© 1997 Oxford University Press

Human Molecular Genetics, 1997, Vol. 6, No. 8 **1275–1282**

Insulin-dependent diabetes mellitus (IDDM) is associated with *CTLA4* polymorphisms in multiple ethnic groups

Michele P. Marron[1], Leslie J. Raffel[2], Henri-Jean Garchon[3], Chaim O. Jacob[4], Manuel Serrano-Rios[5], Maria T. Martinez Larrad[5], Wei-Ping Teng[6], Yongsoo Park[7], Zhi-Xing Zhang[8], Darlene R. Goldstein[2], Yi-Wen Tao[4], Genevieve Beaurain[3], Jean-Francois Bach[3], Hong-So Huang[9], De-Fang Luo[1], Adina Zeidler[4], Jerome I. Rotter[2], Mark C. K. Yang[10], Tamara Modilevsky[4], Noel K. Maclaren[1,+] and Jin-Xiong She[1,11,*]

[1]Department of Pathology, Immunology and Laboratory Medicine, [11]Center for Mammalian Genetics, College of Medicine, [10]Department of Statistics, University of Florida, Gainesville, Florida, USA, [2]Departments of Pediatrics and Medicine, Cedars-Sinai Medical Center and UCLA, Los Angeles, California, USA, [3]Unite de Recherches de l'INSERM U25, Centre de l'Association Claude Bernard, Paris, France, [4]Department of Medicine, University of Southern California School of Medicine, Los Angeles, USA, [5]Facultad de Medicina, Universidad Complutense, Madrid, Spain, [6]The First Hospital of China Medical University, ShenYang, China, [7]Hanyang University Hospital, Soul, Korea, [8]Beijing Red Cross Blood Center, Beijing, China and [9]ChangGun Memorial Hospital, Taipei, Taiwan

Received March 12, 1997; Revised and Accepted May 21, 1997

Linkage disequilibrium (association) analysis was used to evaluate a candidate region near the *CTLA4/CD28* genes using a multi-ethnic collection of families with one or more children affected by IDDM. In the data set unique to this study (Spanish, French, Mexican-American, Chinese and Korean), the transmission/ disequilibrium test (TDT) revealed a highly significant deviation for transmission of alleles at the $(AT)_n$ microsatellite marker in the 3′ untranslated region (P = 0.002) and the A/G polymorphism in the first exon (P = 0.00002) of the *CTLA4* gene. The overall evidence for transmission deviation of the *CTLA4* A/G alleles is also highly significant (P = 0.00005) in the combined data set (669 multiplex and 357 simplex families) from this study and a previous report on families from USA, Italy, UK, Spain and Sardinia. Significant heterogeneity was observed in these data sets. The British, Sardinian and Chinese data sets did not show any deviation for the A/G polymorphism, while the Caucasian-American data set showed a weak transmission deviation. Strong deviation for transmission was seen in the three Mediterranean-European populations (Italian, Spanish and French) (P = 10^{-5}), the Mexican-American population (P = 0.002) and the Korean population (P = 0.03). These results suggest that a true IDDM susceptibility locus (designated *IDDM12*) is located near *CTLA4*.

INTRODUCTION

Insulin-dependent diabetes mellitus (IDDM) is an autoimmune disease characterized by the selective destruction of the insulin-producing β cells in the pancreas. The development of the disease is influenced by the interaction of a large number of susceptibility genes and environmental factors. Two genomic regions have been implicated in IDDM susceptibility by candidate gene analyses using the case–control association approach. *IDDM1* has been localized to the HLA region which contains multiple IDDM susceptibility genes (1–5). *IDDM2* was mapped to the VNTR locus at the 5′ end of the insulin gene (6,7). More recently, affected sibpair analyses using random markers throughout the human genome have suggested 11 additional regions that may contain IDDM susceptibility genes (8–19). However, only three of the newly mapped intervals have been replicated in multiple studies and can be considered as confirmed linkage according to the criteria for complex disease genes proposed by Thomson (10) and Lander and Kruglyak (20). The confirmed intervals include *IDDM4* on 11q13 (8,9,13,14), *IDDM5* at 6q25 (10,14,21), and *IDDM8* at 6q27 (13,14,21). Three other regions, namely *IDDM11* on 14q (16), *IDDM13* on 2q (18) and *IDDM15* on 6q (19), showed significant linkage but independent studies are required for their confirmation. Preliminary indications of linkage for *IDDM3* at

*To whom correspondence should be addressed. Tel: +1 352 392 0677; Fax: +1 352 392 3053; Email: she.pathology@mail.health.ufl.edu
+Present address for NKM: Research Institute for Children, New Orleans

15q26 and *IDDM7* at 2q31 (11–13,22,23) were not confirmed in a study of larger sample size (14).

Mapping complex disease genes is a challenge due to several factors: (i) a large number of genes may be involved; (ii) each gene may only account for a small percentage of the total familial aggregation; and (iii) certain etiological mutations may not exist in all ethnic groups or geographical populations (i.e. genetic heterogeneity). Each of the recently confirmed IDDM genes may explain only 5–10% of the total genetic contribution to IDDM. To demonstrate the effect of such genes on a disease, many hundreds of families are required. This is further complicated by heterogeneity in different populations and/or ethnic groups.

CTLA4 and *CD28* on chromosome 2q33 are likely candidate genes for IDDM and other autoimmune diseases because of their important role in the T-cell proliferative response. The functional importance of these molecules in immune regulation has prompted us to examine their role in IDDM. Nisticò *et al*. (17) reported linkage and association evidence suggesting a susceptibility locus (designated *IDDM12*) in the *CTLA4/CD28* region. Evidence for *IDDM12* was found in Italian and Spanish populations but could not be confirmed in UK and US Caucasian data sets. Although the evidence from the Italian and Spanish data sets was highly significant ($P = 0.0001$), the combined data of all families did not provide sufficient evidence ($P = 0.002$) to confirm the existence of *IDDM12* (17). In addition, comparison of unrelated diabetic patients and controls also revealed significant associations with the *CTLA4* A/G polymorphism in Belgian and German populations (17,24). These results provided evidence supporting an IDDM susceptibility gene in the *CTLA4* region but also raised the question of why the association is only found in some populations. Here, we report highly significant associations between IDDM and two *CTLA4* polymorphisms in several populations as well as genetic heterogeneity among various populations.

RESULTS

Allele sharing by affected sibpairs

A microsatellite marker in the 3′ untranslated region and the +49 A/G polymorphism in exon 1 of *CTLA4* were genotyped for all French, Spanish and Caucasian-American families with two affected siblings (Table 1). Gene sharing by affected sibpairs was calculated using the haplotype data derived from the *CTLA4* A/G and $(AT)_n$ polymorphisms. Analysis of these families revealed a slight increase in haplotype sharing by affected sibpairs compared to the randomly expected 50% distribution (Table 1). The increase was, however, not statistically significant.

Transmission/disequilibrium (TDT) test of the *CTLA4* +49 A/G alleles

The TDT was then used to evaluate the transmission of the two alleles (A and G) at the +49 polymorphic site of *CTLA4* in all simplex and multiplex families that are unique to this study: French (FR), non-overlapping Spanish (SP), Mexican-American (MA), Korean and Chinese families (Table 2). The Mexican-American families showed a significant increase of transmission for the G allele (69.2%) from heterozygous parents to diabetic offspring compared to the randomly expected 50% ($P = 0.002$). The transmission of G allele in the French (59.7%) and Spanish (63.6%) families is also higher than the random expectation even though statistical significance is not reached due to small sample sizes. In the Korean simplex families, there was a significant increase in transmission of the G allele (76.5% and $P = 0.03$) and this transmission rate of the G allele is the highest of all data sets reported to date. In contrast, no such increase in transmission was seen in the Chinese families (44.4%). In the overall data set unique to this study, the transmission of the G allele was significantly higher than random expectation ($P = 0.00002$).

The Caucasian-American (CA) families studied here significantly overlap with the CA families studied in another report (17). Similar to the results in the previous report, there was an increased transmission (53.4%) of the G allele in the 301 Caucasian-American families studied here (Table 2); however, this deviation did not reach statistical significance ($P = 0.1$). Nisticò *et al*. (17) also reported significant increase of transmission of the G allele in the Italian (IT) and Spanish families but not in the UK or Sardinian families. We analyzed the combined data set (669 multiplex families and 357 simplex families) from this study and the previous report. The TDT revealed a significant increase of transmission of the G allele in the overall data set ($P = 0.00005$). It should be noted that genetic heterogeneity is clearly seen among the study populations (Table 2). To better understand the genetic heterogeneity, we combined various data sets according to geographic/ethnic/genetic considerations. The three Mediterranean-European (ME) populations (Italy, Spain and France) revealed a highly significant increase in transmission of the G allele ($P = 10^{-5}$). The ME data was also combined with the Mexican-American (MA) data set because the MA population is an admixture of Spanish and Amerindian genes and most IDDM genes in the MA population may be of Spanish origin (4). The combined ME+MA data subset gave the strongest evidence for association (10^{-7}). It could be argued that the Sardinian (SAR) data set should also be included in this group based on its geographic proximity to the Italian mainland. However, the Sardinian population appears to be genetically distinct from the Italians, Spanish or French (25–27). Even when the Sardinian data set is included, the ME-MA-SAR subset also gave highly significant evidence for association (10^{-6}).

Transmission/disequilibrium test of the *CTLA4* microsatellite alleles

The *CTLA4* microsatellite marker is extremely polymorphic and contains a large number of alleles. We designate the alleles according to the number of repeat units. For example, the smallest allele contains 8 AT repeats and is referred to as allele 8. We have evaluated the transmission of two common alleles (8 and 17) from heterozygous parents to affected children in simplex and multiplex families (Table 3). Allele 8 showed significantly decreased transmission in the Mexican-American data set ($P = 0.01$), the Caucasian-American data set ($P = 0.003$) and the combined data set ($P = 0.0001$). Allele 17 was transmitted significantly more often than randomly expected in these data sets. Most notably, the transmission deviation for allele 17 is significant in the Mexican-American data set ($P = 0.001$). These results are not unexpected since allele 17 is in strong linkage disequilibrium with the G allele and allele 8 with the A allele (data not shown). We also examined the transmission of a group of alleles (alleles 24–33) which show linkage disequilibrium with the A allele. No deviation from random transmission was observed for any of these alleles.

Table 1. Haplotype sharing by affected sibpairs for *IDDM12*

Data sets	1 ibd	0 ibd	PHS	*P*	MLS
French (FR)	22	20	52.4	n.s.	0.0
Spanish (SP)	7	8	46.7	n.s.	0.0
Caucasian-American (CA)	245	209	54.0	n.s.	0.6
Total	274	237	53.6	n.s.	0.6

PHS = percent of haplotype sharing, MLS = maximum lod score, n.s., not significant.

Table 2 Transmission of the G allele from heterozygous parents to diabetic children

Data sets	Transmitted (%)		Not transmitted (%)		*P*	Ref.[a]
French (FR)	40	(59.7)	27	(40.3)	n.s.	0
Spanish (SP)	14	(63.6)	8	(36.4)	n.s.	0
Mexican-American (MA)	45	(69.2)	20	(30.8)	0.002	0
Korean	13	(76.5)	3	(23.5)	0.03	0
Chinese	8	(44.4)	10	(55.6)	n.s.	0
Total (unique families)[b]	174	(62.8)	103	(37.2)	0.00002	0
Caucasian-American (CA)	304	(53.4)	265	(46.6)	0.1	0
SP and Italian (IT)	154	(62.3)	93	(37.7)	0.0001	1
UK	264	(51.0)	253	(49.0)	n.s.	1
Sardinian (SAR)	46	(50.5)	45	(49.5)	n.s.	1
ME (FR, SP and IT)[c]	208	(61.2)	128	(38.1)	10^{-5}	0,1
ME and MA[c]	253	(63.1)	148	(36.9)	10^{-7}	0,1
ME, MA and SAR[c]	299	(60.8)	193	(39.2)	10^{-6}	0,1
Unique + CA	478	(56.5)	368	(43.5)	0.0002	0
Grand total	888	(55.1)	724	(44.9)	0.00005	0,1

[a]Reference: 0 = this study, 1 = Nisticò *et al*. (17).
[b]Total (unique): total for families unique to this study (FR, SP, MA, Korean and Chinese). The results on Caucasian-American families are from this study but the families significantly overlap with another report (17).
[c]ME, Mediterranean-European.

Table 3. Transmission of the *CTLA4* microsatellite markers from heterozygous parents to diabetic children

Data sets[a]	Transmitted (%)		Not transmitted (%)		*P*
Allele 8					
FR and SP	42	(42.4)	57	(57.6)	n.s.
MA	17	(32.7)	35	(67.3)	0.01
Korean	5	(41.7)	7	(58.3)	n.s.
Chinese	7	(53.8)	6	(46.2)	n.s.
Total (unique families)	71	(40.3)	105	(59.7)	0.01
CA	234	(43.5)	304	(56.5)	0.003
Grand total	305	(42.7)	409	(57.3)	0.0001
Allele 17					
FR and SP	29	(58.0)	21	(42.0)	n.s.
MA	34	(73.9)	12	(26.1)	0.001
Korean	9	(39.1)	14	(60.9)	n.s.
Chinese	12	(52.2)	11	(47.8)	n.s.
Total (unique families)	84	(59.2)	58	(40.8)	0.002
CA	229	(53.3)	201	(46.7)	n.s.
Grand total	313	(54.7)	259	(45.3)	0.02
Alleles 24–33					
FR and SP	33	(54.1)	28	(45.9)	n.s.
MA	10	(41.7)	14	(58.3)	n.s.
Korean	4	(40.0)	6	(60.0)	n.s.
Chinese	2	(28.6)	5	(71.4)	n.s.
Total (unique families)	49	(48.0)	53	(52.0)	n.s.
CA	185	(54.4)	155	(45.6)	n.s.
Grand total	234	(52.9)	208	(47.1)	n.s.

[a]Abbreviations for data sets are identical to Table 2.

Table 4. CTLA4 +49 polymorphism genotype frequencies in IDDM and Graves disease[a]

Population	N	Genotypic frequencies			Allelic frequencies	
		AA (%)	AG (%)	GG (%)	A	G
IDDM						
Caucasian:						
Florida patients[b]	244	83 (34.0)	119 (48.8)	42 (17.2)	0.584	0.416
Florida controls	274	97 (35.4)	142 (51.8)	35 (12.8)	0.613	0.387
Spanish patients	89	46 (51.7)	37 (41.6)	6 (6.7)	0.652	0.348
Spanish controls	57	23 (40.3)	29 (50.9)	5 (8.8)	0.658	0.342
Korean:						
Patients	97	2 (2.1)	38 (39.2)	57 (58.8)	0.216	0.784
Controls	112	12 (10.7)	44 (39.3)	56 (50.0)	0.304	0.696
Chinese[c]:						
TW patients	96	5 (5.2)	48 (50.0)	43 (44.8)	0.302	0.698
TW controls	191	19 (9.9)	68 (35.6)	104 (54.5)	0.277	0.723
SY patients	53	6 (11.3)	22 (41.5)	25 (47.2)	0.321	0.679
SY controls	94	6 (6.4)	39 (41.2)	49 (52.1)	0.271	0.729
BJ patients	31	6 (19.4)	12 (38.7)	13 (41.9)	0.387	0.613
BJ controls	94	10 (10.6)	41 (43.6)	43 (45.7)	0.324	0.676
Total Chinese:						
Patients	180	17 (9.4)	82 (45.6)	81 (45.0)	0.322	0.678
Controls	379	35 (9.2)	148 (39.0)	196 (51.7)	0.288	0.712
Graves disease						
SY patients	28	1 (3.6)	11 (39.3)	16 (57.1)	0.232	0.768
SY controls	94	6 (6.4)	39 (41.2)	49 (52.1)	0.271	0.729

[a]No statistically significant difference was found between patient group and ethnically-matched control group in any of the comparisons.
[b]Includes affected siblings from 60 families.
[c]TW, Taiwan; SY, Shen Yang; BJ, Beijing.

Case–control association analyses

To follow up the association evidence for *IDDM12* in the family-based studies, case–control association analyses were carried out using a large collection of patients and ethnically-matched controls from Florida, Spain, Korea, mainland China and Taiwan (Table 4). Since the transmission of allele G was increased in diabetic families, the frequencies for the G allele or GG genotype are expected to be higher in diabetic patients compared with ethnically-matched controls. In the Florida Caucasian population, the frequencies of the GG genotype and G allele were indeed increased in patients compared with controls, but the increases were not statistically significant. However, the percentages of increase in the Florida data set are very similar to those reported in a large Belgian data set (17) and a German data set (24). Surprisingly, there were no differences in the genotypic or allelic frequencies between Spanish sporadic IDDM patients and controls despite strong evidence for transmission deviation in the Spanish IDDM families.

The frequencies of the G allele are much higher in the Korean and Chinese populations than in Caucasian populations (Table 4). The Korean data set revealed increased frequencies of the GG genotype and the G allele in patients compared with controls; however, the differences did not reach statistical significance. In all three Chinese data sets, the frequencies of the GG genotype and G allele were actually lower in IDDM patients compared with ethnically- and geographically-matched controls. These results are not consistent with the increased susceptibility effect on IDDM by the G allele as suggested by family-based studies in other populations. It is important to note that the differences were again not statistically significant and may reflect random variations.

We also analyzed the +49 A/G polymorphism in 28 patients with Graves disease from Northern China. The GG genotype and the G allele have higher frequencies in patients compared to controls, consistent with several previous studies (17,24,28) in Chinese and German populations. Even though the differences did not reach statistical significance, our results are consistent with a possible association between *CTLA4* and Graves disease.

DISCUSSION

Association between IDDM and the *CTLA4* polymorphisms has now been detected in six different populations (Italian, Spanish, French, Mexican-American, Caucasian-American and Korean) using family-based association studies. The overall evidence for association is highly significant in the combined data set from these six populations ($P = 10^{-6}$). However, significant associations were not found in a large data set of 284 UK families (17) or in the small Chinese and Sardinian data sets. Even when these

three data sets are included, the combined evidence for association in a total of 1026 families (669 multiplex and 357 simplex) is still highly significant ($P = 0.00005$), providing strong support for an IDDM susceptibility gene in the region surrounding *CTLA4*. In addition to the family-based studies, highly significant evidence for association was also reported in two independent case–control studies, one in the Belgian population ($P = 10^{-5}$) (17) and the other in the German population ($P = 10^{-5}$) (24). In our case–control studies in the Florida Caucasian and Korean populations, the increases of the GG genotype and G allele in patients compared with controls are consistent with the increase of transmission of the G allele from heterozygous parents to diabetic offspring, even though statistical significance was not reached. However, no evidence for association between IDDM and *CTLA4* was revealed in three Chinese populations or in the sporadic IDDM cases from Spain. Case–control studies are very sensitive to ascertainment of patient and control subjects, as well as sample sizes. Such a study design is not very powerful in detecting small differences between patient and control populations unless a very large data set is studied. Indeed, the differences in the frequency of the GG genotype or G allele between patients and ethnically-matched controls are small (5–10%) but very similar in the Belgian, German, Korean and US Caucasian populations. Such a small difference could well be due to random variation. However, we believe that the observed associations may not be spurious because the GG genotype and G allele consistently has a higher frequency in patients than in controls in these populations. These data suggest that family-based association studies are less susceptible to sample bias and therefore should be the preferred experimental design for association studies.

The overall data reported here and by Nisticò *et al*. (17) suggest that *IDDM12* is a true IDDM susceptibility locus. Among all analyzed populations, only the UK data set had sufficient families to confirm the lack of association between IDDM and *CTLA4*. The sample sizes of the Sardinian and Chinese data sets were too small for any definitive conclusions to be made at this stage. The lack of association between IDDM and *CTLA4* in the UK data set may be explained by at least two hypotheses. First, linkage disequilibrium patterns between the *CTLA4* A/G polymorphism and the *IDDM12* etiological mutation may be different in the UK population than those in the Mediterranean populations. In other words, disequilibrium between *CTLA4* A/G and the etiological mutation may not be as tight in the UK population as in the Italian, French and other populations. This hypothesis was first proposed by Nisticò *et al*. (17) and remains to be tested. Second, the *IDDM12* etiological mutation may have either an extremely high or low frequency (close to monomorphic) in the UK population so that its susceptibility effect cannot be detected by genetic studies. An example of such a situation is illustrated by the insulin (*INS*) gene polymorphisms in IDDM. *INS* polymorphisms have been associated with IDDM in many Caucasian populations (6,29–34) but not in the Chinese population because the frequency of the susceptible *INS* allele is ~95% in Chinese controls and close to 100% in Chinese IDDM (M. Marron and J.X. She, unpublished data). In the case of *IDDM12*, comparison of the Mediterranean, US and the UK populations appears to suggest that the frequency of the *IDDM12* mutation is very low in the UK population. The high transmission rates of the disease-associated *CTLA4* alleles to diabetic offspring in the Mediterranean populations (61.2%) may suggest that the *IDDM12* etiological mutation has a relatively high frequency in those populations. The intermediate transmission rates of the disease-associated alleles in the Caucasian-American data set are consistent with an intermediate frequency of the *IDDM12* mutation. This may be because the Caucasian-American population is an admixture of descendants from the UK with a low frequency of *IDDM12* and Mediterranean-European countries with a higher frequency of *IDDM12*. To test this hypothesis, we analyzed a Mexican-American population and observed a high transmission rate of the disease-associated *CTLA4* alleles. Since the Mexican-American population is an admixture of Spanish and Amerindian genes, our results provide supporting evidence for a Mediterranean-European origin of the *IDDM12* etiological mutation.

Our analyses and interpretation of the *CTLA4* data provide an excellent illustration of the difficulty in mapping complex disease genes when there is genetic heterogeneity. Not only are a large number of families required but a multi-ethnic approach is also essential to identify all disease genes. If the Mediterranean populations had not been analyzed, *IDDM12* would have been overlooked or excluded from further consideration. The lessons learned here should be relevant to most, if not all, other complex diseases.

IDDM12 was detected through TDT analysis, which only yields positive results when significant linkage disequilibrium exists between the marker and the disease gene. The data suggest that *IDDM12* must be very close to the *CTLA4* gene based on the degree of linkage disequilibrium in most of the human genome. It has been shown that the maximum distance for linkage disequilibrium to exist may be <2 cM. Linkage disequilibrium between markers at distances as large as 1–2 cM may only occur in regions with multi-gene families, such as the HLA region. For most of the human genome, linkage disequilibrium in non-isolated populations may exist only between markers within a few hundred kb of DNA sequence (probably <300 kb) (35–37). In some extreme cases, for example, the insulin gene region (*IDDM2*), strong linkage disequilibrium only exists between markers within a region of 4.5 kb of DNA (30). Therefore, *IDDM12* may be located in a region that is less than a few hundred kb flanking *CTLA4*.

CTLA4 and *CD28* are likely candidate genes for *IDDM12* because of their role in the regulation of T cell activation. Co-stimulation of CD28 by B7 in the presence of antigen-MHC signaling through the TCR-CD3 complex leads to T cell proliferation and production of IL-2 and CTLA4 expression. CTLA4 plays a role in limiting the T cell proliferative response. Gribben *et al*. (38) have suggested this may be through antigen-specific induction of the apoptotic pathway. More recent evidence suggests that signaling through CTLA4 may act to down-regulate proliferative responses by inhibiting production of IL-2, IL-2 receptor expression and progression of the cell cycle (39,40). Disruption of the delicate balance between CD28 and CTLA4 interactions with B7 could lead to autoimmune disease by preventing apoptosis or down regulation of activated self-reactive T lymphocytes (41).

Even though *CTLA4* is the most likely candidate gene for *IDDM12*, the etiological mutation has not been found. The exon 1 A/G polymorphism itself is an unlikely candidate because the Thr/Ala substitution is not expected to affect the function of the CTLA4 molecule. Furthermore, the lack of association with the *CTLA4* A/G polymorphism in the UK population provides direct

evidence that the A/G polymorphism is not the etiological mutation. The microsatellite $(AT)_n$ repeat is located in the 3′ untranslated region and could affect mRNA stability. If the $(AT)_n$ repeat is the etiological mutation, alleles in the same range of repeat units would show similar transmission ratios. In our data set, however, alleles with smaller repeat units (allele 8) are transmitted less frequently, alleles within the middle range repeat units (allele 17) are transmitted more often, while alleles in the higher range of repeat units (alleles 24–33) did not show any transmission deviation. These results suggest that the $(AT)_n$ microsatellite is unlikely to be the *IDDM12* etiological mutation. Detailed physical and genetic mapping of the region surrounding *CTLA4* will be required for the identification of the *IDDM12* etiological mutation.

MATERIALS AND METHODS

Subjects

Genomic DNA was obtained for 301 US Caucasian families affected by IDDM (297 multiplex families with two or more affected siblings and four simplex families with one affected child). The US case–control data set included 246 IDDM patients from Florida and 192 matched controls. The French data set included 12 simplex and 32 multiplex families. The Spanish data set included nine simplex and nine multiplex families as well as 89 unrelated patients and 57 controls. We have analyzed a total of 97 Mexican-American families (86 simplex and 11 multiplex) ascertained by the investigators at the LAC/USC Medical Center (4,5,42,43). The Chinese data set consisted of patients and controls from three geographical regions. There were 96 patients and 191 matched controls from Taiwan, 53 patients and 94 controls from ShenYang, and 31 patients and 94 controls from Beijing. We also ascertained one multiplex and 30 simplex Chinese IDDM families. Twenty-eight patients with Graves' disease were obtained from ShenYang. The Korean population consisted of 97 patients with IDDM, 112 matched controls, and 41 simplex families. All together, our data set includes 182 simplex and 350 multiplex IDDM families. All diabetic patients used in this study were diagnosed using the criteria of the National Diabetes Data Group (44). The clinical information on the patients has been described in previous publications (4,5,13,14,23,31,32,42,43,45).

Genotyping

The $(AT)_n$ microsatellite in the 3′ untranslated region of the *CTLA-4* gene was genotyped using radioactive labeling of PCR primers and denaturing polyacrylamide gel electrophoresis as previously reported (13,14,23). PCR was carried out using a forward primer CTLA4–1 (5′-GCCAGTGATGCTAAAGGT-TG-3′) and a reverse primer CTLA4–2 (5′-ACACAAAAACA-TACGTGGCTC-3′). The CTLA4–1 primer was end-labeled using [γ-^{32}P]ATP and T4 polynucleotide kinase. PCR amplifications were performed on 20 ng of genomic DNA (prealiquoted into a 96-well microtitre plate) in 12 µl reaction volume containing 50 mM KCl, 10 mM Tris–HCl, pH 8.3, 1.5 mM $MgCl_2$; 60 µM of each dNTPs. Samples were subjected to 30 cycles of 30 s at 94°C for denaturing, 30 s at 58°C for annealing and 30 s at 72°C for extension, using a Perkin-Elmer-Cetus 9600 thermal cycler. After PCR amplification, 2 vol of sequencing loading solution [0.3% xylene cyanol, 0.3% bromophenol blue, 10 mM EDTA, pH 8.0 and 90% (v/v) formamide] were added. The samples were 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.

A 152 bp fragment containing the +49 A/G polymorphism in exon 1 of *CTLA4* was amplified using a forward primer (CTLA4–5: 5′-AAGGCTCAGCTGAACCTGGT-3′) and a reverse primer (CTLA4–4: 5′-CTGCTGAAACAAATGAAA-CCC-3′). The forward primer was designed with a single base mismatch for the last nucleotide, which corresponds to the +47 position, to introduce a base change in the sequence of the PCR product. The substitution creates a *Bst*EII restriction site in the A allele. Amplification was performed using similar conditions as described for the microsatellite marker. Amplified products (12 µl) were incubated at 60°C overnight using 5 U of *Bst*EII per reaction. Digested products were electrophoresed on a 3.5% agarose gel. The digested A allele yields a fragment of 130 bp and the G allele yields an intact 152 bp fragment.

Data analysis

Data were first analyzed using affected sibpair methods which compare the observed proportion of alleles shared by affected sibpairs (identity by descent or ibd) to the random expectation. Affected sibpairs are expected to share 50% of alleles in regions that are not implicated in the disease. Significant increase of gene sharing compared to the 50% random expectation would indicate linkage. Deviation can be tested using a χ^2 test of equal expectation. Alternatively, maximum likelihood of odds score (MLS) can be calculated using the equation of Risch (46),

$$T = N_1 (\log_{10} [N_1/0.5N]) + N_0 (\log_{10} [N_0/0.5N])$$

where N_0 and N_1 are the number of affected siblings sharing 0 and 1 alleles, respectively, and N is the total number of informative meioses ($N_0 + N_1$).

Intrafamilial association studies were carried out using the transmission/disequilibrium test (TDT) as described by Spielman *et al*. (47). This method compares the number of times a given allele or a group of alleles are transmitted from heterozygous parents to diabetic offspring with the 50% random expectation. It can be applied to multiplex as well as simplex families. It is important to note that significant deviation is observed only when the marker allele and the etiological mutation show significant disequilibrium (association). Case–control association studies were carried out for the genotype frequencies at the +49 A/G polymorphic site using a 2 × 3 contingency χ^2 test for the three genotypes (AA/AG/GG).

ACKNOWLEDGEMENTS

This work was supported by JDFI grant JDF196113 and the NIH grant DK50220 to J.X.S. D.R.G. was supported by a NIH grant ST32 DK07682. C.O.J. and A.Z were supported by JDFI and ADA grants and their molecular core was supported by MO1 RR-43 grant from NCRR. J.I.R. is the Cedars-Sinai Board of Governors' Chair in Medical Genetics. D.R.G. was supportd by NIH grant 5T32 DK07682. We thank the Spanish IDDM study group for their help in the collection of Spanish IDDM patients. M.S.R. was supported by a grant from Fonde de Investigatione Sanitarias FISS #96/2088.

REFERENCES

1. She,J.X. (1996) Susceptibility to type I diabetes: HLA-DQ and DR revisited. *Immunol. Today*, **17**, 323–329.
2. Cucca,F., Lampis,R., Frau,F., Macis,D., Angius,E., Masile,P., Chessa,M., Frongia,P., Silvetti,M., Cao,A., Virgiliis,S.D. and Congia,M. (1995) The distribution of DR4 haplotypes in Sardinia suggests a primary association of Type I diabetes with DRB1 and DQB1 loci. *Hum. Immunol.*, **43**, 301–308.
3. Deng,G.Y., Muir,A., Maclaren,N.K. and 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.
4. Erlich,H.A., Rotter,J.I., Chang,J., Shaw,S., Raffel,L.J., Klitz,W., Beshkov,Y., Costin,R., Pressman,S., Bugawan,T. and Zeidler,A. (1993) HLA class II alleles and susceptibility and resistance to insulin-dependent diabetes mellitus in Mexican-American families. *Nature Genet.*, **3**, 358–364.
5. Erlich,H.A., Rotter,J.I., Chang,J., Shaw,S., Raffel,L.J., Klitz,W., Bugawan,T. and Zeidler,A. (1996) Independent association of HLA-DPB1*0301 with insulin-dependent diabetes mellitus in Mexican-Americans. *Diabetes*, **45**, 610–614.
6. Bennett,S.T., Lucassen,A.M., Gough,S.C.L., Powell,E.E., Undlien,D.E., Pritchard,L.E., Merriman,M.E., Kawaguchi,Y., Dronsfield,M.J., Pociot,F., Nerup,J., Bouzekri,N., Cambon-Thomsen,A., Ronnigen,K.S., Barnett,A.H., Bain,C.C. and Todd,J.A. (1995) Susceptibility to human type 1 diabetes at IDDM2 is determined by tandem repeat variation at the insulin gene minisatellite locus. *Nature Genet.*, **9**, 284–292.
7. Kennedy,G.C., German,M.S. and Rutter,W.J. (1995) The minisatellite in the diabetes susceptibility locus IDDM2 regulates insulin transcription. *Nature Genet.*, **9**, 293–297.
8. Davies,J.L., Yoshihiko,K., Bennett,S., Copeman,J.B., Cordell,H.J., Pritchard,L.E., Reed,P.W., Gough,S.C.L., Jenkins,C., Palmer,S.M., Balfour,K.M., Rowe,B.R., Farrall,M., Barnett,A.H., Bain,S.C. and Todd,J.A. (1994) A genome-wide search for human type 1 diabetes susceptibility genes. *Nature*, **371**, 130–136.
9. Hashimoto,L., Habita,C., Beressi,J.P., Delepine,M., Besse,C., Cambon-Thomsen,A., Deschamps,I., Rotter,J.I., Djoulah,S., James,M.R., Froguel,P., Weissenbach,J., Lathrop,G.M. and Julier,C. (1994) Genetic mapping of a susceptibility locus for insulin-dependent diabetes mellitus on chromosome 11q. *Nature*, **371**, 161–164.
10. Thomson,G. (1994) Identifying complex disease genes: progress and paradigms. *Nature Genet.*, **8**, 108–110.
11. Field,L.L., Tobias,R. and Magnus,T. (1994) A locus on chromosome 15q26 (IDDM3) produces susceptibility to insulin-dependent diabetes mellitus. *Nature Genet.*, **8**, 189–194.
12. Owerbach,D. and Gabbay,K.H. (1995) The HOXD8 locus (2q31) is linked to type I diabetes: Interaction with chromosome 6 and 11 disease susceptibility genes. *Diabetes*, **44**, 132–136.
13. Luo,D.F., Bui,M.M., Muir,A., Maclaren,N.K., Thomson,G. and She,J.X. (1995) Affected sibpair mapping of a novel susceptibility gene to insulin-dependent diabetes mellitus (IDDM8) on chromosome 6q25-q27. *Am. J. Hum. Genet.*, **57**, 911–919.
14. Luo,D.F., Buzzetti,R., Rotter,J.I., Maclaren,N.K., Raffel,L.J., Nistico,L., Giovannini,C., Pozzilli,P., Thomson,G. and She,J.X. (1996) Confirmation of three susceptibility genes to insulin-dependent diabetes mellitus: IDDM4, IDDM5 and IDDM8. *Hum. Mol. Genet.*, **5**, 693–698.
15. Todd,J.A. (1995) Genetic analysis of type 1 diabetes using whole genome approaches. *Proc. Natl. Acad. Sci. USA*, **92**, 8560–8565.
16. Field,L.L., Tobias,R., Thomson,G. and Plon,S. (1996) Susceptibility to insulin-dependent diabetes mellitus maps to a locus (IDDM11) on human chromosome 14q24.3-q31. *Genomics*, **33**, 1–8.
17. Nisticò,L., Buzzetti,R., Pritchard,L.E., Van der Auwera,B., Giovannini,C., Bosi,E., Martinez Larrad,M.T., Rios,M.S., Chow,C.C., Cockram,C.S., Jacobs,K., Mijovic,C., Bain,S.C., Barnett,A.H., Vandewalle,C.L., Schuit,F., Gorus,F.K., Belgian Diabetes Registry, Tosi,R., Pozzilli,P. and Todd,J.A. (1996) The CTLA4 gene region of chromosome 2q33 is linked to, and associated with, type 1 diabetes. *Hum. Mol. Genet.*, **5**, 1075–1080.
18. Morahan,G., Huang,D., Tait,B.D., Colman,P.G. and Harrison,L.C. (1996) Markers on distal chromosome 2q linked to insulin-dependent diabetes mellitus. *Science*, **272**, 1811–1813.
19. Delepine,M., Pociot,F., Habita,C., Hashimoto,L., Froguel,P., Rotter,J., Cambon-Thomsen,A., Deschamps,I., Djoulah,S., Weissenbach,J., Nerup,J., Lathrop,M. and Julier,C. (1997) Evidence of a non-MHC susceptibility locus in type I diabetes linked to HLA on chromosome 6. *Am. J. Hum. Genet.*, **60**, 174–187.
20. Lander,E. and Kruglyak,L. (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nature Genet.*, **11**, 241–247.
21. Davies,L.D., Cucca,F., Goy,J.V., Atta,Z.A.A., Merriman,M.E., Wilson,A., Barnett,A.H., Bain,S.C. and Todd,J.A. (1996) Saturation multipoint linkage mapping of chromosome 6q in type 1 diabetes. *Hum. Mol. Genet.*, **5**, 1071–1074.
22. Copeman,J.B., Cucca,F., Hearne,C.M., Cornall,R.J., Reed,P.W., Ronningen,K.S., Undlien,D.E., Nistico,L., Buzzetti,R., Tosi,R., Pociot,F., Nerup,J., Cornelis,F., Barnett,A.H., Bain,S.C. and Todd,J.A. (1995) Linkage disequilibrium mapping of a type 1 diabetes susceptibility gene (IDDM7) to chromosome 2q31-q33. *Nature Genet.*, **9**, 80–85.
23. Luo,D.F., Maclaren,N.K., Huang,H.S., Muir,A. and She,J.X. (1995) Intrafamilial and case-control association analyses of D2S152 in insulin-dependent diabetes mellitus. *Autoimmunity*, **21**, 143–145.
24. Donner,H., Rau,H., Walfish,P.G., Braun,J., Siegmund,T., Finke,R., Herwig,J., Usadel,K.H. and Badenhoop,K. (1997) CTLA4 alanine-17 confers genetic susceptibility to Graves' disease and to type 1 diabetes mellitus. *J. Clin. Endocrinol. Metab.*, **82**, 143–146.
25. Olivetti,E., Rendine,S., Cappello,N., Curtoni,E.S. and Piazza,A. (1986) The HLA system in Italy. *Hum. Hered.*, **36**, 357–372.
26. Ryder,L.P., Andersen,E. and Svejgaard,A. (1978) An HLA map of Europe. *Hum. Hered.*, **28**, 171–200.
27. Vona,G., Salis,M., Bitti,P. and Succa,V. (1994) Blood groups of the Sardinian population. *Anthropol. Anz.*, **52**, 297–304.
28. Yanagawa,T., Hidaka,Y., Guimaracs,V., Soliman,M. and DeGroot,L.J. (1995) CTLA4 gene polymorphism associated with Graves' disease in a Caucasian population. *J. Clin. Endocrinol. Metab.*, **80**, 41–45.
29. Bell,G.I., Horita,S. and Karam,J.H. (1984) A polymorphic locus near the human insulin gene is associated with insulin-dependent diabetes mellitus. *Diabetes*, **33**, 176–183.
30. Lucassen,A.M., Julier,C., Beressi,J.P., Boitard,C., Froguel,P., Lathrop,M. and Bell,J.I. (1993) Susceptibility to insulin dependent diabetes mellitus maps to a 4.1 kb segment of DNA spanning the insulin gene and associated VNTR. *Nature Genet.*, **4**, 305–310.
31. She,J.X., Bui,M.M., Tian,X.H., Muir,A., Wakeland,E.K., Zorovich,B., Zhang,L.P., Liu,M.C., Thomson,G. and Maclaren,N.K. (1994) Additive susceptibility to insulin-dependent diabetes conferred by HLA-DQB1 and insulin genes. *Autoimmunity*, **18**, 195–203.
32. Bui,M.M., Luo,D.F., She,J.Y., Maclaren,N.K., Muir,A., Thomson,G. and She,J.X. (1996) Paternally transmitted IDDM2 influences diabetes susceptibility despite biallelic expression of the insulin gene in human pancreas. *J. Autoimmun.*, **9**, 97–103.
33. Raffel,L.J., Hitman,G.A., Toyoda,H., Karam,J.H., Bell,G.I. and Rotter,J.I. (1992) The aggregation of the 5' insulin gene polymorphism in type I (insulin-dependent) diabetes mellitus families. *J. Med. Genet.*, **29**, 447–450.
34. Julier,C., Hyer,R.N., Davies,J., Merlin,F., Soularue,P., Briant,L., Cathelineau,G., Deschamps,I., Rotter,J.I., Froguel,P., Boitard,C., Bell,J.I. and Lathrop,G.M. (1991) Insulin-IGF2 region on chromosome 11p encodes a gene implicated in HLA-DR4-dependent diabetes susceptibility. *Nature*, **354**, 155–158.
35. Jorde,L.B., Watkins,W.S., Carlson,M., Groden,J., Albertsen,H., Thliveris,A. and Leppert,M. (1994) Linkage disequilibrium predicts physical distance in the adenomatous polyposis coli region. *Am. J. Hum. Genet.*, **54**, 884–898.
36. Sherrington,R., Melmer,G., Dixon,M., Curtis,D., Mankoo,B., Kalsi,G. and Gurling,H. (1991) Linkage disequilibrium between two highly polymorphic microsatellites. *Am. J. Hum. Genet.*, **49**, 966–971.
37. Jorde,L.B., Watkins,W.S., Viskochil,D., O'Connell,P. and Ward,K. (1993) Linkage disequilibrium in the neurofibromatosis 1 (NF1) region: implications for gene mapping. *Am. J. Hum. Genet.*, **53**, 1038–1050.
38. Gribben,J.G., Freeman,G.J., Boussiotis,V.A., Rennert,P., Jellis,C.L., Greenfield,E., Barber,M., Restivo,J., Ke,X., Gray,G.S. and Nadler,L.M. (1995) CTLA4 mediates antigen-specific apoptosis of human T cells. *Proc. Natl. Acad. Sci. USA*, **92**, 811–815.
39. Walunas,T.L., Bakker,C.Y. and Bluestone,J.A. (1996) CTLA-4 ligation blocks CD28-dependent T cell activation. *J. Exp. Med.*, **183**, 2541–2550.
40. Krummel,M.F. and Allison,J.P. (1996) CTLA-4 engagement inhibits Il-2 accumulation and cell cycle progression upon activation of resting T cells. *J. Exp. Med.*, **183**, 2533–2540.
41. Tivol,E.A., Borriello,A.N., Schweitzer,A.N., Lynch,W.P., Bluestone,J.A. and Sharpe,A.H. (1995) Loss of CTLA4 leads to massive lymphoproliferation and fatal multiorgan tissue destruction, revealing a critical negative regulatory role of CTLA4. *Immunity*, **3**, 541–548.

42. Vadheim,C.M., Zeidler,A., Rotter,J.I., Langbaum,M., Schulman,I.A., Spencer,M.R., Costin,G., Riley,W.J. and Maclaren,N.K. (1989) Different HLA haplotypes in Mexican-Americans with IDDM. *Diabetes Care*, **12**, 477–500.
43. Shaw,S.J., Vadheim,C.M., Rotter,J.I., Costin,G., Beshkov,Y. and Zeidler,A. (1992) Ancestral origin of insulin-dependent diabetes in Mexican-Americans. *Int. J. Epidemiol*, **21**, 725–729.
44. National Diabetes Data Group, (1979) Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. *Diabetes*, **28**, 1039–1057.
45. Huang,H.S., Peng,J.T., She,J.Y., Zhang,L.P., Chao,C.C.K., Liu,K.H. and She,J.X. (1995) HLA-encoded susceptibility to IDDM is determined by DR and DQ genes as well as their linkage disequilibrium in a Chinese population. *Hum. Immunol.*, **44**, 210–219.
46. Risch,N. (1990) Linkage strategies for genetically complex Traits. II. The power of affected relative pairs. *Am. J. Hum. Genet.*, **46**, 229–241.
47. Spielman,R.S., Mcginnis,R.E. and Ewens,W.J. (1993) Transmission test for linkage disequilibrium: The insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am. J. Hum. Genet.*, **52**, 506–512.