Brief Genetics Report

Hepatocyte Nuclear Factor–4γ

cDNA Sequence, Gene Organization, and Mutation Screening in Early-Onset Autosomal-Dominant Type 2 Diabetes

Nattachet Plengvidhya, Anthony Antonellis, Lewis T. Wogan, Andrej Poleev, Michael Borgschulze, James H. Warram, Gerhart U. Ryffel, Andrzej S. Krolewski, and Alessandro Doria

The aim of this study was to investigate whether mutations in hepatocyte nuclear factor (HNF)-4γ, a transcription factor homologous to HNF-4α, contribute to the etiology of early-onset type 2 diabetes. Linkage between diabetes and two polymorphic markers at the HNF-4γ locus (D8S286 and D8S548) was evaluated in 32 multigenerational families with early-onset autosomal-dominant type 2 diabetes unlinked to known maturity-onset diabetes of the young genes. Total logarithm of odds (LOD) scores were strongly negative (–50.3 at D8S286 and –46.2 at D8S548), but linkage could not be excluded in 15 families having LOD scores >–2.0. To screen these pedigrees for HNF-4γ mutations, the gene structure was defined. Because reverse transcriptase–polymerase chain reaction experiments indicated that the first 1,674 bp of the published cDNA sequence (3,248 bp) were a cloning artifact, the correct cDNA sequence was determined by 5′ rapid amplification of cDNA ends (RACE) and primer extension assay. Based on the new cDNA sequence (1,731 bp), 11 exons were found. After screening the 5′ flanking region and all coding exons for mutations, we identified several polymorphisms, one of which affected the amino acid sequence (M190I). However, no mutations segregating with diabetes could be found in these families. We conclude that genetic variability in the HNF-4γ gene is unlikely to play a major role in the etiology of early-onset autosomal-dominant type 2 diabetes. ***Diabetes* 48:2099–2102, 1999**

From the Section on Genetics and Epidemiology (N.P., A.A., L.T.W., J.H.W., A.S.K., A.D.), Research Division, Joslin Diabetes Center, Boston, Massachusetts; the Department of Medicine (N.P., A.S.K., A.D.), Harvard Medical School, Boston, Massachusetts; and the Institute of Cell Biology (A.P., M.B., G.U.R.), University of Essen, Essen, Germany.

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 29 March 1999 and accepted in revised form 6 July 1999.

The nucleotide sequences reported in this paper have been submitted to GenBank with accession numbers Z49826, AF133253, and AF133493-504.

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

BAC, bacterial artificial chromosome; HNF, hepatocyte nuclear factor; LOD, logarithm of odds; MODY, maturity-onset diabetes of the young; NS, not significant; OGTT, oral glucose tolerance test; PCR, polymerase chain reaction; RACE, rapid amplification of cDNA ends; RT, reverse transcription; RT-PCR, reverse transcriptase–polymerase chain reaction; UTR, untranslated region.

A recent breakthrough in diabetes research has been the discovery that mutations in islet transcription factors are a frequent cause of maturity-onset diabetes of the young (MODY). MODY is a monogenic form of diabetes characterized by an early onset and an autosomal-dominant pattern of inheritance (1). Of the five MODY genes thus far identified, four—hepatocyte nuclear factor *(HNF)-1*α , *HNF-4*α, insulin promoter factor (*IPF*)*-1*, and *HNF-1*β—correspond to transcription factors (2–5). Although the mechanisms by which mutations in these genes cause diabetes are not completely understood, an impairment of insulin secretion pathways or abnormalities of β-cell development appear to be involved (6). These findings lead to the hypothesis that sequence differences in other, related transcription factors may contribute to forms of MODY or type 2 diabetes for which the genetic determinants are still un-known. HNF-4γ is a recently identified orphan member of the nuclear receptor superfamily that is highly homologous to HNF-4α, reaching almost complete amino acid identity in the DNA and ligand-binding domains (7). This remarkable homology suggests that HNF-4γ may have functions similar to HNF-4α in glucose homeostasis and that genetic variability in this gene may also result in diabetes. In this study, we investigated whether mutations in HNF-4γ are linked to early-onset type 2 diabetes transmitted with an autosomal-dominant pattern of inheritance. In the process, we also corrected the published *HNF-4*γ cDNA sequence and determined the gene exon-intron organization.

The *HNF-4*γ gene has been previously mapped to chromosome 8q (8). Linkage between diabetes and two polymorphic markers (D8S286 and D8S548) located within 2 cM from the *HNF-4*γ gene was evaluated in 32 multigenerational families with early-onset autosomal-dominant type 2 diabetes unlinked to known MODY genes. The recruitment criteria and clinical characteristics of diabetes in these families have been previously described (9). Included in this study were 231 diabetic and 152 nondiabetic family members. The

Downloaded from http://diabetesjournals.org/diabetes/article-pdf/48/10/2099/363710/10512380.pdf by guest on 06 October 2024

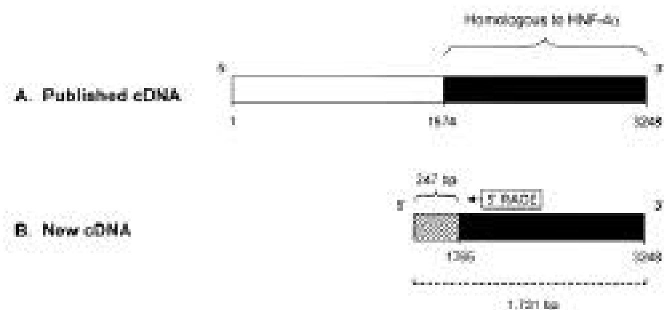

FIG. 1. Schematic representation of the published (*A*) and new (*B*) HNF-4γ cDNA. The black boxes represent the cDNA regions homologous to the *HNF-4α* cDNA. The hatched box represents the new 5′ sequence that was determined by 5′ RACE. The numbering is from the published *HNF-4γ* sequence (7).

mean age at diabetes diagnosis of affected individuals was 37 ± 18 years. Under an autosomal-dominant model with a rare disease allele, the overall multipoint logarithm of odds (LOD) scores were strongly negative for linkage with diabetes (–50.3 at D8S286 and –46.2 at D8S548). Linkage, however, could not be excluded for 15 pedigrees with multipoint LOD scores >–2.0. Positive LOD scores, in one case reaching +2.0, were observed in five families, suggesting that genetic variability at the *HNF-4γ* locus might contribute to diabetes in selected families.

To screen the families with LOD scores >–2.0 for HNF-4γ mutations, the exon-intron structure was determined. The published *HNF-4γ* cDNA sequence consists of 3,248 bp, the first 1,674 of which are not homologous to *HNF-4α* (7) (Fig. 1*A*). These 1,674 bp are 100% identical to sequences on NCBI HTGS clone 257E24 (GenBank AL034424), which has been localized to chromosome 20 rather than 8 and does not contain any *HNF-4γ* cDNA sequence 3′ of position 1,674. This observation, together with the fact that we repeatedly failed to amplify from the 5′ to the 3′ half of the cDNA, indicated that the first 1,674 bp of the published sequence were a cloning artifact. The correct 5′ end of the cDNA was therefore determined by means of 5′ rapid amplification of cDNA ends (RACE) from pancreas RNA. Using a reverse primer at position 1,819 of the published cDNA, we obtained a DNA fragment, the sequence of which is identical to the published cDNA up to position 1,765 and then diverges (Fig. 1*B*). After elongating the fragment with a second round of 5′ RACE, a third RACE experiment did not provide any additional sequence, suggesting that the 5′ end of the cDNA had been reached. The total 5′ sequence determined by 5′ RACE included 247 bp (Fig. 1*B*), resulting in a much shorter cDNA than the published one. As shown in the online appendix A1, the new cDNA consists of 1,731 bp (270 bp of 5′ untranslated region [UTR], 1,227 bp of coding, and 234 bp of 3′ UTR sequence) and predicts a protein of 408 amino acids having 70% identity with HNF-4α1. By reverse transcriptase–polymerase chain reaction (RT-PCR), HNF-4γ was found to be expressed in islets, liver, kidney, brain, and lungs.

After identifying and sequencing bacterial artificial chromosome (BAC) clones containing the whole *HNF-4γ* cDNA, we detected 11 exons spanning about 35 kb (Fig. 2). We also identified an additional exon (exon 2A) corresponding to 91 bp of the published sequence (between position 1674 and 1765), which is homologous to *HNF-4α* but was not present in the RACE sequence (Fig. 1*A* and *B*). This exon was found to be expressed only in the kidney. We also determined 557 bp of the 5′ flanking region, which is extremely GC-rich and does not contain a TATA box (Fig. 3), similar to the *HNF-4α* gene (10). Using promoter prediction software (NNPP at http://www-hgc.lbl.gov/projects/promoter.html), a potential transcription start site was detected 15 bp upstream of the 5′ end of the cDNA that was determined by RACE (Fig. 3). This transcription start site was confirmed by primer extension analysis. Two other potential promoters, however, were detected at positions 112 and 154 bp 5′ of the cDNA end, suggesting that multiple transcripts with different 5′ UTR lengths may exist, as observed for other transcription factor genes (11). As shown in Fig. 3, consensus sequences were found for the same transcription factors that were described in the *HNF-4α* 5′ flanking region, i.e., HNF-1, HNF-6, HNF-3β, activator protein 1 (AP-1), and nuclear factor 1 (NF-1) (10).

The 5′ flanking region, exons, and exon-intron junctions of the *HNF-4γ* gene were screened for mutations in 20 individuals from the 15 families with early-onset type 2 diabetes for which linkage with *HNF-4γ* could not be excluded. The screening was performed by PCR followed by direct sequencing, using the primers provided in the appendix A3. As shown in Table 1, several polymorphisms were identified,

FIG. 2. *HNF-4γ* gene exon-intron organization. Exons that are expressed in pancreas, kidney, liver, brain, and lung are indicated as black (translated portions) and white (untranslated portions) boxes. The hatched box denotes a small exon that was found only in the kidney. The beginning of the opening reading frame is indicated by the arrow. An additional in-frame translation start is present in exon 2A and may determine the presence of an additional 37 amino acids at the HNF-4γ aminoterminus in the kidney. The sequence of exon-intron junctions can be found in the online appendix A2.

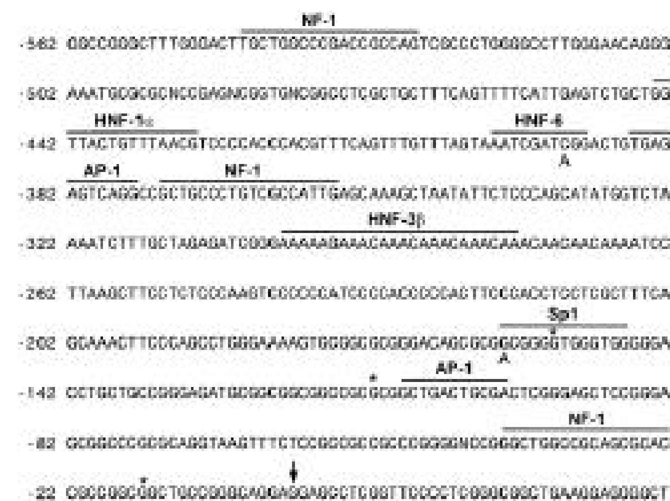

FIG. 3. Sequence of the *HNF-4γ* 5′ flanking sequence. The 5′ end of the cDNA that was determined by 5′ RACE is indicated by the arrow. The asterisks note the transcription start sites predicted by the NNPP software. Potential binding sites for transcription factors that were previously described for *HNF-4α* (10) or that were disrupted by polymorphisms are overlined. The –159 and –393 polymorphisms are indicated by the presence of two bases at the same position.

Downloaded from http://diabetesjournals.org/diabetes/article-pdf/48/10/2099/363710/10512380.pdf by guest on 06 October 2024

TABLE 1
DNA polymorphisms in the HNF-4γ gene among Caucasian individuals

Location	Nucleotide position	Substitution	Amino acid change	Allele frequencies		
				C	DM	FP
Promoter region	–393	C A	—	0.421	0.500	0.460
Promoter region	–159	G A	—	0.416	0.500	0.460
Intron 3	–95	A G	—	0.392	0.456	0.469
Intron 5	+26	C A	—	0.065	0.071	0.125
Exon 6	702	A G	—	0.071	0.060	0.078
Intron 6	+90	T G	—	0.296	0.343	0.281
Exon 7	840	G A	M^{190} I	0.464	0.474	0.500
Intron 7	+15	TT deletion	—	0.388	0.461	0.483
Intron 10	–17	C A	—	0.070	0.059	0.078
Intron 10	–17	C A	—	0.059	0.024	0.048
Exon 11 (3′ UTR)	1,562	A G	—	0.489	0.430	0.435

Nucleotide positions are from Fig. 3 for polymorphisms in the promoter region and from the cDNA in the online appendix for polymorphisms in exons. For intron polymorphisms, numbers indicate the positions relative to the splice donor (+) or acceptor (–) site. C, 96 nondiabetic control subjects; DM, 96 type 2 diabetic subjects; FP, 32 diabetic family probands.

but no mutation segregating with diabetes could be found in these families. Two polymorphisms, in almost complete mutual linkage disequilibrium, were located in the promoter region (G A at –159 and C A at –393) and disrupted potential binding sites for Sp1 and HNF-6, respectively (Fig. 3). However, no significant differences were found in allele frequencies among the 32 family probands, 96 type 2 diabetic subjects, and 96 nondiabetic control subjects (–159A = 0.460 and –393A = 0.460 in family probands vs. 0.500 and 0.500 in type 2 diabetic subjects, and 0.416 and 0.421 in control individuals, not significant [NS]). Another polymorphism (a G-to-A transition in exon 7) resulted in a Met Ile substitution at codon 190 (M190I). Also in this case, however, the allele frequencies in family probands were not significantly different from those in type 2 diabetic subjects or nondiabetic control subjects (190M = 0.500 vs. 0.474 vs. 0.464 in the three groups, respectively, NS). All other polymorphisms did not affect the amino acid sequence or, if they were located in introns, did not disrupt the conserved GT and AG dinucleotides of the splice donor and acceptor sites. No significant differences in allele frequencies were observed among case and control subjects also for these polymorphisms (Table 1).

Our results indicate that genetic variability in the *HNF-4γ* gene does not play a major role in the etiology of early-onset autosomal-dominant type 2 diabetes, although mutations in as-yet-unidentified regulatory regions cannot be excluded at this time. The finding of no association between functional polymorphisms and later-onset type 2 diabetes also suggests that this locus is unlikely to be involved in the development of common forms of type 2 diabetes. It must be emphasized, however, that rare cases of diabetes caused by HNF-4γ mutations cannot be ruled out until this gene is screened for sequence differences in other family collections or in populations of different ethnicity. The information on cDNA and gene structure provided here makes this task possible.

RESEARCH DESIGN AND METHODS

Families. The ascertainment of families with early-onset autosomal-dominant type 2 diabetes has been previously described (9). Diabetes was diagnosed if one of the following criteria was met: *1*) if an individual was treated with insulin or oral agents; *2*) if the results of an oral glucose tolerance test (OGTT) met World Health Organization criteria (12); or *3*) if the level of HbA_{1c} was >7.0% in individuals who declined an OGTT or were not fasting when examined. All families were negative for mutations in the HNF-4α (MODY1) and HNF-1α (MODY3) genes. Analysis with markers HGKCA1 and D7S2428 (flanking the glucokinase locus) also excluded linkage with diabetes (A. Bektas and A.D., unpublished results).

Linkage analysis. Marker genotypes were determined by ^{32}P-labeled PCR followed by denaturing polyacrylamide gel electrophoresis and autoradiography (13). Multipoint parametric analysis was performed using GENEHUNTER (version 1.2.1) software (14), assuming an autosomal-dominant mode of inheritance with a disease allele frequency of 0.001 and four age-related liability classes.

5′ RACE. 5′ RACE was performed from Marathon-Ready pancreas cDNA (Clontech, Palo Alto, CA), using Clontech's Advantage cDNA Polymerase Mix according to the manufacturer's specifications. Three rounds of nested 5′ RACE were performed using forward primers AP1 and AP2 (Clontech) and the following reverse primers: NP/RACE1 (5′-AGTGTTTTCCTGTTGCTCTGTCCCCA-3′) and NP/RACE2 (5′-AGACAGTTGACACCGTTGTCTGTGG-3′), NP/RACE3 (5′-ATA TCTCAGACTGG CGGCTGCTTTCGCT-3′) and NP/RACE4 (5′-TGCTCCTCCA GCGCCTTCTTCCCT-3′), NP/RACE5 (5′-GGCCTCGGGGTATCTGCTTTCAACA-3′) and NP/RACE6 (5′-GGTCCTGAGGCTCAGCCCCT CCTT). The resulting PCR products were subcloned using the TA Cloning Kit (Invitrogen, San Diego, CA) and manually sequenced (AmpliCycle Sequencing kit; Perkin Elmer, Norwalk, CT).

RT-PCR. RNA from different tissues was obtained from Clontech. RT reactions were performed using 1 µg of total RNA, random hexamers as primers, and standard conditions (15). Standard PCR (15) was performed from the RT reaction using forward primers placed in exon 2 (5′-CTGACAGCTTACTCCTTG-3′), and 2A (5′-GGACCCAACTTACACAAC) and a reverse primer in exon 3 (5′-TGGATGCCC CATAGTGTT-3′).

BAC identification and sequencing. The California Institute of Technology human genomic BAC library B (Release IV; Research Genetics, Huntsville, AL) was screened by PCR using primers in the 3′ UTR of the *HNF-4γ* cDNA (F: 5′-AATAGCTGCAAACCAAGC-3′, R: 5′-CCATCTTATAACAGCACC-3′). A positive BAC clone (470-C-10, Fig. 2) was directly sequenced starting from the 3′ end of the cDNA until the last intron was found. Based on this information, new primers to sequence the BAC clone were designed on the cDNA, 5′ of the last exon. This process was repeated until the whole exon-intron structure was resolved. To define exons 1, 2, and 2A, two additional BACs (549-L-2 and 290-I-22, Fig. 2) were identified by screening the same BAC library with primers in exons 3 and 2. Direct sequencing of the BAC clones was performed using the AmpliCycle Sequencing kit (Perkin Elmer) with [^{32}P]-dATP. Intron lengths were determined by long-range PCR (Advantage GC Genomic Polymerase Mix; Clontech) with primers placed in adjacent exons.

Primer extension analysis. Primer extension analysis was performed from 20 µg of human pancreas RNA using the Primer Extension System Kit (Promega, Madison, WI) and primer NP/RACE5. The RT product was electrophoresed on a denaturing polyacrylamide gel and exposed to an autoradiography film (X-OMAT LS; Kodak, Rochester, NY) for 24 h.

Mutation screening. The mutation screening was performed by PCR followed by direct sequencing. A total of 11 DNA fragments were amplified by standard PCR using the primers and annealing temperatures reported in the appendix A3. Fragments were purified from agarose gel and sequenced using the Perkin Elmer's ABI

Downloaded from http://diabetesjournals.org/diabetes/article-pdf/48/10/2099/363710/10512380.pdf by guest on 06 October 2024

Dye Terminator Cycle Sequencing Ready Reaction Kit (PE Applied Bio Systems, Foster City, CA) on an ABI Prism Model 377. Allele frequencies were determined in the 32 family probands, 96 Joslin's type 2 diabetic patients, and 96 nondiabetic control subjects by PCR, dot-blotting, and allele-specific hybridization (13). The recruitment and clinical characteristics of the type 2 diabetic individuals and nondiabetic control subjects have been previously described (16).

ACKNOWLEDGMENTS

This study was supported by National Institutes of Health Grants DK-55523 (to A.D.) and DK-47475 to (A.S.K.), Joslin's Diabetes Endocrinology Research Center Grant DK-36836 (Genetics Core), and Deutsche Forschungsgemeinschaft Grant RY 5/4–1 (to G.U.R.). N.P. was supported by a Fellowship Grant from Siriraj Hospital (Mahidol University, Bangkok, Thailand). We thank Michelle Suprenant for technical help and Dr. Susan Bonner-Weir (Joslin Diabetes Center) for the kind gift of islet RNA. Part of this work was presented at the 59th Scientific Sessions of the American Diabetes Association in San Diego, CA, in June 1999.

REFERENCES

1. Tattersall RB, Fajans SS: A difference between the inheritance of classical juvenile-onset and maturity-onset type diabetes of young people. *Diabetes* 24:44–53, 1974
2. Yamagata K, Oda N, Kaisaki PJ, Menzel S, Furuta H, Vaxillaire M, Southam L, Cox RD, Lathrop GM, Boriraj VV, Chen X, Cox NJ, Oda Y, Yano H, Le Beau MM, Yamada S, Nishigori H, Takeda J, Fajans SS, Hattersley AT, Iwasaki N, Hansen T, Pedersen O, Polonsky KS, Turner RC, Velho G, Chevre JC, Froguel P, Bell GI, et al: Mutations in the hepatocyte nuclear factor-1alpha gene in maturity-onset diabetes of the young. *Nature* 384:455–458, 1996
3. Yamagata K, Furuta H, Oda N, Kaisaki PJ, Menzel S, Cox NJ, Fajans SS, Signorini S, Stoffel M, Bell GI: Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young. *Nature* 384:458–460, 1996
4. Stoffers DA, Ferrer J, Clarke WL, Habener JF: Early-onset type-II diabetes mellitus (MODY4) linked to IPF1. *Nat Genet* 17:138–139, 1997
5. Horikawa Y, Iwasaki N, Hara M, Furuta H, Hinokio Y, Cockburn BN, Lindner T, Yamagata K, Ogata M, Tomonaga O, Kuroki H, Kasahara T, Iwamoto Y, Bell GI: Mutation in hepatocyte nuclear factor-1 beta gene (TCF2) associated with MODY. *Nat Genet* 17:384–385, 1997
6. Stoffel M, Duncan SA: The maturity-onset diabetes of the young (MODY1) transcription factor HNF4alpha regulates expression of genes required for glucose transport and metabolism. *Proc Natl Acad Sci U S A* 94:13209–13214, 1997
7. Drewes T, Senkel S, Holewa B, Ryffel GU: Human hepatocyte nuclear factor 4 isoforms are encoded by distinct and differentially expressed genes. *Mol Cell Biol* 16:925–931, 1996
8. Vaisse C, Kim J, Espinosa R III, Le Beau MM, Stoffel M: Pancreatic islet expression studies and polymorphic DNA markers in the genes encoding hepatocyte nuclear factor-3α, -3β, -3γ, -4γ, and -6. *Diabetes* 46:1364–1367, 1997
9. Doria A, Yang Y, Malecki M, Scotti S, Dreyfus J, O'Keeffe C, Wantman M, Orban T, Warram JH, Krolewski AS: Phenotypic characteristics of early-onset autosomal-dominant type 2 diabetes unlinked to known maturity-onset diabetes of the young (MODY) genes. *Diabetes Care* 22:253–261, 1999
10. Furuta H, Iwasaki N, Oda N, Hinokio Y, Horikawa Y, Yamagata K, Yano N, Sugahiro J, Ogata M, Ohgawara H, Omori Y, Iwamoto Y, Bell GI: Organization and partial sequence of the hepatocyte nuclear factor-4α/MODY1 gene and identification of a missense mutation, R127W, in a Japanese family with MODY. *Diabetes* 46:1652–1657, 1997
11. Semenza GL: *Transcription Factors and Human Disease.* Oxford, U.K., Oxford University Press, 1998
12. Harris MI: Impaired glucose tolerance in the US population. *Diabetes Care* 12: 464–474, 1989
13. Dracopoli NC, Haines JL, Korf BR, Moir DT, Morton CC, Seidman CE, Seidman JG, Smith DR: *Current Protocols in Human Genetics*. New York, John Wiley & Sons, 1994
14. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES: Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 58:1347–1363, 1996
15. Birren B, Green ED, Klapholz S, Myers RM, Roskams J: *Genome Analysis: A Laboratory Manual.* Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press, 1997
16. 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

Downloaded from http://diabetesjournals.org/diabetes/article-pdf/48/10/2099/363710/10512380.pdf by guest on 06 October 2024