

Type 2 Diabetes Locus on 12q15

Further Mapping and Mutation Screening of Two Candidate Genes

Arsun Bektas, Jennifer N. Hughes, James H. Warram, Andrzej S. Krolewski, and Alessandro Doria

We recently reported evidence of a novel type 2 diabetes locus placed on chromosome 12q15 between markers *D12S375* and *D12S1684* (*Diabetes* 48:2246–2251, 1999). Four multigenerational families having logarithm of odds (LOD) scores >1.0 in the original analysis were genotyped for 11 additional markers in this interval to refine this mapping; this allowed us to narrow the linked region to the interval between markers *D12S1693* and *D12S326*. In a multipoint parametric analysis using the VITESSE software, the LOD score for linkage at this location reached 3.1 in one of these families. This interval contains the gene for protein tyrosine phosphatase receptor type R (*PTPRR*)—a protein that may be involved in both insulin secretion and action. After determining *PTPRR* exon-intron structure, we identified several polymorphisms in this gene but no mutation segregating with diabetes. The search for mutations was also negative for carboxypeptidase M (*CPM*)—another candidate gene mapped to this region. In summary, our data provide further evidence for the existence of a type 2 diabetes locus on chromosome 12q15. This locus, however, does not appear to correspond to the *PTPRR* or *CPM*, although a contribution of mutations in regulatory regions cannot be excluded at this time. *Diabetes* 50:204–208, 2001

It is becoming increasingly clear that the etiology of type 2 diabetes is much more heterogeneous than previously thought. Several attempts to identify type 2 diabetes genes by linkage studies have led to conflicting results, indicating that distinct genes are probably involved in different populations (1,2). Even within the same ethnic group, different genes may be involved in different

families (2). The situation does not appear to be simpler for monogenic forms of diabetes, such as those transmitted with an autosomal-dominant pattern of inheritance. Six distinct genes have been described for the best-known form of this group—maturity-onset diabetes of the young (MODY) (3,4)—and several additional loci probably exist. In France and England, ~25% of MODY families have diabetes unaccounted for by known MODY genes, but this proportion seems to be much higher among families with an older age of diabetes diagnosis than classical MODY (5–7). We recently described a novel locus for autosomal-dominant type 2 diabetes on chromosome 12q15, 50 cM from the previously described locus *NIDDM2* (8). All of the evidence of linkage came from four families that had individual LOD scores >1.0. By analyzing obligate recombinants in the two families with highest LOD scores (2.35 and 2.1), we assigned this putative locus to the 6 cM between markers *D12S375* and *D12S1684*. Here we report the further mapping of this diabetes locus and the results of the study of two candidate genes placed in the critical region.

To narrow the linked interval, the four families that had a LOD score >1.0 in the original analysis (families 8, 19, 24, and 32 [8]) were genotyped for 11 microsatellite markers located between *D12S375* and *D12S1684* (Fig. 1). Included were 27 affected and 25 nonaffected family members. The 11 markers were chosen from contig WC12.4 of the W/MIT integrated map (9). Seven of the microsatellites were also in the Marshfield genetic map (Fig. 1). Within each family, all affected members shared at least one portion of this region identically by descent (Fig. 2A), although the shared haplotype differed among families. In family 8, which had a maximum LOD score (Z_{\max}) of 2.35 in the original analysis, the shared haplotype went from *D12S1693* on the centromeric side to *D12S326* on the telomeric side (Fig. 2A). The LOD score for linkage at this location peaked to 3.1 (Fig. 2B). The shared haplotype in family 24, which had a Z_{\max} of 2.15, further moved the telomeric boundary, narrowing the linked region to the interval between *D12S1693* and *D12S1347* (Fig. 2A and B). No additional information was provided by family 19, in which no recombinants were identified. The affected haplotype of family 32 did not overlap with the critical interval defined by families 8 and 24, but extended instead from *D12S1693* toward the centromere, with the LOD score peaking between *D12S1702* and *D12S375* ($Z_{\max} = 1.5$). A possible explanation for this finding is that there are two diabetes loci in this region: one operating in family 8, placed between *D12S1693* and *D12S326*, and the other in family 32, centromeric to *D12S1693*, with diabetes in families 19 and 24

From the Section on Genetics and Epidemiology (A.B., J.N.H., J.H.W., A.S.K., A.D.), Research Division, Joslin Diabetes Center; and the Department of Medicine (A.B., A.S.K., A.D.), Harvard Medical School, Boston, Massachusetts.

Address correspondence and reprint requests to Alessandro Doria, MD, PhD, Section on Genetics and Epidemiology, Joslin Diabetes Center, One Joslin Place, Boston, MA 02215. E-mail: adoria@joslin.harvard.edu.

Received for publication 10 May 2000 and accepted in revised form 25 August 2000.

The nucleotide sequences reported in this article have been submitted to the GenBank Data Bank with accession numbers AF262940–AF262947 and AF263016–AF263029.

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

CP, carboxypeptidase; LOD, logarithm of odds; MODY, maturity-onset diabetes of the young; PCR, polymerase chain reaction; *PTPRR*, protein tyrosine phosphatase receptor type R; Z_{\max} , maximum LOD score.

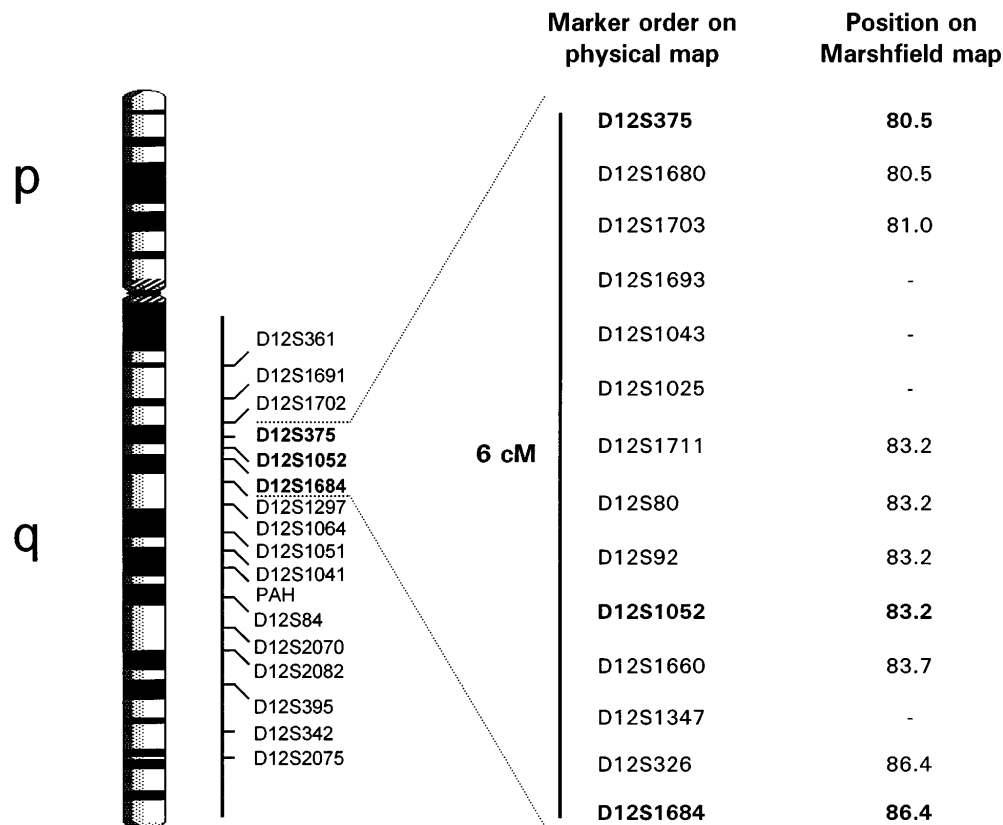


FIG. 1. Microsatellite markers that were genotyped in the study. On the left are markers that were used for the original mapping (8); on the right are those used in the new analysis. Markers that were used in both the original and new analyses are indicated in bold. The marker order is that of contig WC12.4 of the WI/MIT integrated map.

being attributable to either locus. Indeed, it is not unusual that two adjacent loci responsible for the same disorder are initially mapped as a single locus (10). Another possibility is that the recombinants defining either the centromeric boundary in family 8 or the telomeric boundary in family 32 are actually phenocopies. Finally, because family 32 has a Z_{\max} of only 1.5, it is also possible that the linkage observed in this family is entirely due to chance.

One gene placed in the *D12S1693*–*D12S326* interval that caught our attention is protein tyrosine phosphatase receptor type R (*PTPRR*, also known as *NC-PTPCOM1*). Tyrosine phosphorylation is a key component of the cascade of reactions linking the insulin receptor to insulin action in peripheral tissues (11). Furthermore, a tyrosine phosphatase of the same class as *PTPRR* has been shown to be associated with insulin-containing vesicles in β -cells (12). Another interesting aspect was that the major site for *PTPRR* expression was the brain, followed by pancreatic islets and lung (Fig. 3). This distribution pattern is remarkably similar to that of another gene (*NEUROD1*) that has been found to be mutated in MODY and type 2 diabetes (4). Thus, *PTPRR* was a plausible candidate gene for diabetes in these families. The *PTPRR* gene had been assigned to YAC clones 959-C-8 and 788-E-12, which also contain *D12S1025*, but its exon-intron structure had not been resolved. After isolating and sequencing three overlapping BAC clones containing the whole gene, we identified 14 exons ranging in size between 80 and 473 bp (Fig. 4A and Table A1, which can be found in an on-line appendix at www.diabetes.org/diabetes/appendix.htm). Mutation screening of the coding sequence of

probands from the four linked families and other 12 families in which linkage could not be excluded (LOD score > –2.0) revealed several polymorphisms, one of which changed the amino acid sequence (Arg³¹⁴→Lys [Table 1 and Table A2, which can be found in an on-line appendix at www.diabetes.org/diabetes/appendix.htm]). However, no mutations segregating with diabetes were identified. Furthermore, the Arg³¹⁴→Lys polymorphism was similarly frequent among the 32 original family probands (0.266), 173 subjects with type 2 diabetes (0.266), and 181 nondiabetic control subjects (0.226), making *PTPRR* unlikely as a type 2 diabetes locus in these families.

Because there was evidence that the linked interval might extend centromeric to *D12S1693*, or that another diabetes gene might be located beyond this marker, we also analyzed a candidate gene in the vicinity of *D12S375*. In the National Center for Biology Information GeneMap 99 (13), the carboxypeptidase (*CP*) *M* gene was mapped to this region, and we confirmed this location by assigning it to CEPH YAC clones 916-C-11 and 883-H-12, which also contain *D12S375*. *CPM* is a membrane-bound enzyme belonging to the regulatory carboxypeptidase subfamily—a class of proteins that have been implicated in the processing and sorting of polypeptide hormones, either at the site of hormone production or in target tissues (14,15). Of note, a mutation in a member of this family (*CPE*) has been shown to be responsible for diabetes in the *fat/fat* mouse model (16). After isolating BAC clones containing the *CPM* gene, we identified nine exons ranging in size from 98 to 240 bp (Fig. 4B and Table A3, which can be found

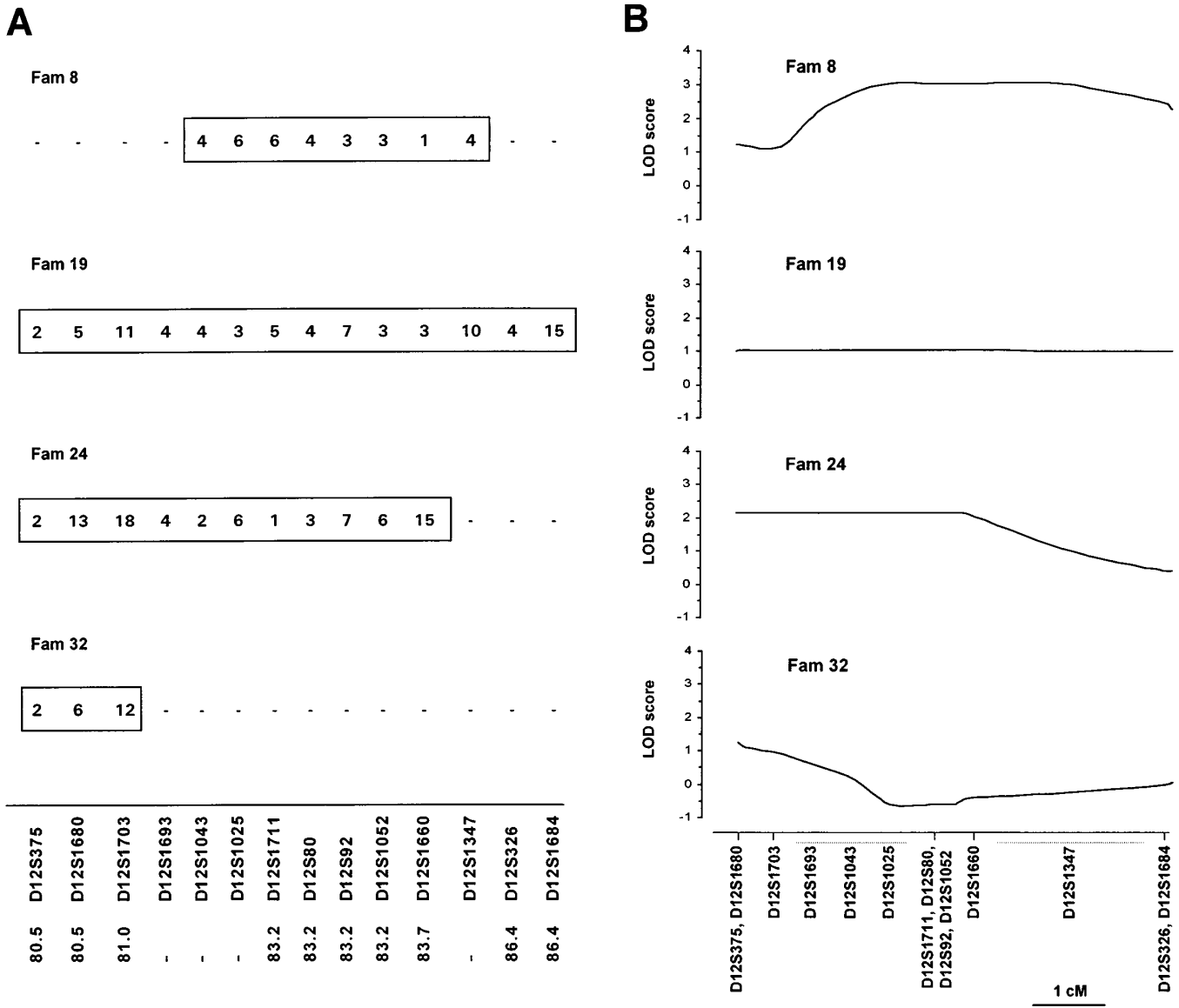


FIG. 2. **A:** Haplotypes shared by all affected individuals within each family. The minus signs indicate regions that did not segregate with diabetes in all affected family members. The allele numbering is arbitrary, but consistent across families. **B:** Multipoint parametric linkage analysis. Marker distances are those indicated in the Marshfield map. For this analysis, markers that are not included in the map (*D12S1703*, *D12S1693*, *D12S1025*, and *D12S1347*) were placed at equal distances from the closest mapped markers.

in an online appendix at www.diabetes.org/diabetes/appendix.asp). In the same 16 families that were screened for *PTPRR*, we identified several silent polymorphisms together with a private Tyr⁹⁹→His mutation that showed incomplete segregation with diabetes in one family (Table 1 and Table A4, which can be found in an online appendix at www.diabetes.org/diabetes/appendix.asp). No mutations, however, were identified in the four linked families, suggesting that *CPM* is unlikely to be the diabetes locus on 12q15.

In summary, we have narrowed the diabetes locus on 12q15 to the region between *D12S1693* and *D12S326*. The LOD score for this location is 3.1 in one family, but it is still possible that the critical interval extends beyond these boundaries. It is unlikely that this locus corresponds to the *PTPRR* or *CPM* genes, although a role of mutations in regulatory regions cannot be excluded at this time. New clues are expected from the upcoming completion of the Human

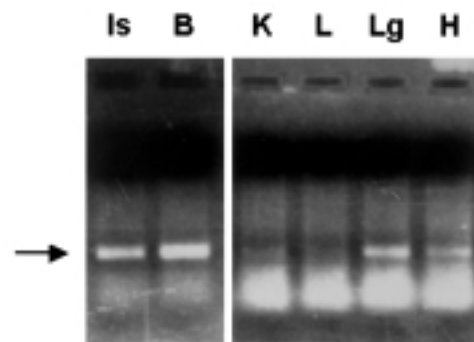


FIG. 3. *PTPRR* expression in different tissues. The presence of *PTPRR* mRNA was determined by standard reverse transcriptase-PCR using GENPAIR primer #31640 (Research Genetics). The arrow indicates the 220-bp band that was amplified from the *PTPRR* cDNA. Is, islets; B, brain; K, kidney; L, liver; Lg, lung; H, heart.

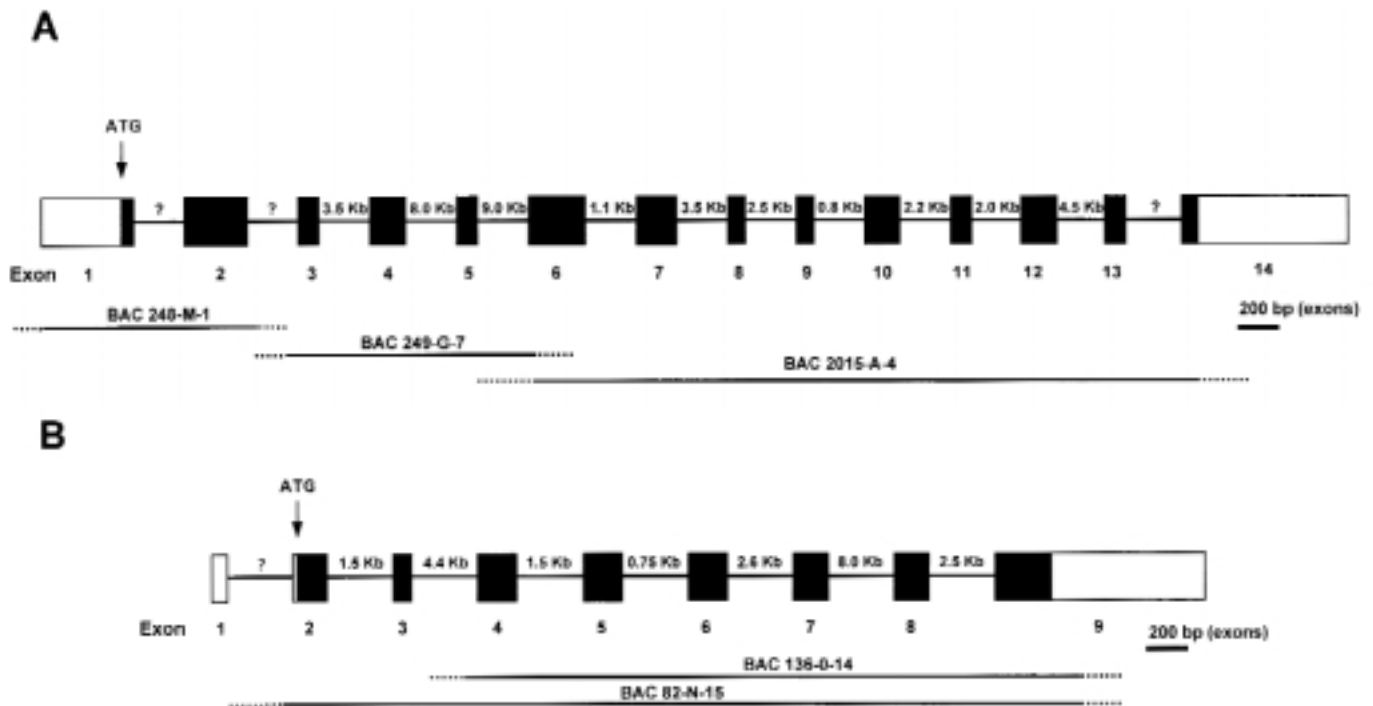


FIG. 4. Exon-intron structure of the *PTPRR* (A) and *CPM* (B) genes. Exons are indicated as translated portions (■) and untranslated portions (□). The beginning of the opening reading frame is indicated by the arrow. BAC clones containing different portions of the genes are indicated below the exons. The sequence of exon-intron junctions can be found in Tables A1 and A3, which can be found in the online appendix. The lengths of *PTPRR* introns 1, 2, 13, and *CPM* intron 1 could not be determined.

Genome Project, which will help us define the exact size of the linked interval, and will provide an inclusive list of the genes placed in this region.

RESEARCH DESIGN AND METHODS

Marker genotyping and linkage analysis. The ascertainment and clinical characteristics of the families with autosomal-dominant type 2 diabetes have been previously described (7). Marker genotypes were determined by ^{32}P -labeled polymerase chain reaction (PCR) followed by denaturing PAGE and autoradiography. To avoid misreading of genotype results, either a standard sequencing marker (Research Genetics, Huntsville, AL) or PCR products of known genotypes were run together with the samples. Genotypes were read separately by two individuals and ambiguous results were repeated. Multipoint parametric linkage analysis was performed using the Vitesse software (17). Because of this software's limitations on the number of markers that can be run at a time, multiple analyses were performed using overlapping sets of four adjacent markers. Individual family LOD scores were calculated assuming an autosomal-dominant mode of inheritance with a disease allele frequency of 0.001, consistent with the rarity of families segregating these forms of diabetes. Similar to previous linkage analyses of MODY, four age-related liability classes (0–10, 11–25, 26–40, and >40 years of age) were assumed. Penetrance of type 2 diabetes in the four age-groups was set at 0.30, 0.50, 0.70, and 0.90, respectively, for the susceptible genotypes DD and Dd. On the basis of the risk of diabetes in the general population, penetrances for the nonsusceptible genotype dd were set to 0.001, 0.005, 0.01, and 0.10. The marker order was that indicated in the Whitehead Institute/Massachusetts Institute of Technology physical map (http://carbon.wi.mit.edu:8000/cgi-bin/contig/phys_map). Intermarker distances (sex-averaged) were those indicated in the Marshfield map (<http://www.marshmed.org/genetics/>). Markers that were not included in the Marshfield map were placed at equal distance from the closest mapped markers. Founder haplotypes were inferred by means of the Genehunter software (18).

YAC library screening. Plates 805–984 of the CEPH Human YAC library were screened by PCR using the CEPH "B" human YAC pools from Research Genetics and gene-specific primers.

Definition of exon-intron boundaries. BAC clone 2015-A-4 was identified as containing the 3' portion of the *PTPRR* gene by means of a basic local alignment search tool (BLAST) search of the GenBank Genome Survey Sequence division, using the *PTPRR* cDNA sequence as a probe. Genomic

BAC clones containing the remaining portions of the *PTPRR* gene and the whole *CPM* gene were identified by PCR screening of the CITB human genomic BAC library (Release IV; Research Genetics) using primers placed in the 3' end of BAC 2015-A-4 and in the 3' untranslated region of the *CPM* cDNA. Direct sequencing of the BAC clones was performed using the Sequitem Excel II DNA sequencing kit (Epicentre, Madison, WI) with ^{32}P -dATP. Intron lengths were determined by long-range PCR (Advantage GC Genomic Polymerase Mix; Clontech, Palo Alto, CA) with primers placed in adjacent exons.

Mutation screening. The coding sequences of *PTPRR* and *CPM* were screened for sequence differences by dideoxy fingerprinting (17). DNA fragments covering the exons and exon-intron boundaries were amplified by PCR from the

TABLE 1
DNA polymorphisms in the *PTPRR* and *CPM* genes among Caucasian individuals

Location	Nucleotide position	Substitution	Amino acid change
<i>PTPRR</i>			
Intron 1	+11	G → A	—
Intron 3	+39	G → A	—
Intron 3	+43	G → A	—
Exon 4	914	T → C	—
Exon 6	1,357	G → A	Arg ³¹⁴ → Lys
<i>CPM</i>			
Exon 3	370	T → C	Tyr ⁹⁹ → His
Exon 3	492	T → C	—
Exon 6	912	G → A	—
Exon 7	1,153	T → C	—

For polymorphisms in exons, nucleotide positions are from the cDNA sequences (GenBank D64053 and J04970). For intron polymorphisms, numbers indicate the positions relative to the splice donor (+) or acceptor (−) site.

DNA of two affected members of families 8 and 24, and one affected member of families 19 and 32, and 12 other pedigrees for which linkage could not be excluded (LOD score > -2.0). Primer sequences and annealing temperatures for each amplification are reported in Tables A2 and A4, which can be found in an on-line appendix. PCR products were purified and subjected to Sanger's dideoxy chain termination reaction using dideoxy GTP in a 10- μ l reaction as described by Sarkar et al. (19). To increase the sensitivity, reactions were performed twice, with the forward and the reverse primer. After adding 20 μ l stop/denaturing solution (7 mol/l urea, 50% formamide, 0.5% bromophenol blue, and 0.5% xylene cyanol) and heating the samples at 95°C for 5 min, 4 μ l were electrophoresed overnight in a nondenaturing 0.75 \times MDE (mutation detection enhancement) gel in 0.5 \times TBE (Tris borate EDTA) on a sequencing apparatus at a constant power of 6 watts at room temperature. Dried gels were autoradiographed overnight. Allele frequencies were determined in the 32 original family probands, 173 Joslin's patients with type 2 diabetes, and 181 non-diabetic control subjects by PCR, dot-blotting, and allele-specific hybridization. The recruitment of the type 2 diabetic individuals and nondiabetic control subjects has been previously described (20). The clinical features of these subjects are reported in Table A5, which can be found in an on-line appendix at www.diabetes.org/diabetes/appendix.htm.

ACKNOWLEDGMENTS

This study was supported by National Institutes of Health grants DK-55523 (A.D.) and DK-47475 (A.S.K.), and Joslin's Diabetes and Endocrinology Research Center Grant DK-36836 (Genetics Core). Human islets for RNA extraction were provided by the Juvenile Diabetes Foundation Center for Islet Transplantation at Harvard Medical School.

Part of this work was presented at the 59th Annual Scientific Sessions (June 2000) of the American Diabetes Association in San Antonio, Texas.

REFERENCES

- Hanis CL, Boerwinkle E, Chakraborty R, Ellsworth DL, Concannon P, Stirling B, Morrison VA, Wapelhorst B, Spielman RS, Gogolin-Ewens KJ, Shepard JM, Williams SR, Risch N, Hinds D, Iwasaki N, Ogata M, Omori Y, Petzold C, Rietzch H, Schroder HE, Schulze J, Cox NJ, Menzel S, Boriraj VV, Chen X, Lim LR, Lindner T, Mereu LE, Wang YQ, Xiang K, Yamagata K, Yang Y, Bell GI: A genome-wide search for human non-insulin-dependent (type 2) diabetes genes reveals a major susceptibility locus on chromosome 2. *Nat Genet* 13:161-166, 1996
- Mahtani MM, Widen E, Lehto M, Thomas J, McCarthy M, Brayer J, Bryant B, Chan G, Daly M, Forsblom C, Kanninen T, Kirby A, Kruglyak L, Munnely K, Parkkonen M, Reeve-Daly MP, Weaver A, Bretton T, Duyk G, Lander ES, Groop LC: Mapping of a gene for type 2 diabetes associated with an insulin secretion defect by a genome scan in Finnish families. *Nat Genet* 14:90-94, 1996
- Froguel P, Velho G: Molecular genetics of maturity-onset diabetes of the young. *Trends Endocrinol Metab* 10:142-146, 1999
- Malecki M, Jhala US, Antonellis A, Fields L, Doria A, Orban T, Saad M, Warram JH, Montminy M, Krolewski AS: Mutations in *NEUROD1* are associated with the development of type 2 diabetes mellitus. *Nat Genet* 23:323-328, 1999
- Vaxillaire M, Boccio V, Philippi A, Vigouroux C, Terwilliger J, Passa P, Beckmann JS, Velho G, Lathrop GM, Froguel P: A gene for maturity onset diabetes of the young maps to chromosome 12q. *Nat Genet* 9:418-423, 1995
- Frayling TM, Bulman MP, Ellard S, Appleton M, Dronsfield MJ, Mackie AD, Baird JD, Kaisaki PJ, Yamagata K, Bell GI, Bain SC, Hattersley AT: Mutations in the hepatocyte nuclear factor-1 α gene are a common cause of maturity-onset diabetes of the young in the U.K. *Diabetes* 46:720-725, 1997
- Doria A, Yang Y, Malecki M, Scotti S, Dreyfus J, O'Keefe C, Wantman M, Orban T, Warram JH, Krolewski AS: Clinical characteristics of early-onset autosomal dominant type 2 diabetes unlinked to known MODY genes. *Diabetes Care* 22:253-261, 1999
- Bektas A, Suprenant ME, Wogan LT, Plengvidhya N, Rich SS, Warram JH, Krolewski AS, Doria A: Evidence of a novel type 2 diabetes locus 50 cM centromeric to *NIDDM2* on chromosome 12q. *Diabetes* 48:2246-2251, 1999
- Hudson TJ, Stein LD, Gerety SS, Ma J, Castle AB, Silva J, Slonim DK, Baptista R, Kruglyak L, Xu S, Hu X, Colbert A, Rosenberg C, Reeve-Daly MP, Rozen S, Hui L, Wu X, Vestergaard C, Wilson K, Bae J, Maitra S, Ganiatsas S, Evans C, DeAngelis M, Ingalls K, Nahf R, Horton L, Oskin M, Collymore A, Ye W, Kouyoumjian V, Zernsteva I, Tam J, Devine R, Courtney D, Renaud M, Nguyen H, O'Connor T, Fizames C, Faure S, Gyapay G, Dib C, Morissette J, Orlin J, Birren B, Goodman N, Weissenbach J, Hawkins T, Foote S, Page D, Lander E: An STS-based map of the human genome. *Science* 270:1945-1954, 1995
- Van Hauwe P, Coucke PJ, Declau F, Kunst H, Ensink RJ, Marres HA, Cremers CW, Djelantik B, Smith SD, Kelley P, Van de Heyning PH, Van Camp G: Deafness linked to DFNA2: one locus but how many genes? *Nat Genet* 21:263, 1999
- Cheatham B, Kahn CR: Insulin action and the insulin signaling network. *Endocr Rev* 16:117-142, 1995
- Wasmeier C, Hutton JC: Molecular cloning of phogrin, a protein-tyrosine phosphatase homologue localized to insulin secretory granule membranes. *J Biol Chem* 271:18161-18170, 1996
- Deloukas P, Schuler GD, Gyapay G, Beasley EM, Soderlund C, Rodriguez-Tome P, Hui L, Matisse TC, McKusick KB, Beckmann JS, Bentolila S, Bihoreau M, Birren BB, Browne J, Butler A, Castle AB, Chiannikulchai N, Clee C, Day PJ, Dehejia A, Dibling T, Drouot N, Duprat S, Fizames C, Fox S, Gelling S, Green L, Harrison P, Hocking R, Holloway E, Hunt S, Keil S, Lijnzaad P, Louis-dit-Sully C, Ma J, Mendis A, Miller J, Morissette J, Muselet D, Nusbaum HC, Peck A, Rozen S, Simon D, Slonim DK, Staples R, Stein LD, Stewart EA, Suchard MA, Thangarajah T, Vega-Czarny N, Webber C, Wu X, Hudson J, Auffray C, Nomura N, Sikela JM, Polymeropoulos MH, James MR, Lander ES, Hudson TJ, Myers RM, Cox DR, Weissenbach JM, Boguski MS, Bentley DR: A physical map of 30,000 human genes. *Science* 282:744-746, 1998
- Tan F, Chan SJ, Steiner DF, Schilling JW, Skidgel RA: Molecular cloning and sequencing of the cDNA for human membrane-bound carboxypeptidase M: comparison with carboxypeptidases A, B, H, and N. *J Biol Chem* 264:13165-13170, 1989
- Cool DR, Normant E, Shen F, Chen HC, Pannell L, Zhang Y, Loh YP: Carboxypeptidase E is a regulated secretory pathway sorting receptor: genetic obliteration leads to endocrine disorders in Cpe (fat) mice. *Cell* 88:73-83, 1997
- Naggert JK, Fricker LD, Varlamov O, Nishina PM, Rouille Y, Steiner DF, Carroll RJ, Paigen BJ, Leiter EH: Hyperproinsulinaemia in obese fat/fat mice associated with a carboxypeptidase E mutation which reduces enzyme activity. *Nat Genet* 10:135-142, 1995
- O'Connell JR, Weeks DE: The VITESSE algorithm for rapid exact multilocus linkage analysis via genotype set-recoding and fuzzy inheritance. *Nat Genet* 11:402-408, 1995
- Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES: Parametric and non-parametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 58:1347-1363, 1996
- Sarkar G, Yoon HS, Sommer SS: Dideoxy fingerprinting (ddF): a rapid and efficient screen for the presence of mutations. *Genomics* 13:441-443, 1992
- Doria A, Caldwell JS, Ji L, Reynet C, Rich SS, Weremowicz S, Morton CC, Warram JH, Kahn CR, Krolewski AS: Trinucleotide repeats at the *rad* locus: allele distributions in NIDDM and mapping to a 3-cM region on chromosome 16q. *Diabetes* 44:243-247, 1995