

A large sample of finnish diabetic sib-pairs reveals no evidence for a non-insulin-dependent diabetes mellitus susceptibility locus at 2qter.

S Ghosh, ... , E Tuomilehto-Wolf, M Boehnke

J Clin Invest. 1998;102(4):704-709. <https://doi.org/10.1172/JCI2512>.

Research Article

In the first reported positive result from a genome scan for non-insulin-dependent diabetes mellitus (NIDDM), Hanis et al. found significant evidence of linkage for NIDDM on chromosome 2q37 and named the putative disease locus NIDDM1 (Hanis et al. 1996. *Nat. Genet.* 13:161-166). Their total sample was comprised of 440 Mexican-American affected sib-pairs from 246 sibships. The strongest evidence for linkage was at marker D2S125 and best estimates of lambdas (risk to siblings of probands/population prevalence) using this marker were 1.37 under an additive model and 1.36 under a multiplicative model. We examined this chromosomal region using linkage analysis in a Finnish sample comprised of 709 affected sib-pairs from 472 sibships. We excluded this region in our sample (multipoint logarithm of odds score ≤ -2) for lambdas ≥ 1.37 . We discuss possible reasons why linkage to 2q37 was not found and conclude that this region is unlikely to be playing a major role in NIDDM susceptibility in the Finnish Caucasian population.

Find the latest version:

<https://jci.me/2512/pdf>



A Large Sample of Finnish Diabetic Sib-Pairs Reveals No Evidence for a Non-Insulin-dependent Diabetes Mellitus Susceptibility Locus at 2qter

Soumitra Ghosh,* Elizabeth R. Hauser,† Victoria L. Magnuson,* Timo Valle,§ Delphine S. Ally,* Zarir E. Karanjawala,* Joseph B. Rayman,* Julie I. Knapp,* Anjene Musick,* Joyce Tannenbaum,* Catherine Te,* William Eldridge,* Shane Shapiro,* Tiffany Musick,* Colin Martin,* Alistair So,* Alyson Witt,* Jillian Blaschak Harvan,‡ Richard M. Watanabe,‡ William Hagopian,‡ Johan Eriksson,§ Stella J. Nylund,§ Kimmo Kohtamaki,§ Eva Tuomilehto-Wolf,§ Liisa Toivanen,§ Gabriele Vidgren,§ Christian Ehnholm,§ Richard N. Bergman,¶ Jaakko Tuomilehto,§ Francis S. Collins,* and Michael Boehnke‡

*Genetics and Molecular Biology Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892; †Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, Michigan 48109-2029; ‡Diabetes and Genetic Epidemiology Unit, Department of Epidemiology and Health Promotion, National Public Health Institute, Helsinki, FIN-00300 Finland; †Department of Medicine, University of Washington, Seattle, Washington 98195; and §Department of Physiology and Biophysics, School of Medicine, University of Southern California, Los Angeles, California 90033

Abstract

In the first reported positive result from a genome scan for non-insulin-dependent diabetes mellitus (NIDDM), Hanis et al. found significant evidence of linkage for NIDDM on chromosome 2q37 and named the putative disease locus NIDDM1 (Hanis et al. 1996. *Nat. Genet.* 13:161–166). Their total sample was comprised of 440 Mexican-American affected sib-pairs from 246 sibships. The strongest evidence for linkage was at marker D2S125 and best estimates of λ_s (risk to siblings of probands/population prevalence) using this marker were 1.37 under an additive model and 1.36 under a multiplicative model. We examined this chromosomal region using linkage analysis in a Finnish sample comprised of 709 affected sib-pairs from 472 sibships. We excluded this region in our sample (multipoint logarithm of odds score ≤ -2) for $\lambda_s \geq 1.37$. We discuss possible reasons why linkage to 2q37 was not found and conclude that this region is unlikely to be playing a major role in NIDDM susceptibility in the Finnish Caucasian population. (*J. Clin. Invest.* 1998. 102:704–709.) Key words: gene mapping • affected sib-pairs • non-insulin-dependent diabetes mellitus • fluorescent genotyping • linkage analysis

Introduction

About 90% of diabetics have non-insulin-dependent diabetes mellitus (NIDDM),¹ a heterogeneous disorder of unknown etiology, which affects ~7% of the adult US population and 4% of the adult Finnish population who are between the ages of 45 and 64 yr (1–3). Increased insulin resistance, impaired β -cell function, and increased hepatic glucose production are the metabolic hallmarks of NIDDM. Epidemiological studies sug-

gest a strong genetic basis for the disease, although there is also well-documented evidence for an environmental component in NIDDM (1, 4). Understanding the genetic basis of NIDDM will help elucidate the molecular mechanisms for disease and lead to a better appreciation of environmental triggers. In turn, this will help scientists and clinicians devise new modes of therapy for disease prevention and cure.

At least two genome scans have been recently completed for NIDDM. Hanis et al. (5) reported the results of a total genome scan using 490 markers at an average resolution of 8.6 cM on a combined data set of up to 440 affected sib-pairs (ASPs) in 246 Mexican-American families from Starr County, Texas. One marker, D2S125 on chromosome 2q37, showed evidence for genome-wide significance for linkage to NIDDM (maximum logarithm of odds [LOD] score or MLS in the combined data set of 4.1, equivalent to $P = 2 \times 10^{-5}$). Similar results were obtained when a model-based approach using MOD (maximizing LOD score over penetrance and allele frequency parameters) scores was used. No obvious candidate genes are known to be mapped to the interval (5). More recently, selected NIDDM pedigrees from Botnia, Western Finland, with the lowest quartile for the mean 30-min insulin levels after oral glucose tolerance tests (OGTT) showed significant evidence for linkage by genome scan to chromosome 12q near the MODY3 locus (6).

There have been so far very few studies that focus on the genetic basis of NIDDM in Finland. Furthermore, the ages of the samples may differ from study to study, making it difficult to compare frequencies of disease. One study that identified twins of the same sex from record linkage reported a risk rate of 16.2% (CI = 12–20%) to dizygotic twins of a proband with NIDDM in individuals above the age of 28 (7). If we assume that the prevalence of NIDDM is ~5% (for the 45–64-yr age group) (2), then the recurrence risk for NIDDM in this population is at least 3.2 (16.2/5). The same study also reported a heritability of 79% for NIDDM in the absence of dominance (the best-fitting model) (7). Taking the above together, Fin-

Address correspondence to Soumitra Ghosh, Room 1W108, Building 9, Genetics and Molecular Biology Branch, National Human Genome Research Institute, National Institutes of Health, 9000 Rockville Pike, Bethesda, MD 20892. Phone: 301-496-7826; FAX: 301-480-9667; E-mail: sghosh@alw.nih.gov Elizabeth R. Hauser's present address is Section of Medical Genetics, Department of Medicine, Duke University, Durham, NC 27710.

Received for publication 15 December 1997 and accepted in revised form 2 June 1998.

The Journal of Clinical Investigation
Volume 102, Number 4, August 1998, 704–709
<http://www.jci.org>

1. *Abbreviations used in this paper:* ASP, affected sibling pairs; FUSION, Finland-United States Investigation of NIDDM Genetics; GAD, glutamic acid decarboxylase; IBD, identical by descent; IBS, identical by state; IDDM, insulin-dependent diabetes mellitus; LOD, logarithm of odds; MLS, maximum LOD score; MOD, maximum LOD score over penetrance and allele frequency parameters; NIDDM, non-insulin-dependent diabetes mellitus; OGTT, oral glucose tolerance test; WHO, World Health Organization.

land appears to be a reasonable choice for performing genetic studies into NIDDM.

We are engaged in a genome-scan approach to identifying NIDDM loci in a cohort of 709 ASPs in 472 Finnish families. Because of the potential importance of the chromosome 2q result, there is a critical need to replicate these findings (8). Therefore we have typed our families with 30 chromosome 2 markers. Our results argue against the presence of a gene near the 2q telomere of substantial effect on NIDDM risk in the Finnish population.

Methods

Family collection and research design. The FUSION (Finland–United States Investigation of NIDDM Genetics) study is an international collaborative effort with the aim of positionally cloning genes predisposing to NIDDM and intermediate quantitative traits in Finnish subjects. The study design and family material of the FUSION study have been described in detail elsewhere (9) and we give only a brief description here. The index cases for the FUSION study were identified from various parts of Finland. Diabetes was diagnosed if the patient was on drug treatment (oral hypoglycemics and/or insulin) and/or blood glucose values from the medical records satisfied the World Health Organization (WHO) criteria for diabetes (10). If patients were untreated or treated by diet alone, and the fasting blood glucose was < 7 mmol/liter, the diagnosis of diabetes was confirmed by OGTT using WHO criteria (10). We selected families for study using the following criteria: (i) age of diagnosis of NIDDM in the index case between 35 and 60 yr; (ii) at least one living affected sibling; and (iii) at least one parent reported to be unaffected. Our age of diagnosis criterion was chosen so as to exclude most cases of insulin-dependent diabetes mellitus (IDDM) and maturity-onset diabetes of the young. All study subjects completed a clinic visit and a battery of questionnaires related to medical history, family history, and lifestyle.

All diabetics had C-peptide and glutamic acid decarboxylase (GAD) antibody measurements performed, in addition to fasting insulin and glucose levels. Families were excluded if probable late-onset IDDM was confirmed in the affected siblings or the first-degree relatives based on the following criteria: fasting C-peptide < 0.3 nmol/liter along with (i) positive GAD antibodies (≥ 0.03 GAD antibody units) and insulin started within 10 yr of diagnosis or (ii) negative GAD antibody and insulin treatment begun within 4 yr of diagnosis. Medical record review was used to exclude probable IDDM in cases where C-peptide or GAD antibody data were unavailable or where there was a normal C-peptide value along with positive GAD antibody and the patient had a duration of diabetes < 10 yr (9). 39 families were excluded because the proband, affected sibling, or another first-degree relative was considered to have probable IDDM.

Linkage studies can be negatively affected by misclassification of half-sibs or unrelated pairs as full sibs. Such misclassification cannot be established absolutely for families limited to a sib-pair, but can be inferred probabilistically. For each sib-pair, we calculated the probability of the marker data assuming the pairs were identical twins, full sibs, half-sibs, or unrelated using genotypes from 209 markers on 17 chromosomes (11). 23 ASPs in which the most likely relationship was either half-sib or unrelated were excluded from the analysis. The final analysis sample then comprised 472 families with 2 or more sibs affected with NIDDM. The 472 sibships included 381 with 2 affected sibs, 82 with 3, 6 with 4, and 1 each with 5, 6, or 7 affected sibs, yielding a total of 709 ASPs. 448 (95%) of these sibships had no parental genotypes, 21 had genotypes for one parent, and 3 had genotypes for both parents. Descriptive statistics for phenotypic measurements in males and females are shown in Table I.

DNA isolation and genotyping methods. DNA samples were isolated from whole blood using a salting-out procedure (GENTRA DNA isolation kit; Gentra, Minneapolis, MN). Each sample was di-

Table I. Phenotype Description of Males and Females from NIDDM-affected Sibling Pairs, Mean \pm SD

Variable	Males	Females
<i>n</i>	523	527
Age at enrollment (yr)	62.9 (7.9)	66.0 (8.2)
Age of diagnosis (yr)	51.3 (9.1)	52.9 (9.2)
Duration of diabetes (yr)	11.7 (7.1)	13.0 (7.8)
% On insulin treatment	37.3	42.7
% Treated with diet alone	10.5	12.7
Fasting plasma glucose (mmol/liter)	10.5 (3.5)	10.3 (3.3)
Fasting serum insulin (pmol/liter)	106 (68)	120 (66)
Fasting serum C-peptide (nmol/liter)	1.58 (0.96)	1.59 (0.92)
BMI (kg/m ²)	29.3 (4.3)	30.4 (5.1)
Height (cm)	173 (6)	159 (6)
Weight (kg)	88 (14)	77 (14)
Waist-to-hip ratio	0.98 (0.06)	0.90 (0.07)

luted to 10 ng/ μ l for amplification before being frozen in 96-well deep plates. Before PCR, DNA from the deep-well plates were aliquoted into 96-well plates (Micro-Amp; Perkin-Elmer Corp., Norwalk, CT) using a microdispenser (HYDRA 96; Robbins Scientific, Sunnyvale, CA) (12).

The fluorescently labeled dinucleotide repeat markers genotyped in this study were modified from a linkage mapping set (ABI PRISM; Perkin-Elmer/Applied Biosystems Division, Foster City, CA) in addition to in-house-designed primer pairs (13). Primer pairs for the following microsatellite markers were used: D2S338, D2S125, D2S140, D2S395, D2S367, D2S113, D2S142, D2S326, D2S117, D2S325, D2S164, IRS1, D2S126, D2S362, D2S206, D2S336, D2S286, D2S162, D2S121, D2S152, D2S305, D2S139, D2S319, D2S168, D2S151, D2S368, D2S383, D2S165, D2S337, and D2S391.

The reverse primers for all markers except D2S337 were modified by replacing the 5' end nucleotide with a guanine (G). This substitution results in a cytosine (C) nucleotide at the 3' end of the extended, fluorescently-labeled complementary strand. A C nucleotide in this position is known to optimally promote non-templated addition by *Taq* polymerase, thereby reducing the miscalling rate for allelic peaks using GENOTYPER (Perkin-Elmer/Applied Biosystems) (13). Singleplex PCR were performed on each DNA in each 96-well plate after aliquoting PCR mixes using a robotic liquid handling system (MultiPROBE 204DT; Packard Instrument Company, Downers Grove, IL). All PCR reactions were performed on PE9600 thermocyclers using AmpliTaq (Perkin-Elmer/Applied Biosystems) under cycling conditions of three-step plus 10-min extension at 72°C (13). This comprised 5 min at 95°C followed by 10 cycles of 94°C for 15 s, 55°C for 15 s, and 72°C for 30 s. This was followed by 20 cycles of 89°C for 15 s, 55°C for 15 s, and 72°C for 30 s. There was a final extension period of 10 min at 72°C.

The exception was marker D2S391, where a touchdown PCR protocol was used to avoid nonspecific products. In this case, one cycle for 5 min at 95°C was followed by 12 cycles of 94°C for 15 s, with annealing starting at 70°C and reducing by 1°C at each cycle down to 59°C, followed finally by 20 cycles (each comprising 89°C for 15 s, 58°C for 15 s, and 72°C for 30 s). This protocol was followed by a 10-min extension period at 72°C.

The pooling of amplified products, gel electrophoresis, and genotype analysis with GENESCAN, GENOTYPER (both from Perkin-Elmer/Applied Biosystems), and ABAS (a program that adjusts for intergel allele size variation and also offers an automated method for allele calling) have been described previously (12). Inheritance checks were performed by a modified version of the program USERM13 (14) of the pedigree analysis package MENDEL (15).

Statistical analysis. To estimate marker order and intermarker distances, we carried out multipoint analysis using CRIMAP (16)

jointly on (i) 210 FUSION families that were extended to include the spouse and one or more offspring of an affected sibling and (ii) the CEPH (Centre d'Étude Polymorphisme Humaine) reference pedigrees (17). We examined the robustness of our findings to marker map misspecification by reanalyzing the data using published distances (18).

As a preliminary analysis, we compared observed allele sharing identical by state (IBS) to that expected under no linkage (19) (Table II). As our primary analysis, we used mode-of-inheritance free ASP analyses to assess evidence for a locus on chromosome 2 using the programs SIBLINK (20), ASPEX (21, 22), and GENEHUNTER (23). We carried out maximum likelihood estimation under the constraints of the "possible triangle" (bounded by lines $z_0 = 0$, $z_1 = 1/2$, and $z_2 = 2z_0$, where z_i is the probability that an ASP shares i genes identical by descent) and for an additive genetic model for which $z_1 = 1/2$ and $z_0 \leq 1/4$ (24). We also carried out exclusion mapping under an additive model for several locus-specific risk ratios λ_s , including the $\lambda_s = 1.37$ estimated by Hanis et al. (5) for their locus on chromosome 2.

In our analyses, we used all possible ASPs from each family. As an approximate correction for the dependence among ASPs constructed from a sibship with s affected individuals, we weighted the

contribution of each ASP by $2/s$, yielding a total weight for the sibship of $s - 1$ (25). Parental genotype information was used in the 24 families for which it was available. Otherwise, IBD was estimated from IBS using gene frequencies estimated by gene counting from the family data. For comparison, we also estimated marker allele frequencies in 231 elderly unaffected Finnish controls; the effect on the results was minimal (data not shown).

Since the distribution of the maximum LOD score is unknown for this problem, we estimated empirical P values for these data by computer simulation assuming 578 independent ASPs (this being the approximately equivalent number of independent ASPs corresponding to our 709 ASPs from 472 families). We also estimated the power of our study to detect the locus identified by Hanis et al. (5).

Results

Genotyping accuracy in FUSION was assessed by genotyping 50 samples that were present in duplicate for each of the 50 markers. The genotyping personnel are blind to the identity of the duplicate samples. Of the 1,424 (94.9% of total) genotypes

Table II. IBS Sharing Statistics and Single-Point Maximum LOD Scores MLS for ASPs

Locus	Pos cM	Het	Number of shared alleles						Mean number of shared alleles		P	MLS
			0		1		2		Obs	Exp		
			Obs	Exp	Obs	Exp	Obs	Exp				
D2S319	0.0	0.751	49	54.8	327	334.3	274	260.8	1.346	1.317	0.115	0.56
D2S162	14.7	0.823	74	78.9	324	338.9	246	226.2	1.267	1.229	0.067	0.49
D2S168	21.5	0.859	72	85.1	335	327.2	208	202.7	1.221	1.191	0.129	0.26
D2S305	34.2	0.762	67	64.9	317	335.1	269	252.9	1.309	1.288	0.192	0.29
D2S165	41.5	0.842	94	83.8	314	334.5	225	214.8	1.207	1.207	0.500	0.02
D2S367	50.1	0.867	104	96.4	350	357.3	218	218.1	1.170	1.181	0.670	0.00
D2S391	65.2	0.730	55	53.8	345	337.4	268	276.7	1.319	1.334	0.729	0.00
D2S337	76.6	0.878	72	83.2	283	290.8	193	173.9	1.221	1.165	0.026*	0.28
D2S286	91.9	0.638	44	39.0	314	321.1	332	330.0	1.417	1.422	0.575	0.01
D2S139	96.6	0.823	73	81.2	365	360.1	243	239.6	1.250	1.232	0.245	0.08
D2S113	104.0	0.759	69	61.5	319	322.3	243	247.1	1.276	1.294	0.764	0.00
D2S121	118.2	0.829	78	80.4	347	345.9	229	227.8	1.231	1.225	0.417	0.05
D2S383	128.9	0.772	60	63.2	312	344.4	292	256.5	1.349	1.291	0.009*	0.47
D2S368	138.7	0.791	55	67.1	309	315.5	244	225.4	1.311	1.260	0.026*	0.72
D2S151	147.6	0.800	55	73.2	357	353.4	262	247.3	1.307	1.258	0.023*	0.47
D2S142	154.9	0.754	51	58.6	342	349.3	285	270.1	1.345	1.312	0.082	0.14
D2S326	170.7	0.865	104	96.5	345	356.1	222	218.4	1.176	1.182	0.591	0.00
D2S152	182.9	0.803	61	65.9	296	304.3	225	211.8	1.282	1.251	0.123	0.14
D2S117	187.3	0.833	89	86.7	378	363.4	221	237.9	1.192	1.220	0.868	0.14
D2S325	197.1	0.832	94	86.3	352	356.2	230	233.5	1.201	1.218	0.745	0.00
D2S164	206.6	0.742	87	58.8	327	336.1	249	268.1	1.244	1.316	0.998	0.00
IRS1	213.0	0.538	36	24.1	277	267.0	348	370.0	1.472	1.523	0.990	0.00
D2S126	214.4	0.813	68	77.1	365	352.1	236	239.8	1.251	1.243	0.375	1.05
D2S362	223.9	0.777	59	63.5	327	332.5	255	245.0	1.306	1.283	0.182	0.08
D2S206	230.4	0.750	63	61.6	343	344.4	270	269.9	1.306	1.308	0.528	0.00
D2S336	235.2	0.822	74	82.8	353	351.4	242	234.9	1.251	1.228	0.174	0.16
D2S338	240.8	0.807	51	67.7	308	311.6	235	214.7	1.310	1.247	0.009*	0.92
D2S125	251.0	0.819	85	74.5	324	331.6	220	223.0	1.215	1.236	0.799	0.00
D2S395	251.5	0.819	76	70.0	303	302.5	197	203.4	1.210	1.232	0.788	0.00
D2S140	252.4	0.736	66	48.9	286	295.1	230	238.1	1.282	1.325	0.954	0.00

Exp, expected; Het, heterozygosity for the marker; Obs, observed; Pos, position. P value reflects a one-tailed test. *Loci with increased mean sharing nominally significant at $P = 0.05$.

that could be scored unambiguously in both original and duplicate typings, three discrepancies were found, leading to a genotype error rate estimate of 0.11%. If any marker produced frequent null alleles or consistent PCR failures of one allele in a heterozygous genotype, one would expect to see an excess of homozygotes. No marker exhibited this phenomenon (data not shown).

Sex-averaged map distances were estimated using CRIMAP (16) assuming the Kosambi mapping function. When FUSION and CEPH data were combined, the most likely order for chromosome 2q telomeric markers was *D2S338*-10.2 cM-*D2S125*-0.5 cM-*D2S395*-0.9 cM-*D2S140*. In contrast to previous studies in which the order of the 2q telomeric markers, notably *D2S125*, was uncertain (5), we were able to order these markers at 1,000:1 relative maximum likelihood. Our resulting order is consistent with the current Genethon map (18).

The results of the multipoint linkage analyses provide no consistent evidence for a major locus for NIDDM in our Finnish families near 2qter based on the best telomeric marker order above (Fig. 1). Indeed, our data exclude linkage (multipoint LOD score < -2) for a locus-specific risk ratio λ of 1.50 or greater for much of the long arm of chromosome 2. At the maximum likelihood estimate of λ of 1.37 of Hanis et al. (5), we exclude 10 cM from the 2q telomere. Our simulation results suggest that this sample has a 93% power at a LOD score critical value of 2 and 77% power at a LOD score critical value of 3 to detect an additive gene effect of $\lambda_s = 1.37$, the value estimated by Hanis et al. (5).

Markers *D2S126* (at 214.4 cM) and *D2S338* (at 240.8 cM) show positive single-locus MLS values of 1.05 and 0.92, respectively, but multipoint LOD scores of 0.03 and 0.18. Further analysis also shows that the regions surrounding these markers can be excluded at a multipoint LOD score of -2 for $\lambda_s \geq 1.39$.

The region between *D2S383* (at 128.9 cM) and *D2S151* (at 147.6 cM) also shows excess sharing in ASPs (Table II) but can be excluded at $\lambda_s \geq 1.67$.

Discussion

The ability to detect significant gene effects in complex diseases is dependent on the phenotypic characteristics of the population, recurrence risk for the individual locus, the true underlying genetic model, sample size, polymorphism of the markers tested and density of the map, sampling design, and the efficiency of the statistical methods employed.

We have performed a multipoint ASP linkage analysis using all possible 709 sib-pairs in our 472 Finnish families and excluded the 2q telomeric region (*D2S338*-*D2S140*) at a $\lambda_s \geq 1.37$. Our study may be the first to exclude the telomeric region on chromosome 2q at the level of effect described in the original report. There are several possible explanations for this exclusion of the region identified by Hanis et al. (5). First, there may be a major gene near the telomere on 2q for NIDDM in Mexican Americans, but that variation in the gene plays no significant role in diabetes risk in Finland. We were able to exclude 2q37 at a λ_s value of 1.37 and above using LOD score = -2 as the exclusion criterion, indicating that our data are at least 100 times more likely in the absence of a disease locus than the presence of a locus for which $\lambda_s \geq 1.37$. Simulation studies conducted by Hauser et al. (20), using interval mapping, showed that the probability of falsely excluding linkage using this criterion was very small ($\sim 0.1\%$) for a range of values of λ_s . This argues that the putative 2q locus plays at most a very minor role in our Finnish sample.

In support of this potential ethnic difference, Hanis et al. (5) also studied 147 non-Hispanic white ASPs ascertained

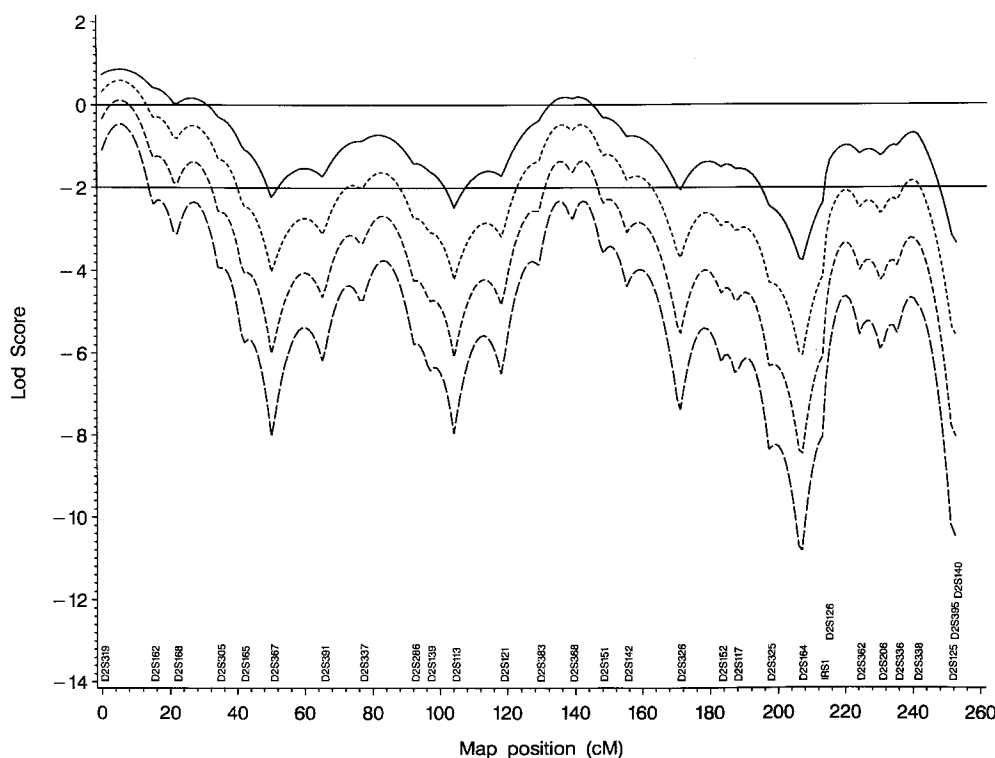


Figure 1. Multipoint exclusion map for chromosome 2 at sibling recurrence risk ratio, λ_s , of 1.25 (—), 1.37 (---), 1.50 (— — —), and 1.60 (— · — · —), corresponding to z_0 values of 0.200, 0.182, 0.166, and 0.156, respectively.

from Dresden, Germany and Chicago, Illinois. In this sample, the marker D2S338 (located 10.2 cM centromeric to D2S125) was shown to have a moderate LOD score with NIDDM (MLS = 1.15), whereas no evidence for linkage to D2S125 was observed. The single locus results reported here in a Finnish (Caucasian) sample also showed the same trend (MLS for D2S338 of 0.92), although the more informative multilocus analysis resulted in negative LOD scores for $\lambda_s > 1$ and the exclusion of this region (multipoint LOD score ≤ -2) around D2S338 for $\lambda_s \geq 1.39$. The possibility exists that a weak susceptibility locus is present near D2S338 in non-Hispanic Caucasians, but larger data sets than ours likely would be required to reach more definitive conclusions. The only other marker which showed a higher single locus MLS was D2S126 (MLS = 1.05), but this location resulted in negative multipoint LOD scores for $\lambda_s \geq 1.03$ and could be excluded in a multipoint analysis at $\lambda_s \geq 1.37$.

Second, sample differences other than ethnic origin might explain the differences in the results. One possibility is the alternative methods used for phenotyping. We used WHO criteria to define diabetes status in our sample, whereas Hanis et al. (5) used the National Diabetes Data Group criteria for the Mexican-American sample. There is a strong similarity between these criteria, and indeed WHO has endorsed the National Diabetes Data Group recommendations (26). Alternatively, Mexican-American diabetic subjects may have higher, underlying primary insulin resistance compared with their Finnish counterparts, which could explain our inability to replicate the results since the 2qter susceptibility gene may promote insulin resistance. Insulin resistance, unlike in the Mexican Americans (27), is not such a strong feature in Swedish (28) and Finnish diabetics (29), in whom impaired insulin secretion also seems to play an important role in the development of NIDDM.

The age of diagnosis was similar for the two populations. Among the 424 individuals in the 330 ASPs of the Mexican-American primary sample, the mean age of diagnosis \pm SD was 50.0 ± 11.6 yr in men and 48.7 ± 10.7 yr in women. In our sample, the corresponding values were 51.3 ± 9.0 yr for males and 52.9 ± 9.2 yr for females. It is therefore unlikely that differences in age of diagnosis accounted for the nonreplication of the Mexican-American sample linkage.

Third, the discrepancies in the two studies could be due to differences in the maps used in the analysis. Hanis et al. (5) report maximum multipoint LOD scores ranging from 2.4 to 4.3, depending on the order of the markers. By combining FUSION data with the Centre d'Étude Polymorphisme Humaine reference genotypes, we were able to order the markers at 1,000:1 odds. In addition, using the Hanis et al. (5) map order did not change the conclusions of our analysis (data not shown). As our map has the same order as the GENETHON map and our genotyping error rates were small, it is unlikely that misspecification of marker order is the cause of our failure to replicate linkage.

Fourth, because many of our families were composed of a single ASP (356 families with no parents) we could not assess their true relationships with certainty. To address this concern, we compared the probability of our marker data assuming the sibling pairs were full sibs to that assuming half-sibs, unrelated pairs, or monozygotic twins across 209 markers on 17 chromosomes (11). On this basis we excluded 23 pairs from our analysis. Including these individuals had little effect on the results.

Thus it is unlikely that we are falsely excluding the telomeric region on 2q due to incorrect inference of relationships amongst putative sibs.

Fifth, the Hanis et al. (5) finding on chromosome 2q could be a false positive. In this regard, intervals without flanking markers (as at telomeres) cause an increase in the rate of false positive linkage (30). Recently, however, data from additional markers more telomeric to D2S125 in the Mexican-American study make it unlikely that an "open-ended effect" is the primary cause for a false positive linkage (31). The Hanis et al. (5) data just manages to reach genome-wide significance using MLS analysis however, and thus could occur by chance in 5% of genome scans. The result would gain more credence if a separate study with a similar population replicated linkage in this region. Only Stern et al. (32) have published results on a similar ethnic group of 444 individuals from 32 families with 83 diabetics. No evidence for linkage has been obtained for an age-of-onset gene for diabetes nor for fasting or 2-hr glucose. However, it is unclear how powerful the Stern et al. study is to detect a locus with a λ_s of 1.37. In a separate study, Hani et al. (33) could not exclude NIDDM1 at a λ_s value of 1.36 in a French sample of 449 affected sib-pairs using criteria for diabetes which were less stringent than the Hanis et al. study in the Mexican Americans. McCarthy et al. (34) were unable to show excess sharing in the same region in 278 affected sib-pairs from the United Kingdom. However, no exclusion mapping was performed in the latter study.

In conclusion, we present evidence that excludes the proposed chromosome 2q locus at $\lambda_s \geq 1.37$ in a Finnish sample of 709 sib-pairs from 472 sibships. The marker D2S125, which was most closely linked to the putative locus, can be excluded at $\lambda_s \geq 1.22$. Our results suggest that this region plays, at most, a very minor role in determining NIDDM risk in the Finnish population. Further studies are needed in similar and different populations.

Acknowledgments

The FUSION project is made possible by intramural funds from the National Human Genome Research Institute (project no. OH95-C-N030) and by R01 HG-00376. Family studies were approved by Institutional Review Boards at the National Institutes of Health (assurance number SPA S-5737-05) and at the National Public Health Institute in Helsinki, Finland. R.M. Watanabe was previously supported by a training grant T32 HG-00040 from the National Institutes of Health and is currently supported by individual postdoctoral award F32 DK-09525. W. Hagopian was supported by a Clinical Research Center M01RR-00037 Clinical Associate Physician Award.

References

1. Bennett, P.H., C. Bogardus, J. Tuomilehto, and P. Zimmet. 1992. Epidemiology and natural history of NIDDM: non-obese and obese. *In* International Textbook of Diabetes Mellitus. K.G.M.M. Alberti, R.A. DeFronzo, H. Keen, and P. Zimmet, editors. John Wiley & Sons, New York. 147-176.
2. Tuomilehto, J., H.J. Korhonen, L. Kartovaara, V. Salomaa, J.H. Stengard, M. Pitkanen, A. Aro, K. Javela, M. Uusitupa, and J. Pitkaniemi. 1991. Prevalence of diabetes mellitus and impaired glucose tolerance in the middle-aged population of three areas in Finland. *Int. J. Epidemiol.* 20:1010-1017.
3. Valle, T., J. Tuomilehto, and J. Eriksson. 1997. Epidemiology of NIDDM in Europids. *In* International Textbook of Diabetes Mellitus. K.G.M.M. Alberti, P. Zimmet, R.A. DeFronzo, and H. Keen, editors. John Wiley & Sons, New York. 125-142.
4. Ghosh, S., and N.J. Schork. 1996. Genetic analysis of NIDDM. The study of quantitative traits. *Diabetes.* 45:1-14.
5. Hanis, C., E. Boerwinkle, R. Chakraborty, D. Ellsworth, P. Concannon,

- B. Stirling, V.A. Morrison, B. Wapelhorst, R. Spielman, K.J. Gogolin-Ewens, et al. 1996. A genome-wide search for human non-insulin-dependent (type 2) diabetes genes reveals a susceptibility locus on chromosome 2. *Nat. Genet.* 13:161–166.
6. Mahtani, M.M., E. Widen, M. Lehto, J. Thomas, M. McCarthy, J. Brayer, B. Bryant, G. Chan, M. Daly, C. Forsblom, et al. 1996. Mapping of a gene for type 2 diabetes associated with an insulin secretion defect by a genome scan in Finnish families. *Nat. Genet.* 14:90–94.
7. Kaprio, J., J. Tuomilehto, M. Koskenvuo, K. Romanov, A. Reunanen, J. Eriksson, J. Stengård, and Y.A. Kesäniemi. 1992. Concordance for Type 1 (insulin-dependent) and Type 2 (non-insulin-dependent) diabetes mellitus in a population-based cohort of twins in Finland. *Diabetologia.* 35:1060–1067.
8. Lander, E.S., and L. Kruglyak. 1995. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat. Genet.* 11:241–249.
9. Valle, T., J. Tuomilehto, R.N. Bergman, S. Ghosh, E.R. Hauser, J. Eriksson, S.J. Nylund, K. Kohtamaki, E. Tuomilehto-Wolf, L. Toivanen, et al. 1998. Mapping genes for non-insulin dependent diabetes mellitus: design of the Finland-United States Investigation of NIDDM genetics (FUSION) study. *Diabetes Care.* 21:949–958.
10. World Health Organization. 1980. World Health Organization: WHO expert committee on diabetes mellitus. Second Report. Technical Report Series, no. 646. Geneva, Switzerland. 7–80.
11. Boehnke, M., and N.J. Cox. 1997. Accurate inference of relationships in sibling pair linkage studies. *Am. J. Hum. Genet.* 61:423–429.
12. Ghosh, S., Z. Karanjawala, E.R. Hauser, D. Ally, J.I. Knapp, J.B. Rayman, A. Musick, J. Tannenbaum, C. Te, S. Shapiro, et al. 1997. Methods for precise sizing, automatic binning of alleles and reduction of error rates in large-scale genotyping using fluorescently-labeled dinucleotide markers. *Genome Res.* 7:165–178.
13. Magnuson, V., D. Ally, S. Nylund, Z. Karanjawala, J. Rayman, J. Knapp, A. Lowe, S. Ghosh, and F.S. Collins. 1996. Substrate nucleotide-determined non-templated addition of adenine by Taq DNA polymerase: implications for PCR-based genotyping and cloning. *BioTechniques.* 21:700–709.
14. Boehnke, M. 1991. Allele frequency estimation from data on relatives. *Am. J. Hum. Genet.* 48:22–25.
15. Lange, K., D. Weeks, and M. Boehnke. 1988. Programs for pedigree analysis: MENDEL, FISHER and dGENE. *Genet. Epidemiol.* 5:471–472.
16. Lander, E.S., and P. Green. 1987. Construction of multilocus genetic linkage maps in humans. *Proc. Natl. Acad. Sci. USA.* 84:2363–2367.
17. Dausset, J., H. Cann, D. Cohen, M. Lathrop, J.-M. Lalouel, and R. White. 1990. Centre d'Étude du Polymorphisme Humain (CEPH): Collaborative genetic mapping of the human genome. *Genomics.* 6:575–577.
18. Dib, C., S. Faure, C. Fizames, D. Samson, N. Drouot, A. Vignal, P. Millasseau, S. Marc, J. Hazan, E. Seboun, et al. 1996. A comprehensive genetic map of the human genome based on 5,264 microsatellite. *Nature.* 380:152–154.
19. Bishop, D.T., and J.A. Williamson. 1990. The power of identity-by-state methods for linkage analysis. *Am. J. Hum. Genet.* 46:254–265.
20. Hauser, E.R., M. Boehnke, S.-W. Guo, and N. Risch. 1996. Affected-sib pair interval mapping and exclusion for complex genetic traits: sampling considerations. *Genet. Epidemiol.* 13:117–137.
21. Risch, N. 1990. Linkage strategies for genetically complex traits. II. The power of affected relative pairs. *Am. J. Hum. Genet.* 46:229–241.
22. Risch, N. 1993. Exclusion mapping for complex disease. *Am. J. Hum. Genet.* 53:A185. (Abstr.)
23. Kruglyak, L., M.J. Daly, M.P. Reeve-Daly, and E.S. Lander. 1996. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am. J. Hum. Genet.* 58:1347–1363.
24. Holmans, P. 1993. Asymptotic properties of affected-sib pair linkage analysis. *Am. J. Hum. Genet.* 52:362–374.
25. Suarez, B.K., and S.E. Hodge. 1979. A simple method to detect linkage for rare recessive diseases: an application to juvenile diabetes. *Clin. Genet.* 15:126–136.
26. National Diabetes Data Group. 1979. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. *Diabetes.* 28:1039–1057.
27. Haffner, S.M., H. Miettinen, and M.P. Stern. 1996. Insulin secretion and resistance in non-diabetic Mexican-Americans and non-hispanic whites with a parental history of diabetes. *J. Clin. Endocrinol. Metab.* 81:1846–1851.
28. Cerasi, E., S. Efendic, and R. Luft. 1973. Dose-response relation between plasma-insulin and blood-glucose levels during oral glucose loads in pre-diabetic and diabetic subjects. *Lancet.* 1:794–797.
29. Eriksson, J., A. Franssila-Kallunki, A. Ekstrand, C. Saloranta, E. Widen, C. Schalin, and L. Groop. 1989. Early metabolic defects in persons at increased risk for non-insulin-dependent diabetes mellitus. *N. Engl. J. Med.* 321:337–343.
30. Risch, N., and L. Giuffra. 1992. Model misspecification and multipoint linkage analysis. *Hum. Hered.* 42:77–92.
31. Bell, G.I., N.J. Cox, T. Lindner, P. Concannon, R.S. Spielman, E. Boerwinkle, and C.L. Hanis. 1997. Genetics of NIDDM in Mexican-Americans of Starr County, Texas: an update. *Diabetic Rev.* 5:277–283.
32. Stern, M.P., R. Duggirala, B.D. Mitchell, L.J. Reinhart, S. Shivakumar, P.A. Shipman, O.C. Uresandi, E. Benavides, J. Blangero, and P. O'Connell. 1996. Evidence for linkage of regions on chromosome 6 and 11 to plasma glucose concentrations in Mexican Americans. *Genome Res.* 6:724–734.
33. Hani, E.H., J. Hager, A. Phillipi, F. Demenais, P. Froguel, and N. Vionnet. 1997. Mapping NIDDM susceptibility loci in French families: studies with markers in the region of NIDDM1 on chromosome 2q. *Diabetes.* 46:1225–1226.
34. McCarthy, M., P. Cassell, P. Saker, M. Armstrong, R. Turner, S. O'Rahilly, M. Walker, A. Hattersley, and G. Hitman. 1997. No excess allele-sharing on chromosome 2q (region of NIDDM1) in a large European sibship collection. *Diabetes.* 46:0201. (Abstr.)