 1. Transmission frequencies of all alleles, regardless of parental origin

	Alleles		Total	Genotypes	
	+	-		+/+	+/- or -/-
Transmitted	106	12	118	48 (81 %)	11 (19 %)
Non-transmitted	85	33	118	31 (52 %)	28 (48 %)
Total	191	45			
Odds ratio		3.3			
(95 % CL)		(1.6-6.8)			
χ^2		11			
p		< 0.005			

The frequencies of the non-transmitted alleles were used to calculate the corresponding genotype frequencies under Hardy-Weinberg equilibrium, to demonstrate their close concordance with case-control studies [2-4]. CL, confidence limits

Table 2. Relative risk conferred by the (+) PstI allele, by sex of the transmitting parent, using the haplotype relative risk method

	Paternal		Total	Maternal		Total
	+	-		+	-	
Transmitted	55	4	59	51	8	59
Non-transmitted	42	17	59	43	16	59
Total	97	21	118	94	24	118
Odds ratio		5.1			2.3	
(95 % CL)		(1.6-15.4)			(0.9-5.7)	
χ^2		8.3			2.6	
p		< 0.01			NS	

CL, Confidence limits

population control subjects in those studies (Table 1). The relative risk of 3.3 was also quite comparable.

Table 2 compares transmission of the two alleles by parental sex. Transmission of paternal (-) alleles to diabetic offspring was 0.19 (4/21), significantly different from the expected 0.5 ($p = 0.007$ by comparison to a binomial distribution). The corresponding maternal rate was 0.33, not significantly different from 0.5. The same information is derived by comparing the two parent-specific, Yates-corrected [9], ORs. The maternal value is much lower and not statistically significant. The paternal and maternal 2×2 tables were then subjected to the heterogeneity test of Woolf, as proposed by Gart [10]. The chi-square for the heterogeneity test was 1.25, which is not significant. This is not unexpected since, although the magnitude of the departure of transmission of maternal alleles from the expected value of 0.5 is not significant, the direction of the departure is in the same direction as that seen for paternal alleles. Therefore, this result does not allow us to rule out an effect of the maternal alleles, but we can conclude that the resulting maternal OR is less than the level of 4, targeted by our sample size.

Discussion

Our results provide independent confirmation of IDDM association with the *INS-IGF2* locus in a mixed North American Caucasian population. The

main problem with association studies in such a context is population stratification, i.e. the possibility that patients and control subjects do not share the same genetic background. Even matching for ethnicity, as was done in the British study [4], does not necessarily protect from such an artifact. Our study is the first to demonstrate this association by the HRR method which eliminates this drawback by using as controls parental alleles that had exactly the same probability of being transmitted to the affected child as those that were actually transmitted. Furthermore, our analysis of allele transmission by parental sex is consistent with the paternal effect found by Julier et al. [2]. The predominant role of paternal alleles may be related to the observation that IDDM is inherited more often from a diabetic father than from a diabetic mother [11], and suggests parental imprinting of the gene(s) involved.

Parental imprinting involves differential behaviour of the two copies of certain autosomal genes, depending on parental sex of origin [12]. We have shown that *IGF2* is one of the few human genes known to be imprinted, with exclusive expression from the paternal copy [13]. This would favour *IGF2* as the gene involved in the IDDM association. *INS*, the only other gene involved in the association, was found to have exclusive paternal expression in mouse yolk sack but not pancreas [14]. *INS* is likely also expressed from both alleles in human pancreas, as is seen at the protein level in carriers of mutant insulin peptides [15]. *IGF2* is expressed by lymphocytes

[16] and pancreatic beta cells [17] and may have important paracrine effects modulating the autoimmune process of IDDM.

On the other hand, in a sib-pair analysis in British IDDM subjects, Bain et al. [4] found significant distortion of haplotype sharing regardless of parental sex, thus questioning the imprinting effect. These authors speculated that genetic differences between the populations studied may account for the discrepancy. Indeed, experimental evidence from studies of parental imprinting of transgenes indicate that, in mice, the imprinting process is genotype-dependent [18, 19]. We have presented evidence of a similar variation in homo sapiens in the imprinting of *IGF2R*, the gene encoding the IGF-II receptor [20]. Differences between the French and British in the prevalence of allelic forms of genes controlling imprinting might explain the discrepancy.

Our paper presents a novel usage of the HRR method for detection and quantitation of parental imprinting effects in the genetics of human disease.

To address the same question, Julier et al. [2] and Bain et al. [4] examined parental effects on linkage by evaluating the frequency at which haplotypes from the parent of each sex are shared by diabetic siblings. Our method evaluates parental effects on association rather than on linkage, and we believe that it has the following advantages: first, affected sibling pairs are not required, an important advantage in disorders less common than IDDM. Second, it gives a meaningful quantitative result, estimating the relative risk of having the phenotype if one carries the associated allele in the chromosome derived from the parent of each sex. This may be important in estimating risk-benefit ratios in preventive interventions (immunosuppression, nicotinamide). Third, in calculating risk, it does not waste information, as all cases are taken into consideration, not only those with heterozygous parents.

The one a priori drawback of the method is that it can only be used in the presence of a known association: a specific allele of the marker locus must be at linkage disequilibrium with the disease-predisposing allele. In the association examined here, disease and marker loci are actually genetically indistinguishable [21]. At present the method is most useful in evaluating associations with precisely defined loci, such as those identified by the candidate gene approach. However, the rapidly improving resolution of human genome maps, currently an average of 0.7 cM [22], will make it increasingly likely to find linkage disequilibrium of anonymous markers with less well defined susceptibility loci, such as those identified by genome-wide linkage searches. At the same time, as imprinting is better understood and imprinted domains of the genome are defined, there will be a greater demand for formal tools, similar to the approach presented here, for detection and

quantitation of its effects in the genetics of human disease.

Acknowledgements. Supported by the Canadian Diabetes Association. A.K. is an Alan Ross Academic Fellow. Ms. J. Dufresne, R.N., was instrumental in recruiting participating families. We also thank Dr. K. Morgan for helpful discussions and Ms. R. Grabs, B.Sc., for technical assistance.

References

1. Risch N (1987) Assessing the role of HLA-linked and unlinked determinants of disease. *Am J Hum Genet* 40: 1–14
2. Julier C, Hyer RN, Davies J et al. (1991) Insulin-IGF2 region on chromosome 11 p encodes a gene implicated in HLA-DR4-dependent diabetes susceptibility. *Nature* 354 (14): 155–159
3. Lucassen AM, Julier C, Beressi JP et al. (1993) Susceptibility to IDDM maps to a 4.8 kb segment of DNA spanning the insulin gene and associated VNTR. *Nature Genetics* 4: 305–310
4. Bain SC, Prins JB, Hearne CM et al. (1992) Insulin gene region-encoded susceptibility to type 1 diabetes is not restricted to HLA-DR4-positive individuals. *Nature Genetics* 2: 212–215
5. Tadokoro K, Fujii H, Inoue T, Yamada M (1991) Polymerase chain reaction (PCR) for detection of Apal polymorphism at the insulin like growth factor II gene (*IGF2*). *Nucleic Acids Res* 19: 6967
6. Field LL, Fothergill-Payne C, Bertrams J, Baur MP (1986) HLA-DR effects in a large German IDDM dataset. *Genet Epidemiol* 1 [Suppl]: 323–328
7. Falk CT, Rubinstein P (1987) Haplotype relative risks: an easy, reliable way to construct a proper control sample for risk calculations. *Ann Hum Genet* 51: 227–233
8. Knapp M, Seuchter SA, Baur MP (1993) The haplotype-relative-risk (HRR) method for analysis of association in nuclear families. *Am J Hum Genet* 52: 1085–1093
9. Fleiss JL (1981) *Statistical methods for rates and proportions*. (2nd edn.) John Wiley, New York
10. Gart JJ (1962) On the combination of relative risks. *Biometrics* 18: 601–610
11. Warram JH, Krolewski AS, Gottlieb MS, Kahn CR (1984) Differences in risk of insulin-dependent diabetes in offspring of diabetic mothers and diabetic fathers. *N Engl J Med* 1: 149–152
12. Hall JG (1990) Genomic imprinting review and relevance to human diseases. *Am J Hum Genet* 46: 857–873
13. Giannoukakis N, Deal C, Goodyer CG, Paquette J, Polychronakos C (1993) Parental genomic imprinting of the human *IGF2* gene. *Nature Genetics* 4: 98–101
14. Giddings SJ, King CD, Harman KW, Flood JF, Carnaghi LR (1994) Allele specific inactivation of insulin 1 and 2, in the mouse yolk sac, indicates imprinting. *Nature Genetics* 6 (3): 310–313
15. Gabbay KH, DeLuca K, Fisher JN Jr, Mako ME, Rubenstein AH (1976) Familial hyperproinsulinemia. An autosomal dominant defect. *N Engl J Med* 294: 911–915
16. Nyman T, Pekkonen F (1993) The expression of IGFs and their binding proteins in normal human lymphocytes. *Acta Endocrinol* 128: 168–172
17. Hill DJ, Hogg J (1992) Expression of insulin-like growth factors and their binding proteins during pancreatic development in rat, and modulation of IGF actions on rat islet DNA synthesis by IGF BPs. *Adv Exp Med Biol* 321: 113–120

18. Allen ND, Norris ML, Surani MA (1990) Epigenetic control of transgene expression and imprinting by genotype-specific modifiers. *Cell* 61: 853–861
19. Sapienza C, Paquette J, Tran TH, Peterson A (1989) Epigenetic and genetic factors affect transgene methylation imprinting. *Development* 107: 165–186
20. Xu Y, Goodyer CG, Deal C, Polychronakos C (1993) Functional polymorphism in the parental imprinting of the human IGF2R gene. *Biochem Biophys Res Commun* 197: 747–754
21. Hodge E (1993) Linkage analysis versus association analysis. Distinguishing between two models that explain disease-marker associations. *Am J Hum Gen* 53: 367–384
22. Cooperative Human Linkage Center, Genethon, University of Utah, Yale University and CEPH (1994) A comprehensive human linkage map with centimorgan density. *Science* 265: 2049–2070