

A gene for maturity onset diabetes of the young (MODY) maps to chromosome 12q

Martine Vaxillaire¹, Valérie Boccio¹, Anne Philippi¹, Corinne Vigouroux¹, Joe Terwilliger², Philippe Passa³, Jacques S. Beckmann⁴, Gilberto Velho⁵, G. Mark Lathrop^{2,5} & Philippe Froguel^{1,3}

Maturity-onset diabetes of the young (MODY) is a subtype of non-insulin dependent diabetes mellitus, with early age of onset. MODY is genetically heterogeneous, associated with glucokinase mutations and a locus on chromosome 20q; in about 50% of cases, its genetic background is unknown. We have studied 12 families in which MODY is unlinked to either glucokinase or chromosome 20q markers, and find significant evidence for linkage with microsatellite markers on chromosome 12q, most likely within a 7 centimorgan interval bracketed by *D12S86* and *D12S342*. The disease was estimated to be linked to this chromosome region in approximately 50% of families in a heterogeneity analysis. These MODY patients exhibit major hyperglycaemia with a severe insulin secretory defect, suggesting that the causal gene is implicated in pancreatic β -cell function.

¹CNRS EP10, Institut Pasteur de Lille (Present address: CEPH, 27 rue Juliette Dodu), 75010 Paris, France

²The Wellcome Trust Centre for Human Genetics, Windmill Road, Oxford OX3 7BN, UK

³Service d'Endocrinologie, Hôpital Saint-Louis, 1 avenue Claude Vellefaux, 75010 Paris, France

⁴Généthon, 1 rue de l'Internationale BP60, 91002 Evry, France

⁵INSERM U358, 12 rue de la Grange aux Belles, 75010 Paris, France

Correspondence should be addressed to P.F.

Maturity-onset diabetes of the young (MODY) is a form of familial non insulin dependent diabetes mellitus (NIDDM), characterized by an early age of onset (childhood, adolescence or young adulthood under 25 years) and autosomal dominant mode of inheritance¹. MODY occurs worldwide and has been found in 13% of the Caucasian NIDDM families we have recently collected in France². It is now recognized that MODY is a genetically heterogeneous disorder. Genetic studies led to the identification of two MODY genes, an as yet unidentified gene tightly linked to the adenosine deaminase gene (*ADA*) on chromosome 20q in one large Michigan family of German origin³ and the glucokinase gene (*GCK*) on chromosome 7 (refs 4,5). Glucokinase mutations appear to be the most common form of MODY in France, present in 56% of our families⁶. So far, we have observed 31 different *GCK* mutations in 35 unrelated families. In contrast, conclusive evidence for linkage to chromosome 20 in the region of *ADA* has yet to be demonstrated in French families.

In order to identify a third MODY susceptibility locus (*MODY3*) that could account for the disease, we have studied twelve French MODY families (Fig. 1), in which diabetes was not genetically linked to *GCK* or to the *ADA* locus. We also excluded linkage to a number of candidate genes involved in glucose homeostasis⁷. In view of our negative results, we subsequently carried out a genome-wide segregation analysis of highly informative microsatellite markers in these 12 families. We now report the localization of a gene for MODY to the long arm of chromosome 12 in six families.

Linkage analysis

The marker loci used in this study consisted exclusively of polymorphic (CA)_n repeats, also termed microsatellites, uniformly distributed across the entire autosomal genome. We analysed 97 microsatellite markers for linkage to MODY in the families selected. These markers⁸ were chosen because of their elevated heterozygosity index (>70%) and even spacing, with a genetic distance of about 20–40 centimorgans (cM) between each. To facilitate microsatellite genotyping, we used a procedure derived from the multiplex sequencing technique of Church and Kieffer-Higgins⁹. Genetic analyses led to the exclusion of 78% of the map from Weissenbach *et al.*⁸ (data not shown).

A chromosome 12q marker, *D12S76* (AFM010th7), exhibited significant evidence of linkage with the MODY phenotype. The maximum lod score was 3.11 at q (max)=0.16 (Table 1). The homogeneity test showed significant evidence that the families were in two groups (χ^2 1df=5.3, $P<0.05$) with 55% of the families estimated to show linkage. The lod score under heterogeneity was 4.27. Two other markers from the same region gave lod scores greater than 2 with or without the assumption of homogeneity (Fig. 2 and Table 2). The conclusions were not modified under a model with high penetrance (see Methodology).

Multipoint analysis and genetic heterogeneity

Based on pairwise results, and the analysis of the disease locus simultaneously with two marker loci (results not shown), we localized the former to a 10 cM interval spanned by the markers *D12S86*–*D12S76*–*D12S342*. The location score results for these three markers and the disease locus are shown in Fig. 3. Under heterogeneity,

the maximum location score was 5.8 (logarithm base 10) at zero recombination from *D12S76*. The 1-*lod* unit interval of confidence extends to a region of approximately 7 cM around *D12S76*, and all placements of the disease locus outside of the interval spanned by *D12S86*–*D12S342* were rejected with odds greater than 1800:1. The likelihood ratio test comparing the maximum location scores under heterogeneity and homogeneity was 12.9 and the estimated proportion of linked families was 0.55.

Discussion

Our study provides strong evidence of linkage between early-onset NIDDM (MODY subtype) and microsatellite

markers on chromosome 12q. It confirms that this monogenic form of type 2 diabetes is genetically heterogeneous. There are now three genes to be implicated in causing chronic hyperglycaemia in MODY patients, but their prevalence seem to be different. The chromosome 20q gene is linked to MODY in only one family so far in the United States. Glucokinase is responsible of about 50% of MODY cases in France, but may be less prevalent in England (Hattersley, pers. comm.) and Japan (Iwasaki *et al.*, unpublished results). From this study, an estimated 55% of 12 families unlinked to *GCK* or chromosome 20q shows linkage with the *MODY3* gene on chromosome 12q, corresponding to about a

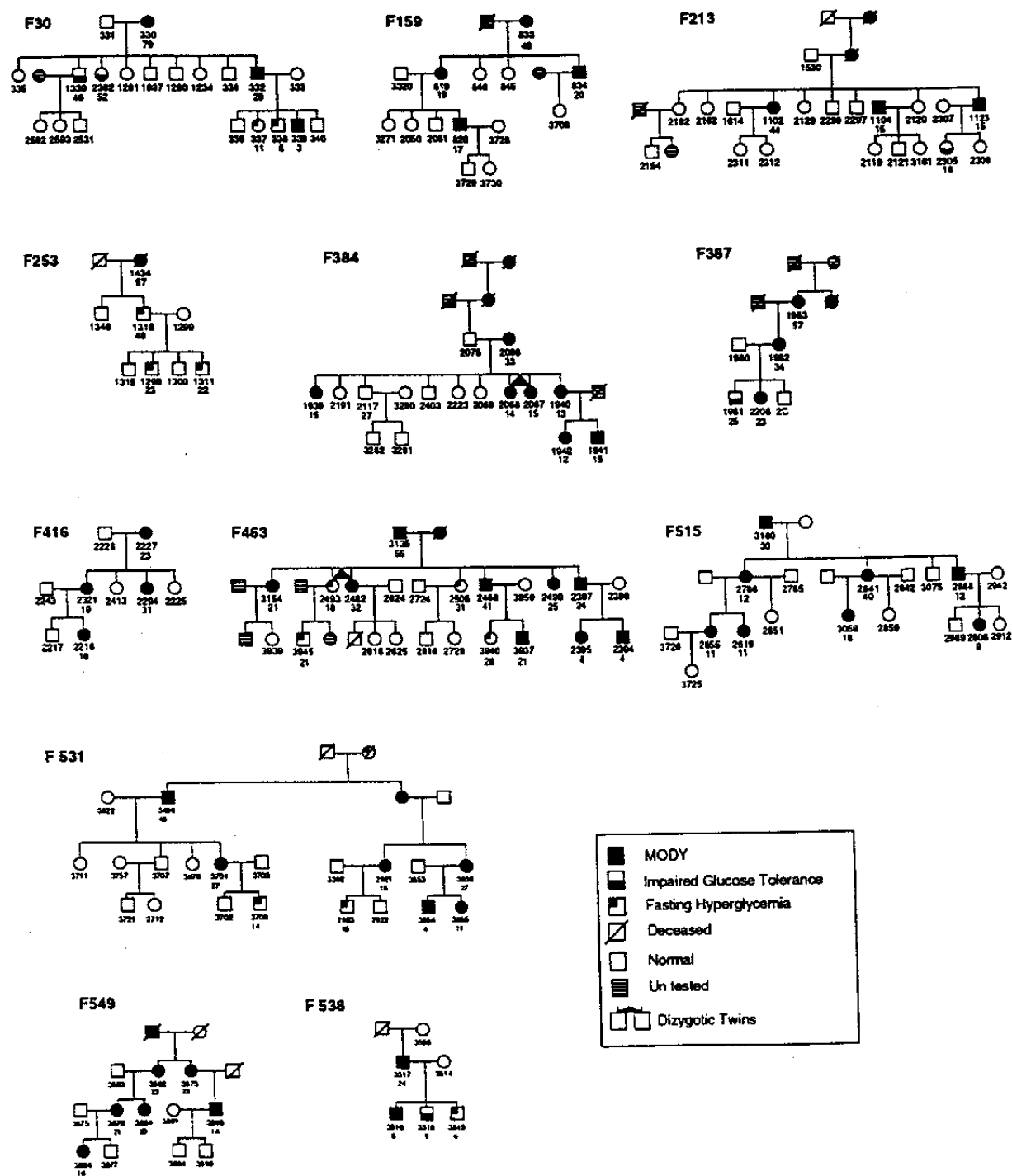


Fig. 1 Pedigrees of 12 MODY families. The numbers under the symbols correspond with CEPH identification numbers. The age at diagnosis and status of each subject are shown. Squares denote male family members and circles are female family members. The first nine pedigrees have been already published⁷.

Table 1 Pairwise lod scores (Z) between *D12S76* and *MODY* in 12 families

	Recombination fraction								
	0.00	0.02	0.04	0.06	0.08	0.10	0.16	0.20	0.30
F30	-3.47	-2.31	-1.75	-1.41	-1.16	-0.97	-0.59	-0.43	-0.18
F159	0.75	0.72	0.68	0.65	0.62	0.58	0.48	0.40	0.22
F213	2.06	1.99	1.92	1.85	1.77	1.70	1.46	1.29	0.83
F253	-3.15	-1.70	-1.33	-1.10	-0.92	-0.78	-0.49	-0.36	-0.14
F384	1.38	1.33	1.28	1.22	1.16	1.11	0.93	0.81	0.48
F387	-2.24	-1.14	-0.86	-0.70	-0.58	-0.49	-0.31	-0.23	-0.10
F416	0.39	0.37	0.35	0.33	0.31	0.28	0.22	0.18	0.09
F463	0.02	0.04	0.07	0.10	0.13	0.15	0.20	0.20	0.15
F515	1.68	1.62	1.57	1.51	1.45	1.39	1.20	1.07	0.71
F531	-3.70	-1.73	-1.23	-0.93	-0.73	-0.58	-0.31	-0.20	-0.07
F538	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F549	0.65	0.61	0.57	0.53	0.49	0.45	0.33	0.26	0.11
Total lod score	-5.62	-0.20	1.26	2.05	2.53	2.82	3.11	2.99	2.12

quarter of all *MODY* families collected in France. Further genetic and epidemiological studies will shed light on the exact prevalence of the *MODY* 3 subtype in different populations. Finally, our results imply that one or more additional *MODY* susceptibility loci remain to be identified.

Our linkage data in 12 *MODY* families show a maximal lod score >3.00 with *D12S76*, and six families present a positive lod score value at a recombination fraction of 0.00. Furthermore, statistically significant evidence of genetic heterogeneity is observed ($P < 0.02$), with a maximal lod score under heterogeneity of 4.27 at a recombination fraction of 0.00. Location score analysis under heterogeneity condition indicates that the most

likely location of *MODY3* lies between *D12S86* and *D12S342* and spans 7 cM. The maximum location score is obtained at the *D12S76* locus (Fig. 3).

MODY linked to chromosome 12q markers shows an incomplete penetrance for the disease status, as was already observed for the chromosome 20q gene and to a lesser extent for *GCK* (Velho *et al.* unpublished results)¹⁰. Eleven subjects from five families present with normal glucose tolerance while having the diabetes susceptibility haplotype at locus *D12S86-D12S76* which is shared by their diabetic relatives (data not shown). Seven of them are younger than 25 years (range 11–21), while the others are aged 29, 33, 34 and 40 years, respectively.

The clinical profile of *MODY* appears to vary between different genetically defined subgroups of patients. All *GCK* variants discovered so far exhibit a decreased enzymatic activity⁵. In patients harbouring a *GCK* variant, clinical studies have shown reduced β -cell responsiveness to glucose with a shift to the right in the glucose/insulin secretion rate dose response curve, and a >50% reduction in insulin secretion rate for a given glucose level¹¹. *GCK* mutations are also associated with a relatively mild form of type 2 diabetes, and less than 1/3 of these patients have severe hyperglycaemia⁶. In contrast, the clinical phenotype resulting from mutations in the *MODY* gene on chromosome 20q closely resembles late-onset type 2 diabetes in its natural history, with subjects rapidly progressing from impaired glucose tolerance to overt diabetes¹. Indeed, the majority of diabetic subjects carrying the chromosome 20q susceptibility haplotype are treated with oral hypoglycaemic agents or insulin.

To determine the clinical profile associated with *MODY3*, we compared data on 33 patients from families which gave positive evidence of linkage to chromosome 12q markers with data from 212 patients from *GCK*-linked families and 40 patients from other families. The comparison suggests that *MODY3* patients present with a more severe form of disease (Table 3). Indeed, the prevalence of overt diabetes, as compared to the prevalence of impaired glucose tolerance, is significantly higher in *MODY3* than in the other two groups. Furthermore, *MODY3* subjects are more often treated by oral hypoglycaemic agents and insulin than the other patients. Fasting and 2 h post-oral glucose load plasma glucose values in the *MODY3* affected individuals are significantly higher than in patients from the other two

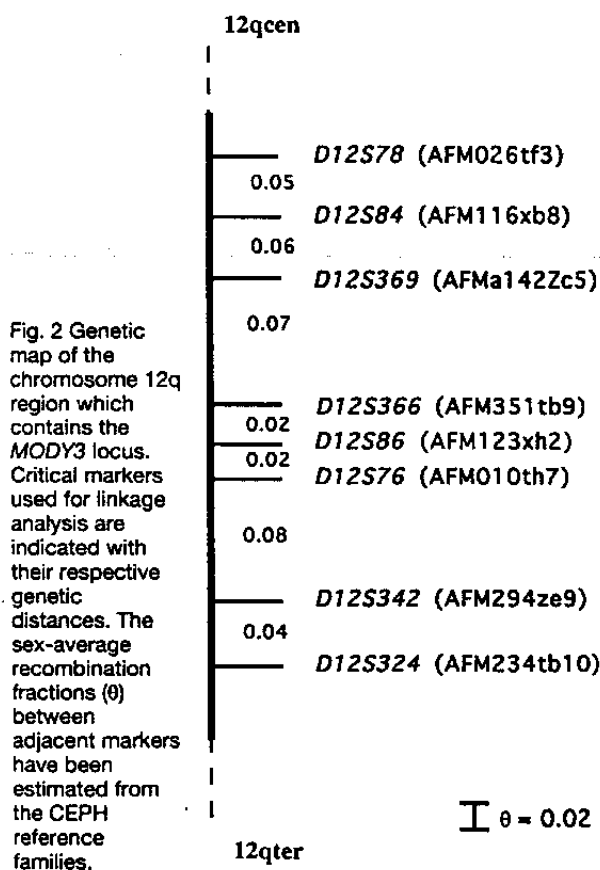


Table 2 Lod scores for eight markers and MODY in 12 families

Locus		Two-point analysis		Heterogeneity test		
		θ	Z_{max}	θ	α	Z_{max}
D12S78	(AFM026tf3)	0.38	0.18	0.00	0.15	0.48
D12S84	(AFM116xb8)	0.30	0.83	0.00	0.25	1.72
D12S369	(AFMa142zc5)	0.26	1.54	0.00	0.30	2.56
D12S366	(AFM351tb9)	0.28	0.88	0.00	0.35	1.05
D12S86	(AFM123xh2)	0.20	2.83	0.02	0.45	3.67
D12S76	(AFM010th7)	0.16	3.11	0.00	0.55	4.27
D12S342	(AFM294ze9)	0.18	2.31	0.10	0.65	2.53
D12S324	(AFM234tb10)	0.26	0.22	0.00	0.35	0.67

groups. It is noteworthy that despite the major hyperglycemia, fasting and 2h post-oral glucose load plasma insulin values are not significantly different in MODY3 subjects compared to the other groups. This suggests that an insulin secretory defect might be responsible for the chronic hyperglycaemia in this form of diabetes.

Regarding the age of diagnosis of chronic hyperglycaemia, no statistically significant difference was observed between the three groups of families (data not shown). However, as this parameter might suffer ascertainment bias, we have also compared the subgroup of subjects younger than 25 years, in whom hyperglycaemia was prospectively searched. A trend towards a later age of diagnosis in MODY3 families was observed (Table 3). These data agree with the trend obtained when the minimal ages of diagnosis in each MODY family were compared between the three groups (Table 3).

Our locus assignment represents the first step towards the identification of the defective gene. Based on the integration of the maps of NIH/CEPH and Weissenbach *et al.*^{3,12}, this novel MODY locus maps to 12q22-qter. The following genes, which may play a role in glucose metabolism and/or glucose induced insulin secretion, have already been mapped to this region: *ACADS* (Acyl-Coenzyme A dehydrogenase), *ALDH2* (Aldehyde dehydrogenase-2), *ATP2A2* (ATPase, Ca⁺⁺-dependent), *ATP2B1* (ATPase, Ca⁺⁺ transporting), *ATP5B* (ATP

synthase H⁺ transporting), *IGF1* (Insulin-like growth factor-1) and *PLA2* (Pancreatic phospholipase A2). As an insulin secretory defect is apparently a main characteristic of MODY patients carrying the chromosome 12q susceptibility haplotype, the MODY3 gene probably governs a crucial step of the pancreatic β -cell function. The identification of the gene defects causing MODY3 will not only provide clues to elucidate this subtype of diabetes, but also may contribute to our understanding of the molecular bases of the other forms of NIDDM, and to define new pharmacological targets.

Methodology

Clinical studies. Families were collected through a multimedia campaign to identify diabetes-prone families for genetic studies. These families were selected for further study because of the presence of NIDDM in at least three consecutive generations (except for family F538) and because at least two subjects were diagnosed with NIDDM before or at 25 years of age, a pattern of inheritance consistent with a diagnosis of MODY. Clinical data were obtained for each subject during a standardized clinical examination at the Hôpital Saint-Louis in Paris or by the subject's personal physician. Subjects were given a standard 75 g oral glucose tolerance test (OGTT) (n=79) and where glucose tolerance testing was not possible, fasting plasma glucose (FPG) samples were obtained (n=93). Individuals were considered affected if they were presently being treated for NIDDM, if the result of the OGTT showed them to have diabetes or impaired glucose tolerance (IGT) (using WHO criteria, that is, diabetes if FPG ≥ 7.8 mM and/or 2 h post glucose load plasma glucose ≥ 11 mM, IGT if 2 h post glucose load plasma glucose ≥ 7.8 mM), or if they had a FPG value ≥ 6.1 mM on two separate measurements. 172 subjects were studied, of

Fig. 3 Multipoint linkage analysis for the two conditions of genetic homogeneity and heterogeneity. Location scores for different positions of the MODY3 locus with respect to D12S86 (Marker 6), D12S76 (Marker 7) and D12S342 (Marker 8). The recombination rates for each interval are indicated above the solid line. - - - -, Homogeneity; ———, heterogeneity.

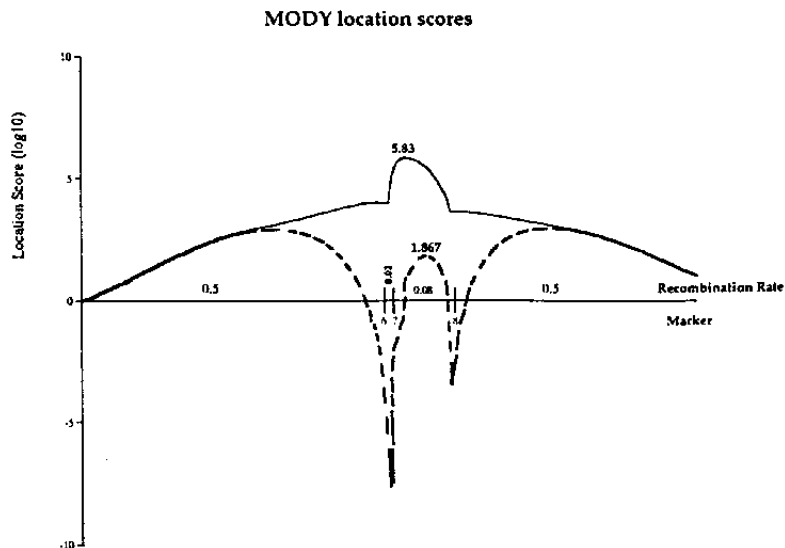


Table 3 Clinical characteristics of MODY patients based on the genetic defect

Characteristic	Chromosome 12q linked families (n=33)	Glucokinase-linked families (n=212)	Glucokinase and 12q unlinked families (n=40)	Statistical significance (P value)
Age (yrs)	40 ± 18 (11–81)	36 ± 21 (2–96)	36 ± 22 (4–81)	0.57
Body-mass index (Kg/m ²)	22.2 ± 2.3	22 ± 4	23.4 ± 4.2	0.15
Sex (M/F)	8/25 ^{a,b}	109/103	23/17	0.009
Glucose tolerance status MFH-IGT/DM (%)	1/32 ^{a,b} (3/97)	115/97 (54/46)	16/24 (40/60)	<0.00001
Age at diagnosis in youngest generation (yrs)	14 ± 2 (11–16)	10 ± 5 (1–21)	12 ± 7 (3–23)	0.06
Minimal age at diagnosis in the families (yrs)	14 ± 2 (11–17)	8 ± 6 (1–30)	10 ± 9 (3–23)	0.08
Fasting glucose (mmol l ⁻¹)	7.9 ± 3.1 ^{a,b}	7.1 ± 1	6.8 ± 1.5	0.006
2 h Glucose (mmol l ⁻¹)	14.8 ± 5.7 ^{a,b}	9.3 ± 2.9	9.8 ± 4.8	< 0.00001
Fasting Insulin (mU l ⁻¹)	11 ± 8	11 ± 7	10 ± 6	0.90
2 h insulin (mU l ⁻¹)	32 ± 25	32 ± 27	31 ± 21	0.99
Treatment				
Diet / OHA-insulin (%)	4/29 ^{a,b} (12/88)	135/77 (64/36)	18/22 (45/55)	<0.00001

Data are expressed as means ± SD (and range). Statistical significance of quantitative traits was assessed with ANOVA test¹⁵. When ANOVA was significant, comparisons between pairs were made using Tukey-Kramer HSD test¹⁶. *P* < 0.05 was considered significant. Statistical significance of qualitative traits was assessed with contingency-table χ^2 test. All statistically significant comparisons between two groups had a *P* < 0.005. ^aStatistically significant difference between chromosome 12 and glucokinase linked-families. ^bComparisons between chromosome 12 linked-families and families not linked to glucokinase or to chromosome 12 markers. MFH, Mild Fasting Hyperglycaemia; IGT, Impaired Glucose Tolerance; DM, Diabetes Mellitus; OHA, Oral Hypoglycaemic Agent.

whom 73 were considered affected for the purposes of the linkage analysis. 56 of the affected individuals had overt NIDDM, five had impaired glucose tolerance, and 12 had mild fasting hyperglycaemia (MFH). Of the diabetic subjects, 55% were being treated using oral hypoglycaemic agents or insulin. The majority of the patients (88%) were lean, with a body-mass index (BMI) below 27 kg/m² and had fasting insulin concentrations in the normal range despite hyperglycaemia. The average age of diagnosis of NIDDM was 23 ± 15 years (range, 3–79 yrs). NIDDM was diagnosed before 25 years of age in 50 subjects and before 10 years of age in nine subjects.

Genotyping studies. Genomic DNA was extracted from peripheral blood samples or from lymphoblastoid cell lines using standard procedures and simple sequence repeat polymorphic markers were genotyped using a PCR based method. All oligonucleotide sequences are available in the Genome Data Bank. PCR reactions were performed in a total volume of 50 µl containing 40 ng of genomic DNA, 50 pmol of each primer, 125 µM dNTPs, 1.5 mM MgCl₂ and 1 U of *Taq* polymerase with 5 µl of 10× specific buffer. Amplifications were carried out using a 'hot-start' procedure⁹. *Taq* polymerase was added after a denaturation step of 5 min at 96 °C. Samples were then processed through 35 cycles of denaturation (94 °C for 40 s) and annealing (55 °C for 30 s) followed by one last step of elongation (2 min at 72 °C). 16 amplification products, obtained with separate primer sets on identical DNA samples, were coprecipitated and comigrated in a single lane of 6% polyacrylamide denaturing gel containing 8 M urea. Separated products were then transferred onto Hybond N+ nylon membranes which were successively hybridized with a nonradioactive labelling procedure, using the ECL system (Amersham, UK)⁹.

Linkage analysis. Two-point and multipoint linkage analyses were performed with the LINKAGE programs¹³ and the program

HOMOG¹⁴ was used to obtain estimates and test statistics under heterogeneity. Marker allele frequencies were estimated from the genotypes observed for the founding pedigree members. For analyses involving two or more marker loci, genotypes were recoded to obtain four alleles at each locus with approximately equal frequencies. The gene frequencies for the recoded alleles were taken to be the sum of the original frequencies for the alleles that were grouped in recoding. Two-locus lod scores were not significantly modified by the recoding of alleles (results not shown). The genetic data were initially analysed using a previously described model⁷ with a frequency for the disease allele of 0.001 and equal female to male recombination rates was assumed. Because of the incomplete penetrance of MODY and its age-dependent expression, lod scores were calculated with four age-dependent liability classes: <10, 10–25, 25–40 and >40 years and with age-related prevalences *P* of 0.0003, 0.0006, 0.0012 and 0.0025, respectively. The frequency of phenocopies among the affected subjects was assumed to be 0.5%, 1%, 10% and 30%, for the four age-dependent liability classes, respectively, and was used to assign the penetrances for the homozygous non-susceptible genotypes (0.0000015, 0.000006, 0.00012 and 0.00075, respectively). The penetrances used for the homozygous and heterozygous susceptible genotypes in each liability class were 0.15, 0.30, 0.54 and 0.88, respectively. When significant evidence of linkage was found, the sensitivity of the results to these assumptions was examined by recalculating lod scores assuming a high penetrance (0.99) in all ages, and no phenocopy. Under the heterogeneity model, the disease was assumed to be due to mutations in a gene linked to the marker or markers in a proportion α of the families, and due to a gene or genes that were unlinked to the marker in a proportion 1- α of families. α was estimated for different recombination rates by a maximum likelihood technique. A lod score was calculated under heterogeneity as the logarithm (base 10) of the ratio of the maximum likelihood

(maximized over α and θ) to the value assuming no linkage. When several loci were analysed simultaneously, the calculations were undertaken assuming distances between the markers based on data from CEPH families. These distances did not differ substantially from those obtained by pairwise analysis between the markers in the MODY families. Re-estimation of the recombination rates between markers simultaneously with the placement of the disease locus and the heterogeneity parameter α did not change the conclusions.

Received 19 December 1994; accepted 3 February 1995.

Acknowledgements

We are grateful to the patients and their families for their participation in this study; Henri Lestrade and Jean-Jacques Robert; Séverine Clauin, Fabienne Dufour, Florence Lethrosne, Mathieu Leroux-Coyau, the Généthron team; Dominique Caterina; and Christian Vaisse for their help and contribution to this work. This work was supported by French Ministry for Research and Technology, Assistance Publique-Hôpitaux de Paris, and the Juvenile Diabetes Foundation International.

1. Fajans, S.S. Scope and heterogeneous nature of MODY. *Diabetes Care* 13, 49-64 (1990).
2. Froguel, P., Velho, G., Cohen, D. & Passa, P. Strategies for the collection of sibling-pair data for genetic studies in type 2 (non insulin-dependent) diabetes mellitus. *Diabetologia* 34, 685 (1991).
3. Bell, G.I. et al. Gene for non-insulin dependent diabetes mellitus (maturity-onset diabetes of the young subtype) is linked to DNA polymorphism on human chromosome 20q. *Proc. natn. Acad. Sci. U.S.A.* 88, 1484-1488 (1991).
4. Froguel, P. et al. The glucokinase locus on chromosome 7p is closely linked to early onset non insulin dependent diabetes mellitus. *Nature* 356, 162-164 (1992).
5. Gidh-Jain, M. et al. Glucokinase mutations associated with non-insulin-dependent (Type 2) diabetes mellitus have decreased enzymatic activity: implications for structure/function relationships. *Proc. natn. Acad. Sci. U.S.A.* 90, 1932-1936 (1993).
6. Froguel, P. et al. Familial hyperglycemia due to mutations in glucokinase: Definition of a subtype of diabetes mellitus. *New Engl. J. Med.* 328, 697-702 (1993).
7. Vaxillaire, M. et al. Search for a third susceptibility gene for maturity-onset diabetes of the young. Studies with eleven candidate genes. *Diabetes* 43, 389-395 (1994).
8. Weissenbach, J. et al. A second-generation linkage map of the human genome. *Nature* 359, 794-801 (1992).
9. Vignal, A. et al. A non-radioactive multiplex procedure for genotyping of microsatellite markers. In *Methods in Molecular Genetics: Gene and Chromosome Analysis* (ed. Adolph, K.W.) 211-221 (Academic Press, San Diego, 1993).
10. Herman, W. et al. Abnormal insulin secretion, not insulin resistance, is the genetic or primary defect of MODY in the RW pedigree. *Diabetes* 43, 40-46 (1994).
11. Velho, G. et al. Primary pancreatic b-cell secretory defect caused by mutations in the glucokinase in kindreds of maturity onset diabetes of the young. *Lancet* 340, 444-448 (1992).
12. NIH/CEPH Collaborative group. A comprehensive genetic linkage map of the human genome. *Science* 258, 67-86 (1992).
13. Lathrop, G.M. & Lalouel, J.M. Easy calculations of lod scores and genetic risks on small computers. *Am. J. hum. Genet.* 38, 460-465 (1984).
14. Ott, J. *Human Genetic Diseases: A Practical Approach* (ed. Davies, K.E.) 19-32 (IRL Press, Oxford, 1991).
15. Snedecor, G.W., Cochran, W.G. In *Statistical methods* (Iowa State University Press, Ames, Iowa, 1989).
16. Kramer, G.Y. Extension of multiple range tests to group means with unequal numbers of replications. *Biometrics* 12, 309-310 (1956).