Am. J. Hum. Genet. 57:911–919, 1995

Affected-Sib-Pair Mapping of a Novel Susceptibility Gene to Insulin-Dependent Diabetes Mellitus (*IDDM8*) on Chromosome 6q25-q27

De-Fang Luo,[1] Marilyn M. Bui,[1] Andrew Muir,[1] Noel K. Maclaren,[1] Glenys Thomson,[3] and Jin-Xiong She[1,2]

[1]Department of Pathology and Laboratory Medicine and [2]Center for Mammalian Genetics, College of Medicine, University of Florida, Gainesville; and [3]Department of Integrative Biology, University of California at Berkeley, Berkeley

Summary

Affected-sib-pair analyses were performed using 104 Caucasian families to map genes that predispose to insulin-dependent diabetes mellitus (IDDM). We have obtained linkage evidence for *D6S446* (maximum lod score [MLS] = 2.8) and for *D6S264* (MLS = 2.0) on 6q25-q27. Together with a previously reported data set, linkage can be firmly established (MLS = 3.4 for *D6S264*), and the disease locus has been designated *IDDM8*. With analysis of independent families, we confirmed linkage evidence for the previously identified *IDDM3* (15q) and *DDM7* (2q). We also typed additional markers in the regions containing *IDDM3, IDDM4, IDDM5,* and *IDDM8*. Preliminary linkage evidence for a novel region on chromosome 4q (*D4S1566*) has been found in 47 Florida families ($P < .03$). We also found evidence of linkage for two regions previously identified as potential linkages in the Florida subset: *D3S1303* on 3q ($P < .04$) and *D7S486* on 7q ($P < .03$). We could not confirm linkage with eight other regions (*D1S191, D1S412, D4S1604, D8S264, D8S556, D10S193, D13S158,* and *D18S64*) previously identified as potential linkages.

Introduction

Insulin-dependent diabetes mellitus (IDDM) is an autoimmune disease with immunologically mediated destruction of the insulin-producing beta cells within the pancreatic islets. As an etiologically and genetically heterogeneous disease, the expression of the IDDM phenotype is influenced by a number of susceptibility genes as well as by environmental factors. The HLA class II region (sometimes referred to as *IDDM1*) actually contains several genes that are involved in IDDM susceptibility (Todd et al. 1987; Thorsby and Ronningen 1992; Caillat-Zucman et al. 1993; Deng et al. 1995). *IDDM2,* located on chromosome 11q15.5 in the insulin gene (INS) region, was initially identified by case-control, then confirmed by intrafamilial, association studies and affected-sib-pair (ASP) analysis (Julier et al. 1991; Bain et al. 1992; Lucassen et al. 1993; Owerbach and Gabbay 1994; She et al. 1994; Bennett et al. 1995).

Genetic contribution to a disease can be evaluated by the degree of familial clustering of a disease estimated by λ_s, the ratio of the risk for siblings of patients to the population prevalence (Risch 1987). While the *HLA* class II genes ($\lambda_s = 3.1$) are the most important susceptibility factors (Davies et al. 1994; Hashimoto et al. 1994), they probably account for <50% of the total genetic contribution to the disease ($\lambda_s \approx 15$) (Risch 1987). Further, *INS* ($\lambda_s = 1.3$) and *HLA* together only explain a portion ($\lambda_s = 4.0$) of the familial clustering, suggesting that other susceptibility factors must exist. Indeed, linkage studies have suggested that ≥10 genes are involved in the expression of insulitis and/or diabetes in the nonobese diabetic (NOD) mouse (Todd et al. 1991; Ghosh et al. 1993). The number of susceptibility genes is probably even higher in human IDDM, in light of the ethnic and genetic heterogeneities apparent in the disease. Recently, several groups have undertaken genome-wide linkage mapping of IDDM susceptibility genes, using families with ASPs. A two-stage approach has generally been used in most studies (Davies et al. 1994; Hashimoto et al. 1994). The first stage involves an initial genome-wide screen using a subset of ASPs and a large number of microsatellite markers, while the second stage is used to confirm and/or replicate linkage with additional families and markers. Such studies to date have identified at least four more non-HLA genomic regions containing IDDM genes (Davies et al. 1994; Field et al. 1994; Hashimoto et al. 1994; Copeman et al. 1995; Owerbach and Gabbay 1995). *IDDM3* is located in a region near *D15S107* and *D15S87* on chromosome 15q26 (Field et al. 1994). Evidence for linkage was suggestive in the initial data set studied and was confirmed in the combined data set. *IDDM4,* located

Received February 2, 1995; accepted for publication June 28, 1995.

Address for correspondence and reprints: Dr. Jin-Xiong She, Department of Pathology and Laboratory Medicine, Box 100275, College of Medicine, University of Florida, Gainesville, FL 32610.

0002-9297/95/5704-0023$02.00

near the *FGF3* locus on chromosome 11q13, has been reported by three different groups (Davies et al. 1994; Field et al. 1994; Hashimoto et al. 1994). Support of *IDDM5* near *ESR* on chromosome 6q was found in two of three data sets analyzed (Davies et al. 1994). Linkage with markers on chromosome 2q was initially reported by Davies et al. (1994), and *IDDM7* was confirmed by a linkage study (Owerbach, Gabbay 1995) and linkage disequilibrium mapping (Copeman et al. 1995). Several other genomic regions may also contain IDDM genes (Davies et al. 1994; Hashimoto et al. 1994). Of these linkages, only the one on 18q (*D18S64*) was found in a second data set. Linkage to kidd blood group near *D18S64* was obtained >10 years ago (Hodge et al. 1981; Hodge et al. 1983), although a second independent study failed to replicate the linkage (Dunsworth et al. 1982). The other linkages have not been replicated in additional data sets. All these studies indicated that mapping of genes for complex diseases such as IDDM is extremely difficult and requires several hundred ASPs as well as a high density map of polymorphic markers (Davies et al. 1994). The large number of genes involved and high degree of heterogeneity of IDDM mandates that families from many ethnic groups and/or geographic regions be studied to ensure the identification of all regions linked to IDDM and confirmation of linkages that have already been reported.

We have carried out a screen for IDDM susceptibility genes, using 100 markers and 25–104 ASP families. We have also analyzed additional families and markers in three regions containing *IDDM3*, *IDDM4*, and *IDDM5* as well as 10 other markers previously identified as potential linkages.

Families, Material and Methods

Study Families

We have obtained genomic DNA from a total of 104 United States families of Caucasian ancestry for our mapping studies. Each family has two affected siblings; no families with more than two affected siblings were included in this analysis. We ascertained 47 of the families from the southeastern United States, mostly north-central Florida (the Florida data set). Cell lines have been established for most of the Florida families and have recently become available to other investigators through the human biological data interchange (HBDI). Forty-nine other families were obtained from the HBDI (HBDI data set), and 8 more were kindly provided by Dr. Richard Spielman (University of Pennsylvania).

Microsatellite-Marker Typing

Microsatellite markers were genotyped using radioactive labeling of PCR primers and denaturing polyacrylamide gel electrophoresis. In brief, one of the PCR primers was end-labeled using γ^{32}P-ATP and T4 polynucleotide kinase. PCR amplifications were performed on 40 ng of genomic DNA (prealiquoted into a 96-well microtitre plate) in 12-μl reaction volume containing 50 mM KCl; 10 mM Tris-Cl; pH 8.3; 1.5 mM $MgCl_2$; and 60 μM of each dNTP. Samples were subjected to 27–30 cycles of 30 s at 94°C for denaturing, 30 s at optimum annealing temperature, and 30 s at 72°C for extension, using a Perkin-Elmer-Cetus 9600 thermal cycler. After PCR amplification, 2 volumes of sequencing loading solution (0.3% xylene cyanol, 0.3% bromophenol blue, 10 mM EDTA, pH 8.0, and 90% (v/v) formamide) are added. The samples are then heated at 95°C for 10 min to denature DNA and 2–4 μl were immediately loaded onto a 6% polyacrylamide DNA sequencing gel. Products from 3–4 different markers with nonoverlapping allele sizes (amplified in separate reactions) can be combined together before loading to genotype multiple markers simultaneously. Alternatively, products of the same marker (but different samples) can be loaded four times (each separated by 30–60 min). Multiplexing of different markers or multiple loading of products from the same marker can greatly increase the efficiency of genotyping. Most microsatellite primers were purchased from Research Genetics. Some of them did not give satisfactory results. New primers for such problematic markers were designed on the basis of published sequences. The newly designed primers usually give better results (primer sequences are available on request).

Data Analysis

The ASP method compares the observed identity by descent (IBD) values with random expectations of parental alleles in ASPs. In meioses from heterozygous parents, random sharing of alleles by ASPs occurs at a 50% frequency, and in fully informative families a 1:2:1 distribution for sharing 2, 1, and 0 alleles would be expected. Parents homozygous for an allele were considered uninformative for ASP analyses. The χ^2 test was used to determine the statistical significance of the excess of gene sharing by ASPs. The χ^2 score was calculated using $(1\ \text{IBD} - 0\ \text{IBD})^2/(1\ \text{IBD} + 0\ \text{IBD})$, with 1 df. Correction for multiple comparisons was not performed, since (1) it is not required in a follow up study of confirmed or suspected linkages, and (2) we wished to identify potential linkages that would then be further studied in additional data sets. To increase the informativeness of the families, polymorphic flanking markers were used to deduce the transmission of alleles from homozygous parents (referred to as "haplotyping"). Markers spaced at <5 cM were used in haplotyping to minimize the possibility of double recombinants.

The maximum lod score (MLS) statistic T was calculated according to the method of Risch (1990) by us-

ing the following equation: $T = N_1(\log_{10}[N_1/0.5N]) + N_0(\log_{10}[N_0/0.5N])$, where N_1 and N_0 are the numbers of ASPs sharing 1 and 0 alleles, respectively, and N is the total number of informative meioses ($N_0 + N_1$). This MLS statistic T is not exactly the same as the MLS reported by Davies et al. (1994).

Genome Screen for Linkage

Initially, 25–47 of the Florida families were analyzed for 100 microsatellite markers distributed randomly throughout the human genome. Regions showing evidence of linkage in the first 25 families (*IL2RB* and *D15S107*), as well as candidate regions on 4q (*D4S393, D4S243,* and *D4S1596*) syntenic to mouse chromosome 3, which contains *Idd3* in the NOD mouse, and markers on 6q (*D6S311, D6S437,* and *D6S264*) where the candidate gene *SOD2* is located, were typed in an additional 22 families from Florida, 49 families from the HBDI collection, and 8 families provided by Dr. Spielman. While completing our analyses on these markers, the linkages of *IDDM3* (15q), *IDDM4* (11q), and *IDDM5* (6q) were reported (Davies et al. 1994; Field et al. 1994; Hashimoto et al. 1994), and preliminary evidence for linkage, requiring confirmation, was reported for *D1S191, D1S412, D4S1604,* and *D8S556* in a French data set (Hashimoto et al. 1994) and *D2S326, D3S1303, D6S264, CFTR* (7q), *D8S556, D10S193, D13S158, D18S64, DXS991,* and *DXS999* in the U.K. 96 data set (Davies et al. 1994). As candidate regions, 10 of these markers (*D1S191, D1S412, D4S1604, D8S556, D2S326, D3S1303, D7S486* (close to *CFTR*), *D10S193, D13S158,* and *D18S64*) were also typed in all families. Additional markers around *IDDM3, IDDM4,* and *IDDM5,* and around linkages identified in our screen, were also typed in all families.

Results

Overall Survey

Only *IL2RB* on 22q ($P < .01$) and *D15S107* on 15q ($P < .04$) showed some evidence for linkage in our initial survey of 25 ASPs. *IL2RB* showed no evidence for linkage when all 104 ASPs were analyzed (table 1). The initial result may represent type I error indicative of the very small sample size used. *D15S107* showed linkage in our overall data set ($P < .008$) (table 2) and in 31 families independent of those studied by Field et al. (1994) (20 in the Florida data set, 3 in the HBDI data set, and 8 provided by Dr. Spielman). Linkage in these 31 families ($P < .02$) thus provide additional evidence for *IDDM3* (table 1).

Linkage evidence for *IDDM4* was not observed in our independent data set (table 1), although positive evidence was obtained in the HBDI data set (data not shown, because of overlap of families with previous reports). Evidence of linkage for *IDDM5* was not observed in the Florida (table 1) or the HBDI (data not shown) families. The linkage of *IDDM7* was also observed in our independent Florida data set ($P < .02$) (table 1).

Both of our candidate gene regions gave positive results. Preliminary linkage evidence was obtained for *D4S1566* in the Florida ($P < .03$) and total ($P < .03$) data sets, although the HBDI data were not significant (table 1). Davies et al. (1994) found linkage with *D6S264* on 6q in one data set (U.K. 96) examined ($P < .01$). We found strong evidence for linkage to *D6S264* in our total data set ($P = .0013$), with significant linkage in both the Florida ($p < .03$) and HBDI ($P < .01$) subsets (table 1). Our findings thus establish linkage of *D6S264,* and the disease locus is now designated *IDDM8* (P. McAlpine, personal communication; also see next section).

Our Florida data set also showed linkage for two markers identified as potential linkages in the only U.K. 96 data set studied by Davies et al. (1994): *D3S1303* ($P < .04$) and *D7S486* ($P < .03$). The HBDI and our total data set did not show linkage. Heterogeneity in evidence for linkage or association was observed between data sets by Davies et al. (1995) and Copeman et al. (1995) in the study of *IDDM4, IDDM5,* and *IDDM7.* It is most likely that the linkages on 3q and 7q are real, but independent confirmation in other data sets is required. Other markers that had previously shown evidence for linkage in one data set (i.e., *D1S191, D1S412, D4S1604,* and *D8S556* [Hashimoto et al. 1994] and *D8S264, D10S193, D13S158,* and *D18S64* [Davies et al. 1994]) showed no evidence for linkage in either the Florida, HBDI, or total data sets (table 1).

Genetic Heterogeneity According to HLA-DR/DQ Status of the Sib Pairs

To test HLA-associated heterogeneity, families were subdivided into two groups according to HLA sharing (identity-by-state of HLA-DR/DQ haplotypes): sib pairs who share two identical HLA haplotypes (HLA 2) and sib pairs who share one or zero HLA haplotype (HLA 1,0). Linkages with markers on 6q (*IDDM8*), 11q (*IDDM4*), and 15q (*IDDM3*) were also analyzed according to the HLA-DR types of the sib pairs (DR3/4, DR3/non-DR4, or DR4/non-DR3). There were variations in proportions of genes shared by ASPs for most markers analyzed here between the HLA 2 and HLA 1,0 categories (table 1) or between HLA-DR genotypes (data not shown). However, there were no particular trends, and none of the comparisons reached statistical significance.

Mapping of IDDM8 on Chromosome 6q

We have analyzed 15 markers on 6q in all 104 families, to further localize *IDDM8.* As shown in table 3

Table 1

Linkage Evidence for Various Marker Loci in Families That Have Not Been Reported in Previous Studies

IDDM Gene and Chromosome	Marker Locus	Data Set[a]	HLA 2[b]				HLA 1.0[b]				Total[b]			
			1	0	PGS	*P*	1	0	PGS	*P*	1	0	PGS	*P*
IDDM3:														
15q	*D15S107*[c]	New	22	9	71.0	.020	13	9	59.1	ns	35	18	66.0	.020
IDDM4:														
11q	*D11S1337*[d]	New	14	10	58.3	ns	11	11	50.0	ns	25	21	54.3	ns
IDDM5:														
6q	*ESR*[e]	Florida	30	31	49.2	ns	15	10	60.0	ns	45	41	52.3	ns
IDDM7:														
2q	*D2S326*[f]	Florida	20	12	62.5	ns	11	3	78.6	.033	31	15	67.4	.018
IDDM8:														
6q	*D6S264*[e]	Florida	27	11	71.0	.0094	11	8	57.9	ns	35	19	64.8	.030
		HBDI	32	21	60.4	ns	14	7	66.7	ns	52	28	65.0	.0073
		Total	62	34	64.5	.0043	27	17	61.4	ns	89	51	63.6	.0013
3q	*D3S1303*[e]	Florida	32	17	65.3	.032	9	7	56.3	ns	41	24	63.1	.035
		HBDI	23	24	48.9	ns	14	13	51.9	ns	37	37	50.0	ns
		Total	56	43	56.6	ns	26	22	54.2	ns	82	65	55.8	ns
4q	*D4S1566*[g]	Florida	32	21	60.4	ns	14	6	70.0	.074	46	27	63.0	.026
		HBDI	34	26	56.7	ns	14	14	50.0	ns	48	40	54.5	ns
		Total	72	48	60.0	.029	30	25	54.5	ns	102	73	58.3	.028
7q	*D7S486*[e]	Florida	34	21	61.8	.08	14	7	66.7	ns	48	28	63.2	.022
		HBDI	23	31	42.6	ns	16	9	64.0	ns	39	40	48.1	ns
		Total	60	54	52.6	ns	32	19	62.8	.069	92	73	55.8	ns
1q	*D1S191*[h]	Total	34	50	40.5	ns	16	11	59.3	ns	50	61	45.1	ns
1q	*D1S412*[h]	Total	35	38	48.0	ns	18	12	60.0	ns	53	50	51.5	ns
4q	*D4S1604*[h]	Total	38	34	52.8	ns	19	11	63.3	ns	57	45	55.8	ns
8p	*D8S264*[e]	Total	42	46	47.7	ns	17	22	43.6	ns	59	68	46.5	ns
8q	*D8S556*[h]	Total	39	36	52.0	ns	24	20	54.5	ns	63	56	52.9	ns
10cen	*D10S193*[e]	Total	40	51	44.0	ns	21	17	55.3	ns	61	68	47.3	ns
13q	*D13S158*[e]	Total	41	45	47.7	ns	17	17	50.0	ns	58	62	48.3	ns
18q	*D18S64*[e]	Total	27	39	40.9	ns	12	24	33.3	.046[i]	39	63	38.2	.018[i]
22q	*IL2RB*[g]	Total	53	49	52.0	ns	32	23	58.2	ns	82	71	53.6	ns

NOTE.—All marker loci were analyzed in 49 HBDI families, 47 "Florida" families, and 8 families provided by Dr. Spielman. The "total" includes all 104 families in the three data sets. All families reported in this table are independent of families used in previously published studies (Davies et al. 1994; Field et al. 1994; Hashimoto et al. 1994; Owerbach and Gabbay 1995). Sib pairs were divided into those sharing two HLA haplotypes (HLA 2) and those sharing 1 or 0 haplotypes (HLA 1, 0).

[a] Data set for *IDDM3* is independent of data sets reported by Field et al. (1994) and includes 20 Florida, 3 HDBI, and 8 Spielman families. Data set for *IDDM4* is independent of the data set reported by Davies et al. (1994), Field et al. (1994), and Hashimoto et al. (1994) and includes 20 Florida and 8 Spielman families.

[b] 1 = 1 IBD; 0 = 0 IBD; PGS = percent of gene sharing (1 IBD/[1 IBD = 0 IBD]); ns = not significant ($P > .1$).

[c] Positive linkage evidence found for the marker in the study by Field et al. (1994).

[d] Positive linkage evidence found for the marker in the studies by Davies et al. (1994), Field et al. (1994), and Hashimoto et al. (1994).

[e] Positive linkage evidence found for the marker in the study by Davies et al. (1994).

[f] Positive linkage evidence found for the marker in the studies by Davies et al. (1994), Copeman et al. (1995), and Owerbach and Gabbay (1995).

[g] Positive linkage evidence found for the marker in the present study.

[h] Positive linkage evidence found for the marker in the study by Hashimoto et al. (1994).

[i] Gene sharing significantly <50%.

Table 2

Mapping of *IDDM3* on Chromosome 15q26

Marker	Distance (cM)	1 IBD	0 IBD	PGS	*P*	MLS
D15S130	0	79	70	53.0	ns	.1
D15S230	10	108	81	57.1	.050	.9
D15S533	13	113	78	59.1	.011	1.4
D15S107	14	114	77	59.7	.0074	1.6
D15S120	17	104	78	57.1	.054	.8
D15S87	22	102	81	55.7	ns	.5
D15S642	31	86	72	54.8	ns	.3

NOTE.—Seven markers on chromosome 15q were analyzed, and haplotyping analysis was performed to increase the informativeness of the markers. Genetic distances from *D15S130* were taken from published maps (Gyapay et al. 1994) for all markers except *D15S533*, which was mapped only to the 15q region. Its location between *D15S230* and *D15S107* was confirmed by several recombinational events in our own data set. Probability (*P*) was estimated using the χ^2 test. Maximum lod score (MLS) was estimated using Risch's equation (1990). PGS = percent of gene sharing; ns = not significant.

and figure 1, the strongest evidence of linkage was obtained at *D6S446* after haplotyping (MLS = 2.8). *D6S264* had a MLS of 2.0 in our data set and 1.4 in the U.K. 96 data set (Davies et al. 1994). These two data sets together gave a MLS of 3.4 for *D6S264*. It is interesting that *ESR*, near which *IDDM5* was previously mapped, had an MLS of only 0.9 in our data set. The MLS for *ESR* was nevertheless slightly higher than those for its flanking markers (table 3 and fig. 1). Since evidence of linkage for *IDDM5* had been obtained in several data sets (Davies et al. 1994) and *ESR* is separated from *D6S264* by ~28 cM, the 6q region appears to contain two distinct diabetes genes. The disease locus near *D6S264* has now been designated *IDDM8*. Indeed, the results obtained in the U.K. 96 data set (Davies et al. 1994) also had distinct peaks at *ESR* and *D6S264*

Table 3

Mapping of *IDDM8* on Chromosome 6q

Marker	Distance (cM)	1 IBD	0 IBD	PGS	*P*	MLS
D6S311	0	88	85	50.9	ns	.0
D6S476	2	101	88	53.4	ns	.2
ESR	4	110	82	57.3	.043	.9
D6S440	6	109	90	54.8	ns	.4
D6S290	7	108	90	54.5	ns	.4
D6S442	10	110	89	55.3	ns	.5
D6S415	13	107	89	54.6	ns	.4
D6S437	15	107	82	56.6	ns	.7
D6S253	22	112	81	58.0	.026	1.1
D6S396	25	109	83	57.1	ns	.8
D6S264	32	117	75	60.9	.0024	2.0
D6S297	35	113	76	59.8	.007	1.6
D6S446	41	116	68	63.1	.0004	2.8
D6S281	42	107	63	63.3	.0007	2.5
TBP	43	79	48	62.2	.006	1.7

NOTE.—IBD values for all markers except *D6S311* and *TBP* are results after haplotyping with flanking markers listed here. The distances were taken from the published maps (Davies et al. 1994; Gyapay et al. 1994) except *D6S476* and *D6S396*, which were mapped using our own data set. The distance between *D6S264* and *D6S281* estimated from our data set is in close agreement with that calculated using the U.K. diabetic families (Davies et al. 1994) but is 12 cM smaller than that from the other published map (Gyapay et al. 1994). Probability (*P*) was estimated using the χ^2 test. Maximum lod score (MLS) was estimated using Risch's equation (1990). PGS = percent of gene sharing; ns = not significant.

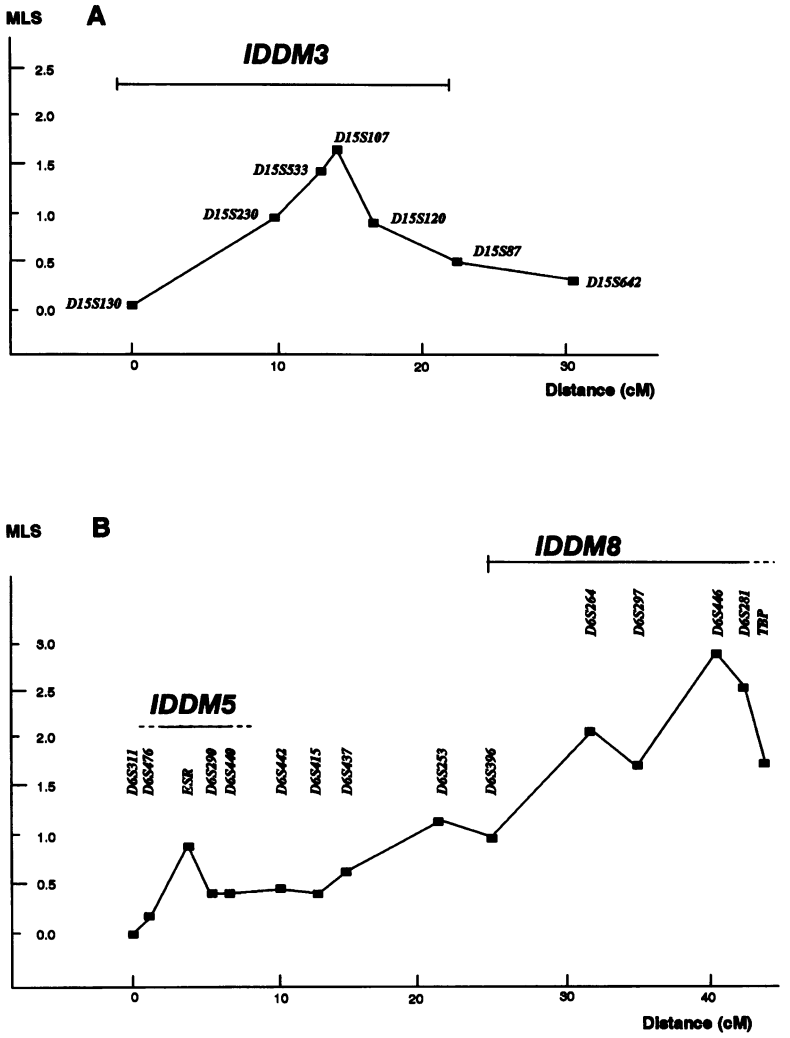

Figure 1 Schematic presentation of the locations of *IDDM3*, *IDDM5*, and *IDDM8*. *A*, MLS for markers on 15q26 in the combined data set (table 2), plotted against genetic distances of the markers. *B*, Plot, based on the data in the combined data set in table 3. A peak MLS was observed at *D6S446*. *ESR* had only slightly higher MLS than its flanking markers. *IDDM8* is probably located in the interval after *D6S396*, since this marker did not show significant evidence of linkage.

and can be interpreted to support the same hypothesis. *IDDM8* may lie in the region after *D6S396*, since this marker did not show evidence of linkage. The telomeric limit of the disease gene could not be determined by using this data set.

Analyses of IDDM3 on Chromosome 15q26 and IDDM4 on Chromosome 11q13

To define the location of *IDDM3* on 15q26, we have analyzed seven microsatellite markers flanking *D15S107* in all 104 sib pairs (table 2 and fig. 1). Haplotyping with flanking markers was performed to increase the informativeness of markers. *D15S107* showed the strongest linkage evidence. Three other markers (*D15S130*, *D15S87*, and *D15S642*) showed no evidence for linkage, even at $P = .05$. These results suggest that *IDDM3* is most likely located in the region surrounding *D15S107* and probably within the 22-cM interval between *D15S130* and *D15S87*.

We analyzed all 104 families with seven microsatellite markers surrounding the *FGF3* locus in order to narrow the interval containing *IDDM4*. Four markers (*D11S1357*, *D11S913*, *D11S1337*, and *D11S901*) have not been studied in previous reports. *D11S1337* had the most significant *P* value (.0041) among all seven markers (table 4). The proportions of gene sharing were lower for markers on both flanking sides of *D11S1337*, suggesting that *IDDM4* may be located in the surrounding region of *D11S1337*.

Discussion

Mapping of genes predisposing to complex disorders such as IDDM is a difficult task. Suarez et al. (1994)

Table 4

Analyses of *IDDM4* on Chromosome 11q

Marker	Distance (cM)	1 IBD	0 IBD	PGS	*P*	MLS
D11S1357	0	89	66	57.4	ns	.7
D11S913	7	104	71	59.4	.013	1.4
D11S1337	8	117	77	60.3	.0041	1.8
FGF3	13	114	84	57.5	.033	1.0
D11S1314	16	119	81	59.5	.0072	1.6
D11S916	18	119	81	59.5	.0072	1.6
D11S901	30	99	68	59.2	.020	1.3

NOTE.—PGS = percent of gene sharing; MLS = maximum lod score; and ns = not significant.

have shown by computer simulation that, if a number of loci, each with a moderately small effect on a disease, are implicated, then linkage will be difficult to detect and replicate, because of heterogeneities expected between data sets within, as well as between, studies. As shown in our own studies and in those previously published (Davies et al. 1994; Field et al. 1994; Hashimoto et al. 1994), the generally accepted lod score of 3 ($P < .001$) for detecting linkage in monogenic diseases could easily miss a signal in a complex disease, even with $\geq$100 ASPs. Therefore, we have reported any linkage evidence with $P < .05$ for both confirmatory and exploratory reasons. In light of the large number of markers tested, such evidence is by itself not strong. It has been suggested that a P value of 3×10^{-5} (MLS = 3.6) is required to confirm a true linkage in genome-wide ASP analyses when parents are also typed (Lander and Schork 1994). Such criteria may be difficult to achieve in complex diseases such as IDDM because of heterogeneity that arise when different data sets are pooled. Other authors have sought a P value of .001 in one data set and then attempted to obtain further evidence ($P < .05$) in other independent, additional data sets (Davies et al. 1994). The rate of false-positive linkage is unknown if such criteria are used; however, we believe that less stringent criteria should be applied for the initial establishment of linkage for complex diseases and that more stringent criteria should then be applied to definitively confirm the existence of disease genes as suggested by Lander and Schork (1994).

Linkage evidence for *IDDM8* in our American Caucasian families (MLS = 2.8 for *D6S446*, and MLS = 2.0 for *D6S264*) and the weaker evidence in the U.K. 96 data set (MLS = 1.4 for *D6S264*) together confirm the presence of a disease locus in the region, using the criteria of Davies et al. (1994). In the combined data set of U.K. and our U.S. families, linkage evidence for *D6S264* almost reached the stringent criteria (MLS = 3.6) suggested by Lander and Schork (1994). Since *D6S264* is 28 cM more telomeric than *ESR* (*IDDM5*), our results suggest that there are probably two distinct IDDM genes on 6q (*IDDM5* near *ESR* and *IDDM8* near *D6S264–D6S446*). This conclusion is also supported by the U.K. data set (Davies et al. 1994). It is worthy to note that the percentage of genes shared by ASPs for *D6S264* was 62.5% in the U.K. data set (Davies et al. 1994), which is very similar to that observed in our USA families (60.9%). *IDDM8* (λ_s = 1.8 for *D6S446*) can explain a higher proportion of the familial clustering of IDDM (i.e., higher λ_s value) than other non-HLA susceptibility genes in this data set (λ_s = 1.4, 1.6, 1.2, and 1.3 for *IDDM3, IDDM4, IDDM5,* and *IDDM7,* respectively). Recently, suggestive evidence indicated that an imprinted gene(s) on chromosome 6 may be involved in transient neonatal diabetes mellitus (TNDM) (James et al. 1995). This gene(s) appeared to be important for pancreatic β-cell development. It will be interesting to see whether this TNDM gene is identical or related to *IDDM8* or *IDDM5,* located on 6q.

By the criteria of Davies et al. (1994), *IDDM3* can now be confirmed as a true linkage. The P value for linkage of *IDDM3* was .001 in the combined data set of three independent subsets (Field et al. 1994), and we have now obtained linkage evidence in a fourth independent data set ($P < .02$). Analyses using additional markers suggested that *D15S107* showed the strongest evidence of linkage in this region. The percentage of gene sharing at *D15S107* in our combined data set (59.7%) was slightly higher than those observed by Field et al. (1994) (53.5%–58.3%). The difference may be due to random variation. However, the percentage of gene sharing at *D15S107* in our Florida data set (62.7%) appeared to be higher than that for our HBDI data set (57.9%), which is not very different from those reported by Field et al. (1994).

Support for *IDDM4* was obtained in the U.K. 96, U.K. 102, 25 Canadian, 133 French, 22 North African, at least 125 other American families (with overlap between studies). Our data from the HBDI families should not be considered as independent confirmation of link-

age, because of overlap of families with previous reports. Our Florida families did not provide additional evidence of linkage for *IDDM4*. There was no significant evidence of linkage for *IDDM5* from our Florida families. These results suggested that there were strong heterogeneities from one data set to another. Some of these heterogeneities may be due to real genetic differences, and some are just due to random chance.

Our analyses also provided additional, independent evidence for *IDDM7* previously mapped to 2q (Davies et al. 1994; Copeman et al. 1995; Owerbach and Gabbay 1995). The families from Florida gave a *P* value of .02 and thus further strengthened linkage evidence for *IDDM7*. Our results provided additional, independent evidence for two other regions (*D3S1303* and *D7S486*); however, the combined evidence in our data set and that from the U.K. data set (Davies et al. 1994) was not strong enough to confirm linkage definitively in these two regions. Our linkage evidence for *D4S1566* was novel and warrants further study in other independent families. Linkage evidence for eight other markers previously documented in individual data sets was observed in neither the HBDI nor the Florida data sets of this study. Some of these linkages may be true and others may be false; however, these regions should still be analyzed in additional families.

IDDM is a complex disease, and it appears from this and other recent studies that genes in five (*IDDM3, IDDM4, IDDM5, IDDM7,* and *IDDM8*) and perhaps more regions, in addition to *IDDM1* and *IDDM2,* are involved in its inherited predisposition. Well-designed association (linkage disequilibrium) analyses using diabetic families, as well as ethnically matched patient and control populations, will be important for the identification of these genes. Their elucidation will be of great help for the prediction of risk for IDDM and for the design of future intervention strategies to prevent the onset of clinical disease.

Acknowledgments

This work was supported by Juvenile Diabetes Foundation grant JDF193182 (to J.-X.S.) and National Institute of Health grants HD19469 (to N.K.M.) and HD12731 (to G.T.). We are grateful to Dr. Richard S. Spielman of the University of Pennsylvania School of Medicine for providing valuable comments on the manuscript and DNA samples from eight IDDM families. We also thank Dr. James Weber of the Marshfield Medical Research Foundation for his help on microsatellite typing techniques. We are also grateful to an anonymous reviewer for providing valuable comments and suggestions on the manuscript.

References

Bain SC, Prins JB, Hearne CM, Rodrigues NR, Rowe BR, Pritchard LE, Ritchie RJ, et al (1992) Insulin gene region-encoded susceptibility to type 1 diabetes is not restricted to HLA-DR4-positive individuals. Nat Genet 2:212–215

Bennett ST, Lucassen AM, Gough SCL, Powell EE, Undlien DE, Pritchard LE, Merriman ME, et al (1995) Susceptibility to human type 1 diabetes at IDDM2 is determined by tandem repeat variation at the insulin gene minisatellite locus. Nat Genet 9:284–292

Caillat-Zucman S, Bertin E, Timsit J, Boitard C, Assan R, Bach JF (1993) Protection from insulin-dependent diabetes mellitus is linked to a peptide transporter gene. Eur J Immunol 23:1784–1788

Copeman JB, Cucca F, Hearne CM, Cornall RJ, Reed PW, Ronningen KS, Undlien DE, et al (1995) Linkage disequilibrium mapping of a type 1 diabetes susceptibility gene (IDDM7) to chromosome 2q31-q33. Nat Genet 9:80–85

Davies JL, Yoshihiko K, Bennett S, Copeman JB, Cordell HJ, Pritchard LE, Reed PW, et al (1994) A genome-wide search for human type 1 diabetes susceptibility genes. Nature 371:130–136

Deng GY, Muir A, Maclaren NK, She J-X (1995) Association of *LMP2* and *LMP7* genes within the major histocompatibility complex with insulin-dependent diabetes mellitus: population and family studies. Am J Hum Genet 56:528–534

Dunsworth TS, Rich SS, Swanson J, Barbosa J (1982) No evidence for linkage between diabetes and the kidd marker. Diabetes 31:991–993

Field LL, Tobias R, Magnus T (1994) A locus on chromosome 15q26 (IDDM3) produces susceptibility to insulin-dependent diabetes mellitus. Nat Genet 8:189–194

Ghosh S, Palmer SM, Rodrigues NR, Cordell HJ, Hearne CM, Cornall RJ, Prins JB, et al (1993) Polygenic control of autoimmune diabetes in nonobese diabetic mice. Nat Genet 4:404–409

Gyapay G, Morissette J, Vignal A, Dib C, Fizames C, Millasseau P, Marc S, et al (1994) The 1993–94 Généthon human genetic linkage map. Nat Genet 7:246–339

Hashimoto L, Habita C, Beressi JP, Delepine M, Besse C, Cambon-Thomsen A, Deschamps I, et al (1994) Genetic mapping of a susceptibility locus for insulin-dependent diabetes mellitus on chromosome 11q. Nature 371:161–164

Hodge SE, Anderson CE, Neiswanger K, Field LL, Spence MA, Sparkes RS, Sparkes MC, et al (1981) Close genetic linkage between diabetes mellitus and kidd blood group. Lancet 2:893–895

Hodge SE, Anderson CE, Neiswanger K, Sparkes RS, Rimoin DL (1983) The search for heterogeneity in insulin-dependent diabetes mellitus (IDDM): linkage studies, two-locus models, and genetic heterogeneity. Am J Hum Genet 35:1139-1155

James RS, Crolla JA, Sitch FL, Jacobs PA, Howell WM, Betts P, Baum JD, et al (1995) An imprinted gene(s) for diabetes. Nat Genet 9:110-112

Julier C, Hyer RN, Davies J, Merlin F, Soularue P, Briant L, Cathelineau G, et al (1991) Insulin-IGF2 region on chromosome 11p encodes a gene implicated in HLA-DR4-dependent diabetes susceptibility. Nature 354:155–158

Lander ES, Schork NJ (1994) Genetic dissection of complex traits. Science 265:2307–2048

Lucassen AM, Julier C, Beressi JP, Boitard C, Froguel P, Lathrop

M, Bell JI (1993) Susceptibility to insulin dependent diabetes mellitus maps to a 4.1 kb segment of DNA spanning the insulin gene and associated VNTR. Nat Genet 4:305–310

Owerbach D, Gabbay KH (1994) Linkage of the VNTR/insulin-gene and type I diabetes mellitus: increased gene sharing in affected sibling pairs. Am J Hum Genet 54:909–912

——— (1995) The HOXD8 locus (2q31) is linked to type I diabetes: interaction with chromosome 6 and 11 disease susceptibility genes. Diabetes 44:132–136

Risch N (1987) Assessing the role of HLA-linked and unlinked determinants of disease. Am J Hum Genet 40:1–14

——— (1990) Linkage strategies for genetically complex traits. II. The power of affected relative pairs. Am J Hum Genet 46:229–241

She J-X, Bui MM, Tian XH, Muir A, Wakeland EK, Zorovich B, Zhang LP, et al (1994) Additive susceptibility to insulin-dependent diabetes conferred by HLA-DQB1 and insulin genes. Autoimmunity 18:195–203

Suarez BK, Hampe CL Van Eerdewegh P (1994) Problems of replicating linkage claims in psychiatry. In: Gershon ES, Cloninger CR (eds) Genetic approaches to mental disorders. American Psychiatric Association, Washington DC, pp 23–46

Thorsby E, Ronningen KS (1992) Role of HLA genes in predisposition to develop insulin-dependent diabetes mellitus. Ann Med 24:523–531

Todd JA, Aitman TJ, Cornall RJ, Ghosh S, Hall JRS, Hearne CM, Knight AM, et al (1991) Genetic analysis of autoimmune type 1 diabetes mellitus in mice. Nature 351:542–546

Todd JA, Bell JI, McDevitt HO (1987) HLA-DQ gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. Nature 329:599–604