Am. J. Hum. Genet. 63:547–556, 1998

Fine Mapping of the Diabetes-Susceptibility Locus, *IDDM4,* on Chromosome 11q13

Yusuke Nakagawa,[1,*] Yoshihiko Kawaguchi,[1,*] Rebecca C. J. Twells,[1] Claire Muxworthy,[1] Kara M. D. Hunter,[1] Amanda Wilson,[1] Marilyn E. Merriman,[1] Roger D. Cox,[1] Tony Merriman,[1] Francesco Cucca,[1] Patricia A. McKinney,[2] Julian P. H. Shield,[3] Jaakko Tuomilehto,[5] Eva Tuomilehto-Wolf,[5] Constantin Ionesco-Tirgoviste,[6] Lorenza Nisticò,[7] Raffaella Buzzetti,[8] Paolo Pozzilli,[9] San-Raffaele Family Study,[10] Geir Joner,[11] Eric Thorsby,[12] Dag E. Undlien,[12] Flemming Pociot,[14] Jörn Nerup,[14] Kjersti S. Rönningen,[13] Bart's-Oxford Family Study Group,[4] Stephen C. Bain,[15] and John A. Todd[1]

[1]The Wellcome Trust Centre for Human Genetics, Nuffield Department of Surgery, University of Oxford, Oxford; [2]Paediatric Epidemiology Group, Research School of Medicine, University of Leeds, Leeds; [3]Institute of Child Health, University of Bristol, Royal Hospital for Sick Children, and [4]Department of Diabetes and Metabolism, University of Bristol, Medical School Unit, South Meads Hospital, Bristol; [5]Diabetes and Genetic Epidemiology Unit, National Public Health Institute, Helsinki; [6]Clinic of Nutrition and Metabolic Disease, Bucharest; [7]Istituto Biologia Cellulare CNR, Monterotondo, [8]Endocrinologia, Istituto Clinica Medica II, University of Roma 'La Sapienza,' and [9]Libero Istituto Universitario Campus Biomedico, Rome; [10]Istituto Scientifico San Raffaele, University of Milan, Milan; [11]Aker Diabetes Research Centre, Aker University Hospital, [12]Institute of Transplantation Immunology, The National Hospital, and [13]Department of Population Health Sciences, National Institute of Public Health, Oslo; [14]Steno Diabetes Center, Gentofte, Denmark; and [15]Department of Medicine, University of Birmingham, Birmingham Heartlands Hospital, Birmingham, United Kingdom

Summary

Genomewide linkage studies of type 1 diabetes (or insulin-dependent diabetes mellitus [IDDM]) indicate that several unlinked susceptibility loci can explain the clustering of the disease in families. One such locus has been mapped to chromosome 11q13 (*IDDM4*). In the present report we have analyzed 707 affected sib pairs, obtaining a peak multipoint maximum LOD score (MLS) of 2.7 ($\lambda_s = 1.09$) with linkage (MLS $\geqslant$ 0.7) extending over a 15-cM region. The problem is, therefore, to fine map the locus to permit structural analysis of positional candidate genes. In a two-stage approach, we first scanned the 15-cM linked region for increased or decreased transmission, from heterozygous parents to affected siblings in 340 families, of the three most common alleles of each of 12 microsatellite loci. One of the 36 alleles showed decreased transmission (50% expected, 45.1% observed [$P = .02$, corrected $P = .72$]) at marker *D11S1917*. Analysis of an additional 1,702 families provided further support for negative transmission (48%) of *D11S1917* allele 3 to affected offspring and positive transmission (55%) to unaffected siblings (test of heterogeneity $P = 3 \times 10^{-4}$, corrected $P = .01$]). A second polymorphic marker, H0570polyA, was isolated from a cosmid clone containing *D11S1917*, and genotyping of 2,042 families revealed strong linkage disequilibrium between the two markers (15 kb apart), with a specific haplotype, D11S1917*03-H0570polyA*02, showing decreased transmission (46.4%) to affected offspring and increased transmission (56.6%) to unaffected siblings (test of heterogeneity $P = 1.5 \times 10^{-6}$, corrected $P = 4.3 \times 10^{-4}$). These results not only provide sufficient justification for analysis of the gene content of the *D11S1917* region for positional candidates but also show that, in the mapping of genes for common multifactorial diseases, analysis of both affected and unaffected siblings is of value and that both predisposing and nonpredisposing alleles should be anticipated.

Received March 11, 1998; accepted for publication June 12, 1998; electronically published July 17, 1998.

Address for correspondence and reprints: Prof. John A. Todd, The Wellcome Trust Centre for Human Genetics, Windmill Road, Headington, Oxford, OX3 7BN, United Kingdom. E-mail: john.todd@well.ox.ac.uk

* These authors contributed equally to this work.

0002-9297/98/6302-0031$02.00

Introduction

Type 1 diabetes is a common multifactorial disease resulting from an interaction of many genes and, probably, of many environmental factors, resulting in the specific immune-mediated destruction of the insulin-producing cells of the pancreas and in life-long insulin deficiency (Tisch and McDevitt 1996). Functional candidate-gene analyses based on case-control association studies and, later, on family-based studies have led to the identification of two *IDDM* loci: (1) *IDDM1* on chromosome 6p21, which, in part, most likely corresponds to func-

tional amino acid variation in the peptide-binding sites of the T lymphocyte antigen–recognition molecules MHC HLA-DR and -DQ (Cucca and Todd 1996; She 1996; Thorsby and Undlien 1996), and (2) *IDDM2* on chromosome 11p15, corresponding to polymorphism of a VNTR locus in the promoter of the insulin gene (*INS*), which affects transcription of *INS* (Bennett et al. 1995; Vafiadis et al. 1997). However, until recently (Nisticò et al. 1996), candidate-gene studies have provided little additional insight into inheritance of type 1 diabetes, since 1984, when the association of the *INS* VNTR with the disease was first discovered (Bell et al. 1984). With improvements in technology and genetic maps (Reed et al. 1994), encouraging gene-mapping results from the spontaneous mouse model of type 1 diabetes (Todd et al. 1991; Wicker et al. 1994) and from collections of large numbers of multiplex families (Bain et al. 1990; Lernmark et al. 1990), it became possible to embark on systematic searches of the whole human genome, for chromosome regions showing evidence of linkage to disease. The first scans showed that *IDDM1*/MHC on chromosome 6p21 was the major locus, contributing $\leqslant 50\%$ of the familial clustering of the disease ($\lambda_s = 3$, the ratio of the expected proportion of affected sib pairs sharing zero alleles identical by descent [IBD], .25, and the observed proportion) (Davies et al. 1994; Hashimoto et al. 1994), with a more modest contribution, <10%, from *IDDM2* ($\lambda_s = 1.25$) (Davies et al. 1994). These studies and a candidate-gene study (Field et al. 1994) also provided some positive evidence for other loci, particularly a locus on chromosome 11q13, designated "*IDDM4*." On the basis of evidence of linkage in several affected-sib-pair data sets studied in at least four independent laboratories, support for the existence of *IDDM4* has been extended, to $P = 1.5 \times 10^{-6}$ (Luo et al. 1996), and it now appears highly likely that a type 1 diabetes gene(s) is encoded by chromosome 11q13.

The problem now is to fine map loci such as *IDDM4* that have modest effects ($\lambda_s < 1.3$) within broad regions of linkage. We have advocated (Copeman et al. 1995) using the transmission/disequilibrium test (TDT) (Spielman et al. 1993) to search for evidence of association in regions for which there is prior evidence of linkage, because it should provide a finer localization of the disease locus and is a more powerful way of detecting effects, owing to the selected analysis of parents heterozygous for test alleles. Families with only one affected sibling can also be used in TDT. However, association mapping is problematic because the association of a chromosome region is dependent on the unknown distribution of the alleles of the markers on predisposing and nonpredisposing chromosomes. We have begun to investigate the application of association mapping in the common, multifactorial disease type 1 diabetes, using data from chromosome 18q21 (the putative locus *IDDM6*) (Merriman et al. 1997, 1998) and, in the present study, data from chromosome 11q13/*IDDM4*. We report evidence of association of the *D11S1917* region with type 1 diabetes. This result has led to the identification of a new member of the LDL-receptor family, close to *D11S1917,* which is a functional candidate gene for type 1 diabetes (Hey et al., in press). Moreover, we demonstrate both the utility of analyzing both affected and unaffected offspring and the importance of considering both positively and negatively transmitted marker alleles and haplotypes.

Subject and Methods

Diabetic Families

All families in this study were Caucasian, and, in each family, at least one affected sibling and both parents were included (table 1). The U.K. data set consisted of 401 multiplex families, 80 simplex families from the Yorkshire region, and 32 simplex families from the south-

Table 1

Sources and Numbers of Multiplex and Simplex Families, according to Stage of Analysis

Analysis and Family Data Set	No. of Families
Initial linkage study:	
U.K. multiplex	236
U.S. multiplex	104
Total	340
Follow-up linkage study:	
U.K. multiplex	165
U.S. multiplex	133
Norwegian multiplex	31
Continental Italy multiplex	38
Total	367
Initial scan for association:	
U.K. multiplex	236
U.S. multiplex	104
Total	340
Follow-up association study:	
U.K. multiplex	165
U.S. multiplex	133
Norwegian multiplex	31
Continental Italy multiplex	38
U.K. simplex:	
Yorkshire	80
Southwest	32
Age <5 years at diagnosis	56
St. Bart's-Oxford Family Study	24
Sardinian:	
Simplex	175
Multiplex	6
Norwegian simplex	375
Continental Italian simplex	62
Finnish simplex	216
Romanian simplex	204
Danish simplex	105
Total	1,702

western region (all three groups have been described elsewhere [Merriman et al. 1997, 1998], 56 simplex families in which all cases had been diagnosed at age <5 years (Wadsworth et al. 1995), and 24 simplex families from the Bart's-Oxford Family Study/Oxford Regional Prospective Study with cases that had been diagnosed at age <21 years (Gardner et al. 1997). The 237 U.S. affected-sib-pair families were obtained from the Human Biological Database Interchange (Lernmark et al. 1990), and each had at least one affected sibling that had been diagnosed at age <29 years. The 181 Sardinian families (175 simplex and 6 multiplex), 406 Norwegian families (375 simplex and 31 multiplex), and the 100 continental Italian families (62 simplex families and 38 multiplex families) have been described elsewhere (Merriman et al. 1997, 1998). The Finnish data set comprised 216 simplex families in which all cases had been diagnosed at age <15 years (Tuomilehto et al. 1992), and the Romanian data set comprised 204 simplex families in which all cases had been diagnosed at age <30 years. The 105 Danish simplex families have been described elsewhere (Pociot et al. 1993). Unaffected siblings were collected when it was possible to do so.

Physical Mapping and Genotyping

Cosmid H0570 was isolated from a gridded flow-sorted chromosome 11–specific library (Nizetic et al. 1994). We attempted to isolate dinucleotide repeats from this clone, according to the method described elsewhere (Copeman et al. 1995). A clone that seemed to be positively hybridized with a $(CA)_n$ dinucleotide probe was sequenced. This clone did not contain dinucleotide repeats but, instead, contained a mononucleotide $(A)_n$ repeat for which PCR primers were designed: H0570polyA forward (5′- TTT CCT CTC TGG GAG TCT CT-3′) and reverse (5′-GGA CAG TCA GTT ATT GAA ATG-3′). Intermarker distance was elucidated by standard restriction enzyme–mapping techniques, and the orientation was determined by examination of the genotypes of multiplex families in which a recombination had occurred within the contig of cosmid clones, including H0570, analyzed elsewhere (Courseaux et al. 1997). Genotyping PCRs using fluorescently labeled primers were performed and analyzed as described elsewhere (Reed et al. 1994).

Analysis of Linkage and Allelic Association

Multipoint LOD score (MLS) values were calculated by the MAPMAKER/SIBS program (Kruglyak and Lander 1995). The *P* values assigned to MLSs were theoretical (Holmans 1993). Transmission, from heterozygous parents to both affected and unaffected offspring, of single microsatellite marker alleles and of two marker haplotypes was assessed by TDT (Spielman et al. 1993), and statistical support for allelic association was determined by the T_{sp} statistic, in which all affected siblings are included in the analysis (Martin et al. 1997). The extent of linkage disequilibrium of an allele or haplotype with disease was quantitated in terms of percentage of transmission, which is the number of times that an allele is transmitted from heterozygous parents to affected and unaffected children, divided by the total number of transmissions, expressed as a percentage. Haplotypes could not be constructed definitely in 161 families, because of F_1 intercross status. In these families, the most likely haplotypes were determined on the basis of tight linkage disequilibrium between the two markers, *D11S1917* and H0570polyA. *D′* values were calculated as described elsewhere (Devlin and Risch 1995).

Results

Linkage Mapping

All the multiplex type 1 diabetic families available in our previous study (Davies et al. 1994) (n = 340 [236 U.K. families and 104 U.S. families]) have been genotyped for 18 microsatellite markers in 25 cM of chromosome 11q13 (fig. 1*A*). Peak evidence of linkage by multipoint analysis was obtained at the marker *D11S1883* (MLS = 1.26; P = .01; λ_s = 1.09). A further 367 multiplex families subsequently became available, and multipoint linkage of chromosome 11q13 was evaluated by use of all 18 markers (fig. 1*B*), providing additional support for *IDDM4*, with peak linkage at *D11S1337*, in the total of 707 families (MLS = 2.7; P = .0003; λ_s = 1.09) (fig. 1*C*). In our original study, we had detected *IDDM4* by conditioning the linkage at the chromosome 11q13 marker locus *FGF3* by means of the allele-sharing status at the *IDDM1*/MHC locus in 282 families (Davies et al. 1994). In the 640 families (of 707) for which *IDDM1* typing was available, peak linkage was at *D11S1337* (MLS = 2.9), and most of the support for linkage still came from the families (n = 294) in which sib pairs shared one or zero *IDDM1*/MHC haplotypes IBD (peak at *FGF3*, 3 cM distal to *D11S1337* [MLS = 2.6]), compared with the remaining families (n = 346), in which sib pairs shared two alleles IBD at *IDDM1* (at *FGF3* [MLS = .6]). In neither the present study nor the previous study (Davies et al. 1994) was there significant evidence for heterogeneity between categories ($P > .05$).

Given recent interest in discordant sib-pair linkage mapping (Risch and Zhang 1995), we evaluated evidence of linkage in discordant affected-unaffected pairs of siblings, compared with evidence from the same type of analysis in affected pairs (table 2). There was no statistically significant support for *IDDM4* in the discor-

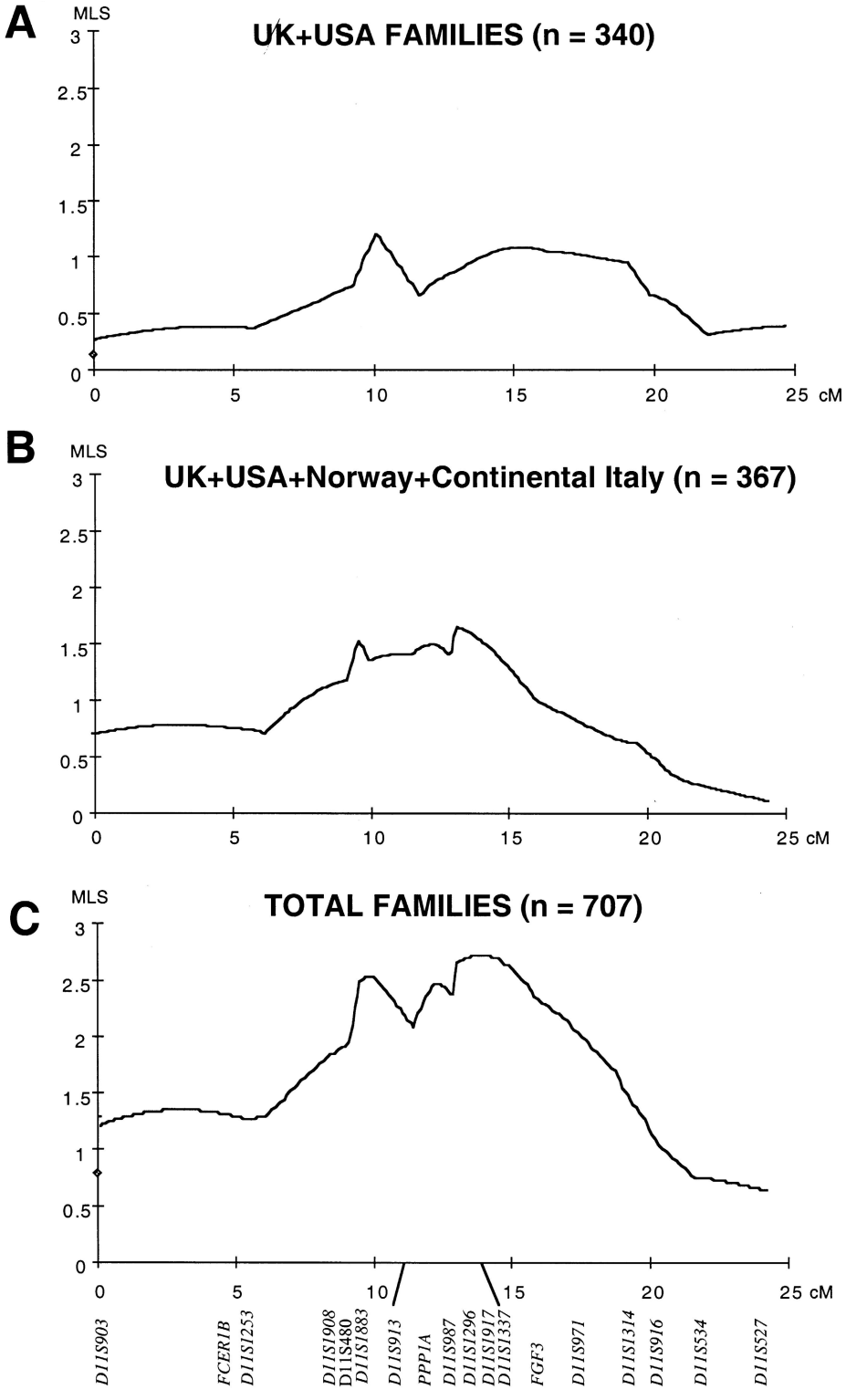

Figure 1 Multipoint linkage analysis of chromosome 11q13 in type 1 diabetic affected-sib-pair families

Table 2

Linkage of Chromosome 11q13 to Type 1 Diabetes in Discordant and Concordant Sib Pairs

Data Set	No. (%) of Sib Pairs Sharing or Not Sharing Allele IBD[a]		P[b]
	One Allele	No Alleles	
Unaffected-affected pairs:			
United Kingdom	46 (47.42)	51	
United States	32 (48.5)	34	
Sardinia	129 (45.9)	152	
U.K. simplex	64 (46.7)	73	
Romania	151 (57.0)	114	.02
Finland	64 (45.7)	76	
Denmark	52 (50.5)	51	
Italy	46 (50)	50	
Total	584 (49.3)	601	
Affected pairs:			
United Kingdom	391 (52.0)	361	
United States	263 (56.0)	207	.01
Norway	28 (54.9)	23	
Italy	62 (60.1)	40	.03
Total	744 (54.1)	631	.002

[a] Data were obtained by combining of results from both marker *D11S1917* and H0570polyA.

[b] Calculated from χ^2 test of allele sharing, against an expected ratio of 1:1 (only values $\leq .05$ are shown).

dant pairs, although six of eight data sets showed that IBD sharing of alleles was <50%.

TDT of the D11S1917 *Region*

The three most common alleles of 12 of these markers (*D11S1908, D11S480, D11S1883, D11S913, PPP1A, D11S987, D11S1296, D11S1917, D11S1337, FGF3, D11S971,* and *D11S1314;* each with maximum MLS $\geq .7$) were analyzed by TDT in 340 families (236 U.K. families and 104 U.S. families). Of the 36 alleles tested, only allele 3 of *D11S1917* showed transmission from heterozygous parents to affected children that was different from the expected 50%: there were 247 (45.1%) cases of transmission and 301 cases of no transmission ($P = .02$, corrected $P = .72$).

The transmission of allele 3 of *D11S1917* then was analyzed in all 2,042 families, extending support to $P = 3 \times 10^{-4}$ in the comparison of transmission to affected offspring versus transmission to unaffected offspring (table 3; 48% and 55% transmission, respectively). We derived a second polymorphic marker from a *D11S1917*-positive cosmid, H0570 (Courseaux et al. 1997), a mononucleotide microsatellite repeat, $(A)_n$, named "H0570polyA." Allele 2 of the H0570polyA locus showed strong linkage disequilibrium with allele 3 of *D11S1917* ($D' = .94$; $P \leq 1 \times 10^{-100}$); hence the transmission of allele 2 of H0570polyA was evaluated specifically in the 2,042 families (table 4). Marker H0570polyA was more strongly associated with type 1 diabetes than was *D11S1917:* 46.7% transmission of allele 2 to affected siblings (TDT $P = .003$ and T_{sp} $P = .004$) and 54.7% transmission of allele 2 to unaffected siblings ($P = .004$), with a significant test of heterogeneity ($P = 4.8 \times 10^{-5}$). The D11S1917*03-H0570polyA*02 (3-2) haplotype also was associated with type 1 diabetes (T_{sp} $P = .002$). The transmission of the 3-2 haplotype also was significantly different between affected and unaffected siblings (heterogeneity $P = 1.5 \times 10^{-6}$; table 4). Correction of this P value by the number of loci and alleles tested ($n = 36$) and by the number of data sets analyzed ($n = 8$) gives $P = 5.4 \times 10^{-5}$ and $P = 4.3 \times 10^{-4}$, respectively.

Having obtained substantial support for an association of the *D11S1917*-H0570polyA region with type 1 diabetes and having identified a nonpredisposing haplotype, we evaluated the transmission of the other common haplotypes of these markers, in the expectation that one or more of them should have >50% transmission to affected siblings and perhaps <50% transmission to unaffected siblings (table 5). Five haplotypes represented 97.3% of all haplotypes in the families: 3-2 (26.6% frequency in all children of 2,012 families; $D' = .94$), 2-3 (28.6%; $D' = .12$; $P < 1.4 \times 10^{-20}$), 2-1 (20.7%; $D' = .61$; $P < 1 \times 10^{-100}$), 1-3 (20.1%; $D' = .6$; $P < 1 \times 10^{-100}$), and 3-3 (1.95%; $D' = -.63$; $P < 1 \times 10^{-100}$). Three of these haplotypes—2-3, 2-1, and 1-3—were slightly positively transmitted to affected siblings, at frequencies of 51.3%, 51.9%, and 51.6%, respectively, and were slightly negatively transmitted to unaffected siblings, at frequencies of 46.3% ($P = .04$), 49.0%, and 49.0%, respectively. This pattern of transmission appears to be compensatory for that of the 3-2 haplotype, which has the opposite pattern. The fourth haplotype, 3-3, although much rarer than the others, provided interesting results suggesting that it is being positively transmitted more often than are the other three haplotypes, with 62.4% ($P = .001$) transmission to affected siblings and 42.6% transmission to unaffected siblings (test of heterogeneity between transmission to affected and transmission to unaffected siblings, $P = .008$). The 62.4% transmission of the 3-3 haplotype was significantly different from the average 51.6% transmission of the 2-3, 2-1 and 1-3 haplotypes combined (χ^2 test of heterogeneity, $P = .006$).

Discussion

In the largest linkage study of the chromosome 11q13 region in type 1 diabetes to date, we have obtained in 707 affected-sib-pair families a peak MLS of 2.7 (marker *D11S1337; P* = .0003) and a peak λ_s of 1.09. This compares with the results of a previous study of 596 affected-sib-pair families, which produced MLS = 5.0 at marker

Table 3

Transmission of Allele 3 of *D11S1917* to Affected and Unaffected Offspring in 2,042 Type 1 Diabetic Families

	No. (%) of Families[a]					
	Affected Offspring		Unaffected Offspring			
	T	NT	T	NT	P[b]	Het P[c]
United Kingdom	373 (47.2)	417	75 (55.1)	61		
United States	174 (48.3)	186	33 (57.9)	24		
Norway	151 (48.2)	162	158 (53.6)	137		
Sardinia	62 (48.4)	66	55 (62.5)	33	.02	.04
Romania	93 (52.8)	83	67 (51.9)	62		
Finland	68 (46.6)	78	37 (49.3)	38		
Italy	51 (41.8)	71	46 (61.3)	29	.05	
Denmark	42 (52.5)	38	47 (54.7)	39		
Total	1,014 (47.9)	1,101	518 (55)	423	.002	.0003

[a] T = transmission; NT = no transmission; and Het P = heterogeneity P.

[b] From TDT using all affected sibs (only values $\leqslant .05$ are shown).

[c] From 2 × 2 contingency-table heterogeneity test of data for affected offspring versus data for unaffected offspring (only values $\leqslant .05$ are shown).

D11S1296 (100 kb centromeric of *D11S1337;* authors' unpublished data) (Luo et al. 1996). That study included data from 331 U.K., U.S., and French families reported in prior studies (Davies et al. 1994; Hashimoto et al. 1994), and these U.K. and U.S. families are included within the 707-family data set of the present study. For the *D11S1917* region in the 707 families analyzed here, the number of cases in which one and zero alleles were shared IBD were 744 (54.1%) and 631, respectively. If we add the corresponding results for the French families, 173 and 128 (Hashimoto et al. 1994), then the total MLS = 3.2 at 54.7% IBD sharing (by the formula $\text{MLS} = N_1[\log_{10}(N_1/0.5N)] + N_0[\log_{10}(N_0/0.5/N_0)]$, where N_1 and N_0 are the number of sib pairs sharing one and zero alleles IBD, respectively, and $N = N_1 + N_0$). These data, combined with the TDT results (tables 4 and 5), strongly support the existence of *IDDM4* within 11q13, but the effect at $\lambda_s = 1.09$ is modest, and the average odds ratio for the etiological, predisposing allele, when it is eventually identified, cannot be >1.5 in this total sample. Moreover, this assumes that the linked region contains only one disease locus, which is not what we are finding in genetic analysis of type 1 diabetes in inbred strains of mice (Podolin et al. 1997, 1998). It is noted, however, that in certain ethnically homogeneous populations, the λ_s value might be much greater, depending on the allele frequencies at *IDDM4* and at other interacting loci and on the effects of unknown environmental factors. Efforts must be made to collect very large numbers of families (or cases and genetically matched controls) from such homogeneous populations in which the effect of a particular locus is exaggerated, to permit detailed fine mapping and disease-gene identification. The magnitude of effect obtained here by use of a mixture of populations is, however, likely to be typical of susceptibility genes responsible for the development of type 1 diabetes and other common multifactorial diseases.

In our analyses, the initial detection of *IDDM4* was dependent on conditioning of the linkage data for chromosome 11q13 by the sharing status at *IDDM1*/MHC. Although there is no significant heterogeneity, the bias in linkage of chromosome 11q13 to disease when one or zero alleles are shared at *IDDM1*/MHC is still observed in 640 families. Our results suggest that, in the initial detection of potentially interesting chromosome regions, conditioning of marker data by the sharing status at other unlinked loci is a worthwhile strategy. We would not have continued to study the chromosome 11q13 region if we had not conditioned the data by *IDDM1*.

We have found evidence for a common haplotype for which the transmission to affected and unaffected siblings is <50% and >50%, respectively, suggesting that this haplotype contains an allele that is nonpredisposing for or even protective against type 1 diabetes. In the same way, the *IDDM1* MHC class II HLA-DQB1*0301 allele is negatively transmitted to affected siblings and is protective against type 1 diabetes (F. Cucca and J. A. Todd, unpublished data). In addition, a rare *INS* VNTR class III haplotype shows a distinctive pattern of <50% transmission to affected siblings (Bennett et al. 1995), indicating its protective role with regard to disease.

For these markers in these families, the most powerful statistic is a test of heterogeneity between the affected and unaffected data sets, yielding $P = 1.5 \times 10^{-6}$ in support of a difference in transmission of the protective 3-2 haplotype to affected versus unaffected offspring (table

Table 4

Transmission of H0570polyA Allele 2 and the D11S1917*03-H0570polyA*02 Haplotype to Affected and Unaffected Offspring in 2,042 Type 1 Diabetic Families

	No. (%) of Families												
	Allele 2							D11S1917*03-H0570polyA*02 Haplotype					
	Affected Offspring			Unaffected Offspring				Affected Offspring			Unaffected Offspring			
	T	NT	*P*	T	NT	*P*	Het *P*	T	NT	*P*	T	NT	*P*	Het *P*
United Kingdom	356 (45.8)	421	.02	78 (56.1)	61		.02	327 (45.9)	386	.05	75 (57.3)	56		.02
United States	178 (46.6)	204		34 (54.0)	29			162 (47.0)	183		31 (57.4)	23		
Norway	140 (48.6)	148		136 (51.7)	127			111 (47.8)	121		114 (54.0)	97		
Sardinia	63 (47.0)	71		60 (61.2)	38	.03	.03	52 (45.6)	62		54 (65.9)	28	.004	.005
Romania	95 (53.7)	82		66 (53.2)	58			83 (51.2)	79		60 (54.1)	51		
Finland	66 (45.2)	80		38 (49.4)	39			57 (44.5)	71		35 (50)	35		
Italy	39 (36.4)	68	.005	42 (60.9)	27			37 (37.4)	62		37 (60.7)	24		.004
Denmark	40 (48.8)	42		51 (56.7)	39			35 (49.3)	36		42 (58.3)	30		
Total	977 (46.7)	1,116	.003	505 (54.7)	418	.004	4.8×10^{-5}	864 (46.4)	1,000	.002	448 (56.6)	344	.0002	1.5×10^{-6}

Note.—See footnotes to table 3.

Table 5

Transmission of Haplotypes 2-3, 2-1, 1-3 and 3-3 of Markers *D11S1917* and H0570polyA in 2,042 Type 1 Diabetic Families

	No. (%) of Families																			
	United Kingdom		United States		Sardinia		Norway		Romania		Finland		Italy		Denmark		Total			Het
	T	NT	T	NT	T	NT	T	NT	T	NT	T	NT	T	NT	T	NT	T	NT	*P*	*P*
Affected:																				
2-3	376	348	187	176	65	71	123	105	65	69	69	63	37	45	37	35	959 (51.3)	912		
2-1	278	267	126	117	71	54	87	85	69	71	55	52	52	42	38	31	776 (51.9)	719		
1-3	276	235	146	138	33	37	106	105	40	50	75	74	43	25	24	32	743 (51.6)	696		
3-3	37	31	30	18	4	1	11	6	2	1	9	3	6	0	7	4	106 (62.4)	64	.001	
Unaffected:																				
2-3	58	68	20	38	48	52	104	123	43	46	37	40	23	25	40	40	373 (46.3)	432	.04	.02
2-1	52	47	21	23	43	52	85	80	49	47	29	31	15	25	34	36	328 (49.0)	341		
1-3	44	55	18	16	25	27	110	107	30	34	45	36	8	9	28	36	308 (49.0)	320		
3-3	3	6	5	5	1	4	7	6	1	2	3	3	0	1	6	8	26 (42.6)	35		.008

Note.—See footnotes to table 3.

4). At this stage of the analysis, it appears that the other common haplotypes show close to 50%, or "neutral," transmission. Interestingly, there is evidence for a rare (frequency 2%–3%) haplotype (3-3) that may be positively transmitted or predisposing (table 5).

We also note that the TDT *P* values for the transmission of the disease-associated alleles and haplotypes are almost identical to the *P* values for the modified TDT statistic, T_{sp} (Martin et al. 1997), which takes into account the presence of increased allele sharing in sib pairs and which allows the data from the second sib to be included, thereby giving a completely valid and powerful test of association. For example, the 3-2 haplotype TDT *P* value for 2,042 families was .002, and the T_{sp} *P* was .002. Hence, even though the 2,042 families include 707 affected-sib-pair families, which show evidence of allele sharing and linkage (fig. 1), TDT of all sibs in these families gives results identical or nearly identical to those from a valid test of allelic association or linkage disequilibrium, such as the T_{sp} statistic. This is because the degree of increased allele sharing in the 707 unrelated affected sib pairs is very modest and does not introduce any significant bias into the TDT of all siblings. If our study had used a few large multigeneration families and if there had been pronounced allele sharing, then the T_{sp} *P* values would be expected to be much larger than those from the TDT using all sibs in the analysis. We recommend, in the analysis of numerous unrelated, affected-sib-pair families, the use of a T_{sp}-like test as a valid test of association, rather than the use of TDT and one affected sibling per family, which is less powerful. Having shown, by means of the T_{sp} statistic, evidence of true association, we find that it is both (1) still convenient to calculate TDT values and (2) more powerful, in our current data set, to use TDT data in tests of heterogeneity, between affected and unaffected offspring, of transmission frequencies that are clearly underpinned by the association of the region with disease.

We now have to extend the physical and genetic maps flanking the *D11S1917*-H0570polyA region, to determine both how much of the chromosome is associated with type 1 diabetes and whether this is the only region, under the linkage curve, that shows association with the disease. It is conceivable that the evidence of linkage is due to more than one susceptibility locus, as we have found to be the case in the NOD mouse and its congenic derivatives: the *Idd3* locus on chromosome 3, originally defined as one peak of linkage (Todd et al. 1991), is now known to comprise four separate loci (*Idd3, Idd10, Idd17,* and *Idd18*), all within a 30-cM region of chromosome 3 (Podolin et al. 1997, 1998). Therefore, the entire 15-cM region of linkage on 11q13 must be scanned more comprehensively by analysis of more markers in more families, in a search for other regions of potential association. The results presented here and elsewhere (Merriman et al. 1997, 1998) clearly show that moderately polymorphic microsatellites are useful markers for defining the association of a chromosome region with disease, even for disease chromosomes that are common throughout Europe and that are, therefore, presumably ancient.

Given both the necessary scale of these studies and the likelihood of complex association-mapping data, it is essential to evaluate the disease association of functional candidate genes positioned in the region of association. We have shown that the H0570polyA locus is only 3 kb 5′ of a novel gene, designated "*LRP5*." *LRP5* is a member of the LDL-receptor gene family and is both a positional and functional candidate gene for type 1 diabetes (Hey et al., in press). Future experiments will include identification of polymorphisms in or near this gene and of others in the region that have effects

on this gene's structure or expression and that could account for the transmission patterns of the haplotypes that we have reported here.

Acknowledgments

We thank the Wellcome Trust, the British Diabetic Association, the Medical Research Council, the Juvenile Diabetes Foundation, the National Institutes of Health (grant DK 37957), the Novo Nordisk Foundation, the Norwegian Diabetes Association, and the Italian Telethon (grant E400) for support; and we thank our colleagues, including J. Carr-Smith, B. Rowe, A. Barnett, C. Smyth and E. Wadsworth, for help and advice. Type 1 diabetic families were gratefully received from the British Diabetic Association, the Norwegian Study Group for Childhood Diabetes, the San-Raffaele Family Study (E. Bosi, M. Rocco Pastore, V. Lampasona, A. Sergi, E. Capiello, G. Vitali, L. Esposito, and M. Ferrari), the Danish Study Group for Diabetes in Childhood, and the Human Biological Data Interchange.

References

Bain SC, Todd JA, Barnett AH (1990) The British Diabetic Association—Warren Repository. Autoimmunity 7:83–85

Bell GI, Horita S, Karam JH (1984) A polymorphic locus near the human insulin gene is associated with insulin-dependent diabetes mellitus. Diabetes 33:176–183

Bennett ST, Lucassen AM, Gough SL, 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

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

Courseaux A, Szepetowski P, Fernandes M, Serizet C, Kawaguchi Y, Grosgeorge J, Perucca-Lostanlen D, et al (1997) Framework YAC contig anchored into a 3.2-Mb high-resolution physical map in proximal 11q13. Genomics 40: 13–23

Cucca F, Todd, JA (1996) HLA susceptibility to type 1 diabetes: methods and mechanisms. In: Browning M, McMichael AJ (eds) HLA/MHC: genes, molecules and function. BIOS Scientific, Oxford, pp 383–406.

Davies JL, Kawaguchi Y, Bennett ST, 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

Devlin B , Risch N (1995) A comparison of linkage disequilibrium measures for fine-scale mapping. Genomics 29: 311–312

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

Gardner SG, Bingley PJ, Sawtell PA, Weeks S, Gale EAM, Barts-Oxford Study Group (1997) Rising incidence of insulin dependent diabetes in children aged under 5 years in the Oxford region: time trend analysis. BMJ 315:713–717

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

Hey PJ, Twells RCJ, Phillips MS, Nakagawa Y, Brown SD, Kawaguchi Y, Cox R, et al. Cloning of a novel member of the low density lipoprotein receptor family. Gene (in press)

Holmans P (1993) Asymptotic properties of affected-sib-pair linkage analysis. Am J Hum Genet 52:362–374

Kruglyak L , Lander ES (1995) Complete multipoint sib-pair analysis of qualitative and quantitative traits. Am J Hum Genet 57:439–454

Lernmark Å, Ducat L, Eisenbarth G, Ott J, Permutt MA, Rubenstein P, Spielman R (1990) Family cell lines available for research. Am J Hum Genet 47:1028–1030

Luo D-F, Buzzetti R, Rotter JI, Maclaren NK, Raffel LJ, Nisticò L, Giovannini C, et al (1996) Confirmation of three susceptibility genes to insulin-dependent diabetes mellitus: *IDDM4, IDDM5* and *IDDM8*. Hum Mol Genet 5:693–698

Martin ER, Kaplin NL , Weir BS (1997) Tests for linkage and association in nuclear families. Am J Hum Genet 61: 439–448

Merriman T, Eaves IA, Twells RCJ, Merriman ME, Danoy PAC, Muxworthy C, Hunter KMD, et al (1998) Transmission of haplotypes of microsatellite markers rather than single marker alleles in the mapping of a putative type 1 diabetes susceptibility gene (*IDDM6*). Hum Mol Genet 7: 517–524

Merriman T, Twells R, Merriman M, Eaves I, Cox R, Cucca F, McKinney P, et al (1997) Evidence by allelic-association dependent methods for a type 1 diabetes polygene (*IDDM6*) on chromosome 18q21. Hum Mol Genet 6:1003–1010

Nisticò L, Buzzetti R, Pritchard LE, Van der Auwera B, Giovannini C, Bosi E, Martinez Larrad MT, et al (1996) The *CTLA-4* gene region of chromosome 2q33 is linked to, and associated with, type 1 diabetes. Hum Mol Genet 5: 1075–1080

Nizetic D, Monard S, Young B, Cotter F, Zehetner G, Lehrach H (1994) Construction of cosmid libraries from flow-sorted human chromosomes 1, 6, 7, 11, 13 and 18 for reference library resources. Mamm Genome 5:801–802

Pociot F, Norgaard K, Hobolth N, Anderson O, Nerup J (1993) A nationwide population-based study of the familial aggregation of type 1 (insulin-dependent) diabetes mellitus in Denmark. Diabetologia 36:870–875

Podolin PL, Denny P, Armitage N, Lord CJ, Hill NJ, Levy ER, Peterson LB, et al (1998) Localization of two insulin-dependent diabetes (*Idd*) genes to the *Idd10* region on mouse chromosome 3. Mamm Genome 9:283–286

Podolin PL, Denny P, Lord CJ, Hill NJ, Todd JA, Peterson LB, Wicker LS, et al (1997) Congenic mapping of the insulin dependent diabetes (*Idd*) gene, *Idd10*, localizes two genes mediating the *Idd10* effect, and eliminates the candidate gene *Fcgr1*. J Immunol 159:1835–1843

Reed PW, Cucca F, Jenkins S, Merriman M, Wilson A, McKinney P, Bosi E, et al (1994) Chromosome-specific micro-

satellite sets for fluorescence-based, semi-automated genome mapping. Nat Genet 7:390–395

Risch N, Zhang H (1995) Extreme discordant sib pairs for mapping quantitative trait loci in humans. Science 268: 1584–1589

She J-X (1996) Susceptibility to type 1 diabetes: HLA-DQ and DR revisited. Immunol Today 17:323–329

Spielman RS, McGinnis RE, Ewens WJ (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am J Hum Genet 52:506–516

Thorsby E, Undlien D (1996) The HLA associated predisposition to type 1 diabetes and other autoimmune diseases. J Pediatr Endocrinal Metab Suppl 9:75–88

Tisch R, McDevitt H0 (1996) Insulin-dependent diabetes mellitus. Cell 85:291–297

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 351:542–547

Tuomilehto J, Lounamaa R, Tuomilehto-Wolf E, Reunanen A, Virtala E, Kaprio E, Akerblom HK (1992) Epidemiology of childhood diabetes mellitus in Finland—background of a nationwide study of type 1 (insulin-dependent) diabetes mellitus. Diabetologia 35:70–76

Vafiadis P, Bennett ST, Todd JA, Nadeau J, Grabs R, Goodyer CG, Wickramasinghe S, et al (1997) Insulin expression in human thymus is modulated by INS VNTR alleles at the *IDDM2* locus. Nat Genet 15:289–292

Wadsworth E, Shield JPH, Hunt L, Baum D (1995) Insulin dependent diabetes in children under 5: incidence and ascertainment validation for 1992. BMJ 310:700–703

Wicker LS, Todd JA, Prins J-B, Podolin PL, Renjilian RJ, Peterson LB (1994) Resistance alleles at two non-major histocompatibility complex-linked insulin-dependent diabetes loci on chromosome 3, *Idd3* and *Idd10,* protect nonobese diabetic mice from diabetes. J Exp Med 180:1705–1713