J Hum Genet (2000) 45:184–187

SHORT COMMUNICATION

Robert A. Hegele · Anthony J.G. Hanley
Bernard Zinman · Stewart B. Harris · Carol M. Anderson

Disparity between association and linkage analysis for *HNF1A* G319S in type 2 diabetes in Canadian Oji-Cree

Received: October 29, 1999 / Accepted: January 5, 2000

Abstract In parallel experiments designed to find the genetic determinants of type 2 diabetes in Oji-Cree, we identified several linked chromosomal regions, using genomic scanning, in addition to a private diabetes-associated mutation, namely *HNF1A* G319S, using candidate gene sequencing. The genome scan did not identify the region harboring *HNF1A* as being linked with diabetes. Also, the *HNF1A* mutation, when used directly in sib-pair linkage analysis, was not linked with diabetes. However, *HNF1A* G319S was very strongly associated with diabetes, predicted the clinical severity of diabetes, and performed well as a diagnostic predictive test for diabetes in the Oji-Cree. Despite the failure of linkage analysis to identify *HNF1A* as a determinant of type 2 diabetes, we feel justified in interpreting that G319S has a very important pathogenic role in Oji-Cree diabetes, based upon the highly suggestive association studies. The probable etiologic heterogeneity of type 2 diabetes in the Oji-Cree created a situation in which association analysis was much more sensitive to detect a relationship between *HNF1A* S319 and diabetes than was linkage analysis. The effectiveness of linkage analysis will probably be limited in study samples that have an even greater complexity of genetic background and/or disease etiology. Thus, the absence of linkage does not always mean that a genomic variant is not an important determinant of a complex disease. Furthermore, our experience confirms the value of using several complementary strategies to identify susceptibility genes for a complex disease.

R.A. Hegele (✉) · C.M. Anderson
Blackburn Cardiovascular Genetics Laboratory, John P. Roberts Research Institute, 406-100 Perth Drive, London, Ontario, N6A 5K8 Canada
Tel. +1 519-663-3461; Fax +1 519-663-3789
e-mail: robert.hegele@rri.on.ca

A.J.G. Hanley · B. Zinman
Samuel Lunenfeld Research Institute and Department of Medicine, University of Toronto, Toronto, Ontario, Canada

S.B. Harris
Thames Valley Family Practice Research Centre, University of Western Ontario, London, Ontario, Canada

Key words Complex disease · Polygenic disease · Gene environment

Introduction

Approximately 16,000 Oji-Cree inhabit an area the size of France in the far north-west of Ontario. The prevalence of type 2 diabetes and impaired glucose tolerance among Oji-Cree adults living on the isolated Sandy Lake reserve is about 40%, which is the third highest of any group in the world (Harris et al. 1997). As recently as 50 years ago, type 2 diabetes was virtually unknown in these people (Harris et al. 1997). The ancestors of the Sandy Lake Oji-Cree lived a nomadic, hunting-gathering existence. Since the development of the reservation and residential school systems, the lifestyle of these people has changed drastically, from physically active to very sedentary. The traditional diet of low fat, low nutrient-dense foods has been replaced by high-energy processed foods. The high prevalence of type 2 diabetes suggests that these people are genetically predisposed to the disease, whose expression has been unmasked by recent lifestyle changes (Harris et al. 1997). We have used complementary analytic approaches to define susceptibility genes for type 2 diabetes in the Oji-Cree.

One approach has involved DNA sequencing of candidate genes for type 2 diabetes, followed by analysis of the association of new DNA variants with diabetes. In one experiment, we sequenced the *MODY3* gene, also called *HNF1A*, in three adult Oji-Cree subjects with typical type 2 diabetes (Hegele et al. 1999b). We found a novel private amino acid variant, G319S, in two of the three diabetic subjects. The diabetes phenotype in the Oji-Cree resembles typical type 2 diabetes mellitus (DM) that is associated with obesity and hyperinsulinemia. It is definitely distinct from maturity-onset-diabetes-of-the-young (MODY), for which *HNF1A* is the causative gene (Velho and Froguel 1998). The G319S mutation resides within the proline II rich domain of the transactivation site of HNF-1α and affects a glycine residue that has been conserved throughout evolu-

tion. We found that the overall *HNF1A* S319 allele frequency in the Oji-Cree was 0.119 (107/902), which was significantly different ($P < 10^{-13}$) from its complete absence from more than 1,000 alleles of subjects from six other ethnic groups (Hegele et al. 1999b).

We found that the S319 allele frequency in the diabetic Oji-Cree was significantly higher than that in the non-diabetic Oji-Cree (0.209 [49/234] vs 0.087 [58/668]; $P < 10^{-6}$). Compared with G319/G319 subjects, the relative risk for diabetes was 4.0 (95% confidence interval [CI], 2.7 to 6.0) and 2.0 (95% CI, 1.4 to 2.7) for S319/S319 and S319/G319 subjects, respectively. Also, compared with G319/G319 subjects, S319/S319 and S319/G319 subjects were affected with diabetes at a significantly lower age and smaller body mass. Thus, candidate gene sequencing and association analysis helped to identify a very promising DNA variant, which can now be further evaluated to understand its role in the susceptibility for type 2 diabetes in the Oji-Cree (Hegele et al. 1999b).

In a parallel series of experiments, we performed genome wide screening in a subgroup of Sandy Lake Oji-Cree sib pairs affected with type 2 diabetes (Hegele et al. 1999a). We were surprised that no marker on chromosome 12, particularly 12q24, which harbors *HNF1A,* showed even suggestive linkage with type 2 diabetes. We first surmised that the inability to detect linkage with *HNF1A* was due to an insufficient number of sib pairs (Hegele et al. 1999a). We subsequently realized that we had in hand a marker from the above association analysis; namely, *HNF1A* G319S, that was likely to be a susceptibility allele for diabetes. We reasoned that entering genotypes of *HNF1A* G319S into an expanded affected sib-pair linkage analysis might be helpful to understand why there was failure to detect linkage with chromosome 12 markers, as S319 appeared to be a bona fide susceptibility allele for type 2 diabetes in the Oji-Cree.

Materials and methods

Study subjects

The community of Sandy Lake, Ontario is located about 2,000 km northwest of Toronto, in the subarctic boreal forest of central Canada. Seven hundred and twenty-eight members (72% of the total population) of this community aged 10 years and above participated in the Sandy Lake Health and Diabetes Project (Harris et al. 1997). Several complete clinical descriptions of the entire study sample have already been published (Harris et al. 1997; Hegele 1998a,b,c). The project was approved by The University of Toronto Ethics Review Committee and the Sandy Lake First Nations Band Council.

Biochemical analyses and diagnostic criteria

Plasma samples were obtained with informed consent. Exclusion criteria were inadequate blood samples available for all biochemical and/or genetic determinations. Subjects gave plasma samples after fasting overnight for 12 h. Blood was centrifuged at 2,000 rpm for 30 min and the plasma was stored at −70°C. Concentrations of fasting plasma glucose and insulin were determined as described (Harris et al. 1997). A standard 75 g oral glucose tolerance test (OGTT) was then administered and a second blood sample was collected after 2 h for plasma glucose determination. Subjects were excluded from the OGTT if they had physician-diagnosed diabetes and/or if they were currently receiving treatment with insulin and/or oral hypoglycemic agents, or if they had a fasting blood glucose level exceeding 11.1 mmol/l. Subjects who were pregnant at the time of recruitment had their OGTT deferred until 3 months postpartum. Type 2 diabetes, impaired glucose tolerance, and normal glucose tolerance (non-diabetic) were diagnosed using pre-1997 criteria (Harris et al. 1997).

Genetic analyses

We had previously determined genotypes for *HNF1A* G319S (Hegele et al. 1999b). From the total sample of 728 Oji-Cree subjects, we used 92 affected individuals to create 97 sib pairs. Of the 92 affected subjects, 65 had frank type 2 diabetes and 27 had impaired glucose tolerance. The affected sibships, with their *HNF1A* G319S genotypes, are shown in Fig. 1. The observed frequency of the *HNF1A* S319 allele in the 97 affected sib pairs was 0.20, which was the same as that observed in subjects with type 2 diabetes from the entire Sandy Lake sample (Hegele et al. 1999b);

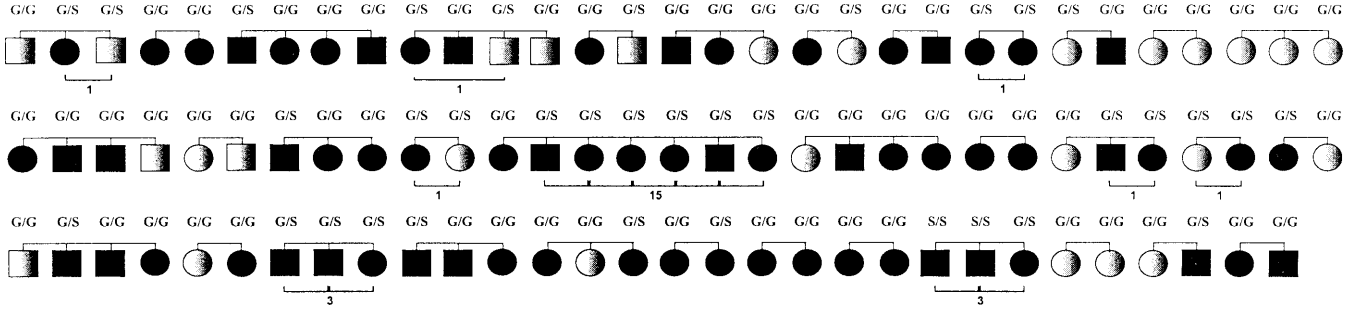

Fig. 1. *HNF1A* G319S genotypes in Oji-Cree affected sib pairs. The *solid symbols* represent subjects with type 2 diabetes and the *shaded symbols* represent subjects with impaired glucose tolerance. Genotypes are shown *above the sibships*. The affected sib pairs, and the number of sib pairs that are concordant for having the *HNF1A* S319 allele are indicated by *brackets and numerals below the sibships*

35.9% of subjects (33/92) in the sib-pair analysis had either the G319/S319 or S319/S319 genotype. The percentages of affected sib pairs concordant for having *HNF1A* S319, concordant for not having *HNF1A* S319, and discordant for the *HNF1A* genotype were 27.8% (27/97), 39.2% (38/97), and 33.0% (32/97), respectively. Analysis of the linkage of *HNF1A* S319 to diabetes in the affected sibships was performed using SIBPAL (version 2.8) from SAGE (version 2.2, Department of Epidemiology and Biostatistics, Case Western Reserve University).

Results

As linkage calculations depend upon the allele frequencies entered into the analysis, three analyses, using different *HNF1A* S319 allele frequencies, were performed. Specifically, the *HNF1A* S319 allele frequencies from the non-diabetic Sandy Lake subjects (0.09), the 97 sib pairs selected for the present analysis (0.20), and the estimate used previously (Hegele et al. 1999a) for the sib-pair analysis (0.30), were entered as the allele frequencies for the locus parameter for each analysis. For each analysis, the z-score was <1.0 and each P value was not significant (Table 1).

Discussion

These results suggest that a marker such as *HNF1A* S319, which is common in the population and is the functional basis for diabetes susceptibility, can be associated with, but not linked with, the disease phenotype. Compared with association analysis, linkage analysis with common, identical-by-state alleles is thus markedly less sensitive, even when the marker is identical with the probable mechanistic basis for the diabetes susceptibility. This is consistent with results derived from simulated data (Risch and Merikangas 1996). The apparent confounding factor in the Oji-Cree is the etiologic and/or genetic heterogeneity for type 2 diabetes. Such heterogeneity is indicated by the fact that only approximately 40% of Oji-Cree with type 2 diabetes were either homozygous or heterozygous for *HNF1A* S319. This created many classes of affected sib pairs according to *HNF1A* genotype, as indicated in Fig. 1. It was only those sib pairs concordant for having S319 that contributed to the non-parametric linkage score. This type of sib pair represented only a minority (27/97) of the total. The linkage was probably diluted by the other types of sib pairs; namely, those discordant for *HNF1A* S319 or concordant for not having *HNF1A* S319.

Table 1. Summary of linkage analysis using *HNF1A* G319S genotype in affected Oji-Cree sib pairs ($n = 97$) with type 2 diabetes

S319 frequency	z-score	P value
0.30 (Estimate[a])	0.848	0.20
0.20 (Present sib-pairs [$n = 97$])	0.706	0.28
0.09 (Non-diabetic Sandy Lake subjects)	0.428	0.33

[a] Estimate for sib-pair analysis used by Hegele et al. (1999a)

Before we embarked on these studies, we had assumed that finding linkage between diabetes and a common genetic variant would be relatively straightforward in the Sandy Lake Oji-Cree, who were descended from as few as 20 founders (Hegele et al. 1998a). Furthermore, the geographic isolation of the reserve has ensured minimal European admixture, as confirmed by admixture estimates ranging from 1.6% to 9.4%, based upon the observed frequencies of *HFE* Y282 and *F5* Q506 (Hegele et al. 1998b; 1998c). Because of the relative homogeneity of both the genetic background and the environmental attributes in the Sandy Lake Oji-Cree, we had assumed there would be a strong likelihood of finding a major genetic susceptibility locus for diabetes, even if there was heterogeneity of susceptibility at the genetic level. However, comparison of the analyses shown above indicates that the use of *HNF1A* S319 directly in the linkage analysis would not have resulted in its identification as a significantly linked genetic variant, using the maximal number of affected sib pairs that could be obtained from this study sample. This is not to say that linkage analysis would not be a reasonable approach in another situation. Rather, complementary strategies will be necessary to identify susceptibility genes for complex diseases.

We are presently evaluating the in-vitro function of *HNF1A* S319 in order to better understand the mechanism of the association. We are also planning intervention studies in young asymptomatic *HNF1A* S319 carriers to determine whether progression of diabetes can be affected in young people with this genotype. It was the association analysis, not the linkage analysis, that brought us to this point in our studies and has allowed us to pursue new lines of experimentation. Despite the apparently favorable attributes of the Oji-Cree sample, the probable etiologic heterogeneity of type 2 diabetes created a situation in which association analysis was much more sensitive to detect a relationship between *HNF1A* S319 and diabetes than was linkage analysis. This suggests that the effectiveness of linkage analysis will be limited in study samples that have even greater complexity of genetic background and/or etiology of disease. Our experience also confirms the value of using several complementary strategies to identify susceptibility genes for complex diseases. Finally, this real-life example indicates the constraints encountered when actual human samples are used in genetic analyses. In the entire Oji-Cree sample of more than 700 subjects, there were fewer than 15 units suitable for traditional transmission disequilibrium testing (TDT). Thus, while certain analytic strategies, such as TDT, might hold out great theoretical appeal, our experience would indicate that the final choice of an analytical strategy to identify causative genes represents a compromise that makes effective use of the actual human material obtained.

Acknowledgments We would like to acknowledge: the chief and council of the community of Sandy Lake; the Sandy Lake community surveyors; the Sandy Lake nurses and the staff of the University of Toronto Sioux Lookout program.

Linkage analysis was run using the SIBPAL version 2.8 subroutine

of the SAGE version 2.2 statistical package for genetic epidemiology (Department of Epidemiology and Biostatistics, Case Western Reserve University).

This work was supported by Grants from the National Institutes of Health (1-R21-DK44597–01), the Ontario Ministry of Health (no. 04307), the Canadian Diabetes Association, the Medical Research Council of Canada (no. MA13430), and general support from the Blackburn Group. Dr. Hegele is a Career Investigator of the Heart and Stroke Foundation of Ontario (no. CI-2979). Dr. Harris is a Career Investigator of the Ontario Ministry of Health. Mr. Hanley was supported by Health Canada through a National Health Research and Development Program Research Training Award.

References

Harris SB, Gittelsohn J, Hanley AJG, Barnie A, Wolever TMS, Gao J, Logan A, Zinman B (1997) The prevalence of NIDDM and associated risk factors in native Canadians. *Diabetes Care* 20:185–197

Hegele RA, Harris SB, Hanley AJ, Azouz H, Connelly PW, Zinman B (1998a) Absence of association between genetic variation of the beta 3-adrenergic receptor and metabolic phenotypes in Oji-Cree. *Diabetes Care* 21:851–854

Hegele RA, Harris SB, Zinman B (1998b) Hemochromatosis and diabetes mellitus. *Ann Intern Med* 129:587

Hegele RA, Harris SB, Cao H, Hanley AJ, Zinman B (1998c) Factor V Leiden (F5 Q506) and vascular disease in Canadian Oji-Cree. *Diabetes Care* 21:1203

Hegele RA, Sun F, Harris SB, Anderson C, Hanley AJG, Zinman B (1999a) Genome-wide scanning for type 2 diabetes susceptibility in Canadian Oji-Cree, using 190 microsatellite markers. *J Hum Genet* 44:10–14

Hegele RA, Cao H, Harris SB, Hanley AJG, Zinman B (1999b) The hepatic nuclear factor with early-onset type 2 diabetes in Canadian Oji-Cree. *J Clin Endocrinol Metab* 84:1077–1082

Risch N, Merikangas K (1996) The future of genetic studies of complex human diseases. *Science* 273:1516–1517

Velho G, Froguel P (1998) Genetic, metabolic and clinical characteristics of maturity onset diabetes of the young. *Eur J Endocrinol* 138:233–239