Brief Genetics Report

Physical and Genetic Mapping of *IDDM8* on Chromosome 6q27

David Owerbach

Genome-wide mapping studies have provided evidence of a type 1 diabetes susceptibility gene (*IDDM8*) that is located on chromosome 6q27. However, association studies of *IDDM8* have so far been negative. The purpose of this investigation was to determine a linkage disequilibrium (LD) map in the chromosome 6q27 region and to better localize *IDDM8*. A physical map of nearly 1 Mb containing the chromosome 6 telomere was constructed, and polymorphic markers spanning this region were defined. Haplotypes composed of the markers in LD were tested for association with type 1 diabetes in 266 families. A microsatellite marker allele and multiple haplotypes were associated with *IDDM8*, which suggests localization of this type 1 diabetes susceptibility gene to the terminal 200 kb of chromosome 6. *Diabetes* 49:508–512, 2000

Genome-wide mapping studies have provided evidence for several susceptibility loci involved in type 1 diabetes (1–3). However, in these studies, the genes are localized to a very broad region of 5–20 Mb on a specific chromosome. A major challenge in the genetics of multigenic diseases, such as type 1 diabetes, is to progress beyond initial genome searches to actually find the disease susceptibility genes. Association studies in families have been useful in localizing *IDDM6* (4,5) and *IDDM10* (6), although the identity of these susceptibility genes still remains unknown. *IDDM8*, which is located on chromosome 6q27, shows clear evidence for linkage with type 1 diabetes with a maximal logarithm of odds score of 5.0 (7–9). However, association studies of *IDDM8* have so far been negative. The purpose of this study was to better localize *IDDM8*.

From the Harry B. and Aileen Gordon Diabetes Research Laboratory, Molecular Diabetes and Metabolism Section, Department of Pediatrics, Baylor College of Medicine, Houston, Texas.

Address correspondence and reprint requests to Dr. David Owerbach, the Harry B. and Aileen Gordon Diabetes Research Laboratory, Molecular Diabetes and Metabolism Section, Department of Pediatrics, Baylor College of Medicine, Houston, TX 77030. E-mail: davido@bcm.tmc.edu.

Received for publication 3 June 1999 and accepted in revised form 17 November 1999.

Additional information can be found in an online appendix at www.diabetes.org/diabetes/appendix.asp.

BAC, bacterial artificial chromosome; EST, expressed sequenced tags; HC5, human proteasome complex-5 subunit; LD, linkage disequilibrium; PAC, P1-bacteriophage–derived artificial chromosome; PCR, polymerase chain reaction; SSCP, single-strand conformation polymorphism; *T*, transmission frequency; TBP, TATA-binding protein; Y1, YAC 2158; YAC, yeast artificial chromosome.

A physical map of the chromosome 6q terminal region is shown in Fig. 1. Yeast artificial chromosome (YAC) 2158 contains the chromosome 6q telomere and defines the end of the chromosome (11). From the sizes of YAC 2158 (280 kb), P1-bacteriophage–derived artificial chromosome (PAC) 5724 (105 kb), and bacterial artificial chromosome (BAC) 62J14 (125 kb), it is estimated that the 5′ end of BAC 62J14 is ~50 kb from the telomere. The approximate distance from the telomere to the end of BAC 62D7 is 845 ± 100 kb. Microsatellite polymorphisms have previously been described for TATA-binding protein (TBP), D6S1590, D6S446, D6S281, and AFM234 (10). In Fig. 1, the location of these genetic markers and 12 expressed sequenced tags (ESTs) are shown on the physical map. Currently, approximately half of the chromosome 6 region shown in Fig. 1 (400 kb) has been sequenced at the Sanger Genome Centre; the sequence data can be obtained from ftp://ftp.sanger.ac.uk/human chromosome 6.

In addition to the five previously described microsatellite polymorphisms, five new marker polymorphisms were identified and characterized (Table 1). Microsatellite markers 191, M122K, and 137 were identified by the presence of multiple CA repeats found in the Sanger database. Upon subsequent testing, it was found that 191, M122K, and 137 (located 141, 190, and 842 kb from the telomere) have four, six, and three major alleles, respectively (Table 1). Moreover, two markers, 62J (located 167 kb from the telomere) and 7419 (located 520 kb from the telomere), were identified in cosmid clones and were determined to be polymorphic by single-strand conformation polymorphism (SSCP) analysis. Allele frequencies are shown for the eight microsatellite and two SSCP markers in Table 1.

In the initial analysis of LD, alleles at M122K/D6S1590 and D6S281/AFM234 that had a frequency >5% were analyzed (30 allele combinations) (Table 1). The significance of LD is determined by χ^2 contingency tests using a 2×2 table format. To account for multiple testing, χ^2 values >10.6 ($P < 0.001$) were considered significant. As shown in Table 2, three combinations of M122K/D6S1590 alleles and two combinations of D6S281/AFM234 alleles were in LD. Additional marker alleles flanking the M122K/D6S1590 and D6S281/AFM234 were analyzed next for LD. Alleles at 191, 62J, and D6S446 were in LD with M122K/D6S1590, whereas alleles at 7419 and 137 were in LD with D6S281/AFM234 (Table 2). Indeed, LD was strong, even between D6S281 and 137 (D′, 0.63; χ^2, 89.6), a physical distance of ~142 kb.

Surprisingly, LD was not detected between D6S446 and 7419, a physical distance of 75 kb (Fig. 1 and Table 2). However, the recombination rate between D6S446 and 7419 was approx-

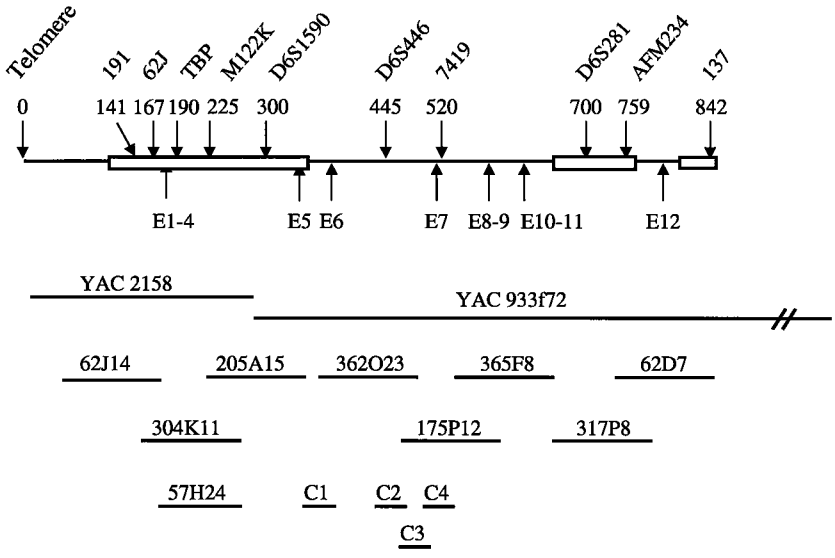

FIG. 1. Physical map of chromosome 6q terminal region. Boxed regions contain nucleotide sequences in the Sanger Genome Centre database. The EST are as follows: E1, D12280; E2, stSG6383; E3, WI-11564; E4, HC5; E5, stSG4590; E6, WI-8780; E7, stSG7419; E8, stSG8638; E9, SGC35081; E10, stSG6307; E11, WI-9654; and E12, stSG21773. Locations on map (0-842) are indicated in kilobase pairs.

imately 10-fold greater than expected (0.67%). This high rate of recombination is likely due to a hot spot of recombination in the region. The possibility of grossly understating the distance between D6S446 and 7419 (75 kb) seems unlikely, because cosmid and BAC clones produce supportive mapping information and are from independent sources (Fig. 1).

To examine association of markers with type 1 diabetes, the transmission disequilibrium test (TDT) was performed by use of marker-haplotypes that demonstrated LD (18 haplotypes) (Table 2). Subsequently, two groups of markers were tested: one that spanned 191 to D6S446 and one that spanned 7419 to 137. Haplotypes in the region of 7419 to 137 (Table 2) were not associated with type 1 diabetes (transmission frequency [T] = 41.3–53.4%). In contrast, multiple haplotypes were associated with type 1 diabetes from the region spanning 191 to D6S446 (T = 58.2–68.1%, $P < 0.05$) (Table 3). Furthermore, the disease-associated haplotypes were randomly transmitted to unaffected siblings (Table 3), which ruled out the possibility of a general transmission disequilibrium to all siblings. Therefore, the recombinant hot spot between D6S446 and 7419 fortuitously defined an outer border, which suggests localization of *IDDM8* within 500 kb of the 6q telomere.

Although 18 haplotypes (Table 2) were tested by TDT, many of these analyses were not independent; therefore, a statistical correction for multiple testing (n = 18) was not used. For example, it can be deduced that an ancestral diabetes–extended haplotype contains the alleles 2-U-5-1-1-4 (191-62J-TBP-M122K-D6S1590-D6S446) (Table 3). Of the six different TDT analyses performed, three reached statistical significance, and the 191-62J-TBP end of the extended haplotype showed the strongest associations. These results suggest that *IDDM8* is located closer to the telomeric side of the haplotype.

Further support for the location of *IDDM8* comes from TDT analyses of the 5-3 haplotype (191-M122K, χ^2 = 5.9) and the 191-5 allele alone (191, χ^2 = 6.3). Allele 5 of microsatellite 191 segregates with type 1 diabetes, whereas M122K-allele 3 does not. These results suggest that *IDDM8* is located within 225 kb of the telomere, between 6qter and M122K. Furthermore, two independent haplotypes appear to be associated with type 1 diabetes susceptibility (191-62J-TBP [2-U-5] and

TABLE 1
Marker allele frequencies and sizes

	Marker allele									
	191	62J	TBP	M122K	D6S1590	D6S446	7419	D6S281	AFM234	137
Allele size										
U	—	0.13 (305)	—	—	—	—	0.70 (170)	—	—	—
L	—	0.87 (305)	—	—	—	—	0.30 (170)	—	—	—
1	—	—	—	0.09 (178)	0.14 (213)	0.50 (201)	—	—	—	0.18 (267)
2	0.12 (216)	—	0.07 (191)	—	0.50 (217)	0.21 (203)	—	0.20 (205)	—	—
3	0.51 (218)	—	0.17 (194)	0.27 (182)	0.32 (219)	—	—	—	0.72 (253)	—
4	0.20 (220)	—	0.25 (197)	0.06 (184)	—	0.07 (207)	—	0.52 (209)	0.09 (255)	—
5	0.08 (222)	—	0.34 (200)	0.11 (186)	—	0.17 (209)	—	0.13 (211)	0.06 (257)	—
6	—	—	0.08 (203)	0.32 (188)	—	—	—	—	—	0.61 (277)
7	—	—	—	0.10 (190)	—	—	—	—	—	—
8	—	—	—	—	—	—	—	—	—	—
9	—	—	—	—	—	—	—	—	—	0.15 (283)

Data are frequencies >0.05 from the parents and allele sizes (bp). U, upper allele; L, lower allele.

TABLE 2
LD among parental marker alleles

Haplotype no.	Marker allele										D′	χ^2
	191	62J	TBP	M122K	D6S1590	D6S446	7419	D6S281	AFM234	137		
1	—	—	—	1	1	—	—	—	—	—	0.59	21.2
2	—	—	—	1	1	4	—	—	—	—	0.74	15.0
3	—	U	—	1	1	—	—	—	—	—	0.72	24.6
4	2	—	—	1	1	—	—	—	—	—	0.73	28.2
5	—	U	5	—	—	—	—	—	—	—	0.46	25.9
6	2	U	5	—	—	—	—	—	—	—	0.49	10.8
7	—	—	—	3	2	—	—	—	—	—	0.20	15.5
8	—	U	—	3	2	—	—	—	—	—	0.72	120
9	4	—	—	3	2	—	—	—	—	—	0.55	71.2
10	5	—	—	3	—	—	—	—	—	—	0.48	16.1
11	—	—	—	6	3	—	—	—	—	—	0.38	53.4
12	3	—	—	6	3	—	—	—	—	—	0.24	41.5
13	—	—	—	—	—	—	—	4	3	—	0.10	10.8
14	—	—	—	—	—	—	U	4	3	—	0.08	16.4
15	—	—	—	—	—	—	—	4	3	6	0.10	15.2
16	—	—	—	—	—	—	—	2	5	—	0.60	19.2
17	—	—	—	—	—	—	—	2	5	9	0.70	37.4
18	—	—	—	—	—	—	—	2	—	9	0.63	89.6

Allele sizes are defined in Table 1. Only data for haplotypes having a χ^2 value >10.6 are shown.

191-M122K [5-3]). It is possible that both haplotypes share a common sequence associated with type 1 diabetes. It is also plausible that different mutations exist on each haplotype, possibly in different parts of the same gene. Nucleotide sequence analysis of both haplotypes will be necessary to determine these possibilities.

TABLE 3
TDT in diabetic patients and healthy siblings

Haplotype no.	Marker allele						T	NT	%T	χ^2
	191	62J	TBP	M122K	D6S1590	D6S446				
D1	—	—	—	1	1	—	40	43	48.2	—
D2	—	—	—	1	1	4	21	16	56.8	—
D3	—	U	—	1	1	—	23	15	60.5	—
D4	2	—	—	1	1	—	26	13	66.7	4.3
S4	2	—	—	1	1	—	8	16	33.3	—
D5	—	U	5	—	—	—	76	51	59.8	4.9
S5	—	U	5	—	—	—	26	31	45.6	—
D6	2	U	5	—	—	—	32	15	68.1	6.1
S6	2	U	5	—	—	—	11	16	40.7	—
D7	—	—	—	3	2	—	149	143	51.0	—
D8	—	U	—	3	2	—	89	64	58.2	4.1
S8	—	U	—	3	2	—	29	26	53.7	—
D9	4	—	—	3	2	—	78	68	53.4	—
D10	5	—	—	3	—	—	48	27	64.0	5.9
S10	5	—	—	3	—	—	9	11	45.0	—
D11	—	—	—	6	3	—	176	137	56.2	—
D12	3	—	—	6	3	—	111	115	49.1	—
D13	5	—	—	—	—	—	62	37	62.6	6.3
S13	5	—	—	—	—	—	10	15	40.0	—
D14	2	—	—	—	—	—	73	77	48.7	—
D15	—	U	—	—	—	—	134	109	55.1	—
D16	—	—	5	—	—	—	220	217	50.3	—
D17	—	—	—	3	—	—	189	166	53.2	—
D18	—	—	—	1	—	—	78	83	48.4	—
D19	—	—	—	—	1	—	103	104	49.8	—
D20	—	—	—	—	—	4	67	63	51.5	—

Allele sizes are defined in Table 1. D, diabetic patients; NT, not transmitted; S, healthy siblings; T, transmitted; %T, % transmitted ([T/T + NT] × 100).

A case-control analysis also was performed for the 191-allele 5. Allele frequencies in 130 unrelated healthy control subjects were compared with those of 199 type 1 diabetic probands (first-diagnosed diabetic in each family). The diabetic and control populations had similar HLA genotypes; the control population was selected on the criteria of having HLA-DR3 and/or DR4 and lacking HLA-DR2. In the case-control analysis, 19.1% of the diabetic probands, but only 7.7% of the healthy control subjects, were positive for allele 5 (χ^2 = 8.2, $P < 0.004$). Although the TDT and case-control tests share overlapping populations (diabetic probands), the two tests were measured according to different parameters: allele frequency in the case-control tests versus allele transmission in the TDT, respectively. The case-control study, therefore, provides additional supportive evidence of a diabetes susceptibility gene that is located between 6qter and M122K.

Previous genetic studies of linkage on chromosome 6q27 could only localize *IDDM8* to a region of perhaps 3–10 million bp (7–9). The association studies described in this article appear to localize *IDDM8* to a region of ~200 kb. The localization of these regions provided the framework to find the specific diabetes susceptibility sequences and to study their roles in the disease pathogenesis. D12280, stSG6383, WI-11564, TBP, and HC5 (human proteasome complex-5 subunit) all mapped within this region and should be considered as candidate genes for type 1 diabetes susceptibility. These results further demonstrate that TDT analyses of haplotypes that are defined by markers in LD are useful tools for better localizing susceptibility genes in complex diseases, such as type 1 diabetes.

RESEARCH DESIGN AND METHODS

All families in this study (188 from the human biological data interchange and 78 from the study described by Owerbach and Gabbay [3]) comprised Caucasian Americans with nondiabetic parents and 1–3 siblings affected with type 1 diabetes. Unaffected healthy siblings also were included in this study (n = 111). In addition, lymphoblastoid cell lines from 130 control subjects, who were unrelated to the aforementioned families, were referred by the National Marrow Donor Program (Minneapolis, MN). These control subjects were selected on the bases of being Caucasian, >18 years old, nondiabetic, HLA-DR3$^+$ and/or HLA-DR4$^+$, and HLA-DR2$^-$ (HLA phenotypes associated with type 1 diabetes).

YAC 2158 (Y1) (11 was purchased from the American Tissue Culture Collection (ATCC no. 99420; Rockville, MD). The 3′ end of this YAC that was subcloned into a PAC vector (57H24) was purchased from Genome Systems (St. Louis, MO). YAC 933f7 (Centre d'Étude du Polymorphisme Humaine) was purchased from the Human Genome Center at the Baylor College of Medicine (Houston, TX). YAC 933f7 is 1,120 kb; the entire length is not shown in Fig. 1. A BAC library was purchased from Research Genetics (Huntsville, AL) and was screened by a protocol provided with the library, and individual BAC clones were purchased from the Human Genome Center at the Baylor College of Medicine (62J14, 304K11, 205A15, 362023, 175P12, 365F8, 317P8, and 62D7). A cosmid library was constructed from YAC 933f7 by use of the SuperCos1 cosmid vector and a protocol from the supplier (Stratagene, La Jolla, CA). End-labeled RNA probes were generated from T3 and T7 sites flanking the cosmid inserts by using a RNA transcription kit from Stratagene. These probes were used to map overlapping cosmids (C1–C4) (Fig. 1) or BACs. In addition, *Not*I-*Hin*dIII fragments containing the end sequences of the BACs were subcloned into pBluescript II plasmid vectors (Stratagene), and the nucleotide sequence was determined by use of the Sequenase version 2.0 sequencing kit from USB (Cleveland, OH). Polymerase chain reaction (PCR) primer sequences were identified and were used subsequently to isolate overlapping BAC clones.

TBP, M122K, D6S1590, and D6S446 were typed in all 266 families. Because of limited amounts of some DNAs, 62J, 7419, D6S281, AFM234, and 137 were typed in 226 families, and 191 was typed in only 202 families. Nucleotide sequences of the PCR primers that were used can be found in the online appendix at www.diabetes.org/diabetes/appendix.asp. As previously described (12,13), microsatellite and SSCP typing were performed by using 10-µl reactions, a 96-well PCR format, and ^{32}P-labeled dCTP. PCR reactions were for 27 cycles: 1 min at 94°C, 1 min at 55–60°C, and 1 min at 72°C.

LD was calculated between two markers by comparing the expected frequency determined from the individual allele frequencies (multiplied together) with those of actual data of two markers present together on the same haplotype (parents). Only informative haplotypes were ascertained. Because of allele heterozygosity in both parents, and the pattern of segregation of haplotypes in the families, not all alleles could be assigned to haplotypes. Once a pair of markers was shown to be in LD, the specified haplotype was treated as a single unit, and additional markers were tested for LD in the pairwise fashion previously described.

LD was also examined by calculating a D′ value for LD by using a formula previously described (14). D′ values of 1, 0, and –1 indicated that two markers were always associated, randomly associated, and never associated, respectively. D′ values between 0 and +1 indicated intermediate levels of association of the two marker alleles. χ^2 analysis was used to determine the significance of the D′ values. Only markers having LD with χ^2 values >10.6 are shown.

The transmission disequilibrium test used the formula $(T - NT)^2/(T + NT)$ to determine χ^2 (15).

ACKNOWLEDGMENTS

This work was supported by grants from the Juvenile Diabetes Foundation (D.O.) and grants from the Harry B. and Aileen B. Gordon Foundation to Dr. Kenneth H. Gabbay.

I thank Krystal Davis, Jennifer Watt, Linda Wible, and Jason Sweeton for their excellent technical assistance; and Drs. Janet Hegland (National Marrow Donor Program, Minneapolis, MN) and Kenneth H. Gabbay for generously providing lymphoblastoid cell lines for the analyses. Dr. David Powell of the Baylor College of Medicine is acknowledged for helpful discussions.

REFERENCES

1. Davies JL, Kawaguchi Y, Bennett ST, Copeman JB, Cordell HJ, Pritchard LE, Reed PW, Gough SCL, Jenkins SC, Palmer SM, Balfour KM, Rowe BR, Farrall M, Barnett AH, Bain SC, Todd JA: A genome-wide search for human type 1 diabetes susceptibility genes. *Nature* 371:130–136, 1994
2. Hashimoto L, Habita C, Beressi JP, Delepine M, Besse C, Cambon-Thomsen A, Deschamps I, Rotter JI, Djoulah S, James MR, Froguel P, Weissenbach J, Lathrop GM, Julier C: Genetic mapping of a susceptibility locus for insulin-dependent diabetes mellitus on chromosome 11q. *Nature* 371:161–164, 1994
3. Owerbach D, Gabbay KH: The search for type 1 diabetes susceptibility genes: the next generation. *Diabetes* 45:544–551, 1996
4. Merriman T, Twells R, Merriman M, Eaves I, Cox R, Cucca F, McKinney P, Shield J, Baum D, Bosi E, Pozill P, Nistico L, Buzzetti R, Joner G, Ronningen K, Thorsby E, Undlien D, Pociot F, Nerup J, Bain S, Barnett A, Todd J: Evidence by allelic association-dependent methods for a type 1 diabetes polygene (*IDDM6*) on chromosome 18q21. *Hum Mol Genet* 6:1003–1010, 1997
5. Merriman TR, Eaves IA, Twells RCJ, Merriman ME, Canoy PAC, Muxworthy CE, Hunter KMD, Cox RD, Cucca F, McKinney PA, Shield JPH, Baum JD, Tuomilehto J, Tuomilehto-Wolf E, Ionesco-Tirgoviste C, Joner G, Thorsby E, Undlien DE, Pociot F, Nerup J, Ronningen KS, Bain SC, Todd JA: 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, 1998
6. Reed P, Cucca F, Jenkins S, Merriman M, Wilson A, McKinney P, Bosi E, Joner G, Ronningen K, Thorsby E, Undlien D, Merriman T, Barnett A, Bain S, Todd J: Evidence for a type 1 diabetes susceptibility locus (*IDDM10*) on human chromosome 10p11-q11. *Hum Mol Genet* 6:1011–1016, 1997
7. Luo DF, Bui MM, Muir A, Maclaren NK, Thomson G, She J-X: Affected-sib-pair mapping of a novel susceptibility gene to insulin-dependent diabetes mellitus (*IDDM8*) on chromosome 6q25–27. *Am J Hum Genet* 57:911–919, 1995
8. Davies JL, Cucca F, Goy JV, Atta ZA, Merriman ME, Wilson A, Barnett AH, Bain SC, Todd JA: Saturation multipoint linkage mapping of chromosome 6q in type 1 diabetes. *Hum Mol Genet* 5:1071–1074, 1996
9. Luo DF, Buzzetti R, Rotter JI, Maclaren NK, Raffel LJ, Nistico L, Giovannini C, Pozzilli P, Thomson G, She J-X: Confirmation of three susceptibility genes to insulin-dependent diabetes mellitus: *IDDM4*, *IDDM5* and *IDDM8*. *Hum Mol Genet* 5:693–698, 1996
10. Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J: A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* 380:152–154, 1996

11. Macina RA, Morii K, Hu X-L, Negorev DG, Spais C, Ruthig LA, Riethman HC: Molecular cloning and RARE cleavage mapping of human 2p, 6q, 8q,12q and 18q telomeres. *Genome Res* 5:225–232, 1995
12. Owerbach D, Gabbay KH: The HOXD locus (2q31) is linked to type I diabetes: interaction with chromosome 6 and 11 disease susceptibility genes. *Diabetes* 44:132–136, 1995
13. Owerbach D, Naya FJ, Tsai M-J, Allander SV, Powell DR, Gabbay KH: Analysis of candidate genes for susceptibility to type 1 diabetes: a case-control and family-association study of genes on chromosome 2q21-35. *Diabetes* 46:1069–1074, 1997
14. Devlin B, Risch N: A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics* 29:311–322, 1995
15. Spielman RS, McGinnis RE, Ewens WJ: Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 52:506–516, 1993