Maturity-onset diabetes of the young (MODY) 5 is caused by mutations in the TCF2 gene encoding the transcription factor hepatocyte nuclear factor-1β. However, in 60% of the patients with a phenotype suggesting MODY5, no point mutation is detected in TCF2. We have hypothesized that large genomic rearrangements of TCF2 that are missed by conventional screening methods may account for this observation. In 40 unrelated patients presenting with MODY5 phenotype, TCF2 was screened for mutations by sequencing. Patients without mutations were then screened for TCF2 rearrangements by the quantitative multiplex PCR of short fluorescent fragments (QMPSF). Among the 40 patients, the overall detection rate was 70%: 18 had point mutations, 9 had whole-gene deletions, and 1 had a deletion of a single exon. Similar phenotypes were observed in patients with mutations and in subjects with large deletions. These results suggest that MODY5 is more prevalent than previously reported, with one-third of the cases resulting from large deletions of TCF2. Because QMPSF is more rapid and cost effective than sequencing, we propose that patients whose phenotype is consistent with MODY5 should be screened first with the QMPSF assay. In addition, other MODY genes should be screened for large genomic rearrangements. Diabetes 54:3126–3132, 2005

Maturity-onset diabetes of the young (MODY) is characterized by the occurrence of nonketotic diabetes of early onset, typically before the age of 25, caused by primary insulin-secretion defects and inherited as an autosomal dominant trait. Currently, heterozygous mutations in six different genes have been identified as a cause of MODY. These genes encode the enzyme glucokinase (MODY2 subtype) and the following transcription factors: hepatocyte nuclear factor-4α (HNF-4α; MODY1), HNF-1α (TCF1; MODY3), insulin promoter factor 1 (MODY4), HNF-1β (TCF2; MODY5), and neurogenic differentiation factor 1 (MODY6) (1).

In 20–40% of the patients presenting with clinical and family history consistent with MODY, no mutation in the known MODY genes are found (2,3). Part of these so-called MODY-X cases may be caused by mutations in still unidentified genes. Alternatively, some MODY-X cases could result from complex molecular alterations in the known MODY genes that are missed by conventional screening methods.

This hypothesis is supported by the observation that large genomic rearrangements account for up to 20% of the molecular defects responsible for other monogenic diseases (4–7). PCR amplification of individual exons followed by sequencing is currently the standard screening method for MODY mutation analysis. However, in the case of large genomic deletions involving one or several exons, this method would yield false-negative results due to the amplification of the single wild-type allele.

MODY5 encompasses a wide clinical spectrum comprising diabetes, pancreas atrophy with subclinical exocrine deficiency, progressive non-diabetic nephropathy, kidney and genital malformations, and liver test abnormalities (8). Sequence variations in the TCF2 gene cause MODY5, and all cases described so far have been associated with either heterozygous point mutations or deletion/insertion of a
found large deletions in the TCF2 gene in almost one-half of the adult patients with a MODY5 phenotype but no point mutations. All primer sequences and amplification conditions can be found in the standardized protocol. Creatinine clearance was calculated according to the Cockcroft formula. Imaging studies of the kidneys consisted of ultrasonography or computed tomography scan. Genital tract abnormalities (presence of cysts and/or reduced kidney size and/or pelvic abnormalities) and/or impaired renal function (creatinine clearance <0.5 ml/min per 1.73 m²) in the absence of persistent proteinuria (albumin excretion >0.5 g/24 h) and of diabetic retinopathy.

Clinical history of diabetes and of renal involvement was recorded using a standardized protocol. Creatinine clearance was calculated according to the Cockcroft formula. Imaging studies of the kidneys consisted of ultrasonography or computed tomography scan. Genital tract abnormalities were assessed by ultrasonography. All participants gave written informed consent.

### Molecular analysis

The minimal promotor, the coding regions, and exon-intron boundaries of TCF2 gene were screened for mutations by direct sequencing as described previously (9). Subjects who were negative for mutations were then screened for TCF2 rearrangements by QMPSF analysis.

QMPSF consists of the simultaneous amplification of multiple short exonic fragments. A limited number of cycles limits the PCR to the exponential phase of the amplification process and enables thus obtaining of a semiquantitative estimation of each PCR product. The analysis is based on the comparison of the peak heights generated from the tested DNA sample and control DNA (4). A heterozygous exon deletion will lead to a twofold reduction of the height of the corresponding peak. Primers were designed for the promoter and the exons 1 to 6 of TCF2 gene. A 5' extension, consisting of a rare combination of 10 nucleotides preceding the exon-specific sequence, was added to primers as described previously (6). The forward primer of each pair was 5' end-labeled with 6-FAM fluorochrome. Amplified DNA fragments were separated using the POP6 polymer on an ABI PRISM Genetic Analyzer 3100 sequencer (Applied Biosystems) according to the manufacturer’s instructions. Data were analyzed using the Genescan 3.7 software (Applied Biosystems). In all experiments, we included two control DNA, one with a complete TCF2 heterozygous deletion and the other with a known heterozygous TCF2 mutation (5,23).

To confirm complete TCF2 deletions, we used real-time quantitative PCR based on SYBR-Green I fluorescence. Real-time PCR was performed using an ABI 7900 Sequence Detection System (Applied Biosystems). Primers were designed for the nine exons of TCF2 gene, and the PCR was carried out in a 20-µl reaction using the SYBR-green I PCR master mix (Applied Biosystems) using 300 ng/ml of each primer and 20 ng DNA. We used as reference the β-globin gene (HBB). The number of copies was determined using the 2−ΔΔCT method where ΔΔCT = CTTCF2amplification − CTreference gene (24).

To define the boundaries of TCF2 deletions, we used a panel of seven markers: D17S1818, D17S1787, D17S934, and D17S1868 located upstream of TCF2 gene and D17S932, D17S937, and D17S939 located downstream of TCF2 gene. We also designed five sequence-tagged sites (STSs) in the neighboring genes TRIP2, AATF, MRPL55, and MEL15 and in the genomic contig AC113211. To confirm single-exon deletion, a second set of exonspecific primers was designed to exclude a false-positive result due to unreported polymorphisms affecting primer amplification. The deletion size was determined by long-range PCR using the Expand long Template PCR kit (Roche Diagnostics, Meylan, France) according to the manufacturer’s instructions. All primer sequences and amplification conditions can be found in the online appendix (available at http://diabetes.diabetesjournals.org).

### RESULTS

#### TCF2 alterations

We have screened 40 unrelated subjects presenting with a clinical phenotype consistent with MODY5.
MODY5 (Table 1). A molecular alteration of TCF2 was found in 70% \((n = 28)\) of them. The molecular defect was either a point mutation \((n = 18)\) or a gross genomic rearrangement \((n = 10)\). In all cases of genomic rearrangement but one, a complete heterozygous deletion of TCF2 was observed (Fig. 1A). All deletions were confirmed using the real-time quantitative PCR method for the nine exons of TCF2 (data not shown). Genotyping of neighboring microsatellite markers and STSs revealed that the deletion encompasses a genomic region of at least 1.2 megabases. The 5' breakpoint was located between TCF2 and the STS in the AC113211 contig, which is 131 kilobases away from TCF2. The 3' breakpoint was located between D17S1872 and the STS in exon 5 of TRIP3 gene (Fig. 2).

A single-exon deletion of exon 5 was observed in the remaining case of genomic rearrangement (Fig. 1B). We designed specific primers in the introns 4 and 5 and amplified by long-range PCR the rearranged fragment. Direct sequencing of the PCR product showed that the deletion extended from nucleotide \(-294\) in intron 4 to nucleotide +704 in intron 5, with loss of 1,159 nucleotides plus an insertion of 55 nucleotides at the deletion junction. A sequence homology search (www.ncbi.nih.gov/BLAST/) revealed that the inserted sequence was identical to the complementary sequence of a fragment located in intron 4, between IVS4+19314 and IVS4+19369, and inserted in inverse orientation (Fig. 3). Repetitive sequence elements and short repeats are known to be involved in gross genomic deletions through the formation of secondary structures between single-stranded DNA ends at breakpoint junc-

**FIG. 1.** Detection of TCF2 exonic deletions by QMPSF in patients with MODY5. QMPSF consists of the simultaneous amplification of multiple short exonic fragments labeled with a fluorochrome. The analysis is based on the comparison of the peak heights generated from the tested DNA and control DNA samples. The horizontal line indicates the size of PCR products expressed in bp; the vertical line indicates the intensity of fluorescence expressed in arbitrary units. Peak numbers correspond to TCF2 exons, and the “C” peak corresponds to the exon 3 of TCF1 used as a standard sample for adjusting peak heights. Fluorescence profiles of the patient (red) and those of a control subject (blue) are superimposed. A: A complete TCF2 deletion characterized by a twofold reduction of the height of all peaks. B: Deletion of exon 5 detected by a twofold reduction of the corresponding peak.

**FIG. 2.** Schematic representation of the TCF2-deleted region. Boundaries of the deletion were defined by genotyping microsatellite markers and STSs located within neighboring genes TRIP3, AATF, MRPL45, and MEL18 or in the genomic contig AC113211. The common minimal deletion encompasses a genomic region of 1.2 Mb delimited by TRIP3 and TCF2. Seven genes, AATF, ACACA, LHX1, TADA2L, DUSP14, DDXS2, and API1GBP1, and two predicted proteins, Q8N8D2 and Q96TB3, are located within the deleted region.
tions (25). To understand the mutational mechanism underlying the exon 5 deletion, we searched for repetitive sequence elements at the deletion breakpoints using the Repeat Masker program (http://www.repeatmasker.org). We found that both deletion breakpoints were located in repetitive sequences: the 5' deletion breakpoint in a human LINE-2 (long interspaced element-1, L2) element, and the 3' deletion breakpoint in a MIR (medium interspaced repeat) element. We also observed that this rearrangement involved two sequences, TATT and AAC, located at the start and the end of the inserted sequence, respectively, and homologous to the breakpoint junctions (Fig. 3).

Sixteen different point mutations were identified in 18 unrelated subjects (Table 2). These mutations included 11 missense mutations resulting in amino acid substitutions, 1 splice-site mutation, 3 nonsense mutations, and 1 frameshift mutation leading to a truncated protein. We have previously reported eight of these mutations (8). The other eight distinct mutations including six missense mutations are novel and affect residues conserved in human, pig,
LARGE GENE DELETIONS IN MODY5

TABLE 3
Main clinical characteristics of the patients with MODY5 due to mutations or deletions of TCF2

<table>
<thead>
<tr>
<th></th>
<th>Mutation</th>
<th>Deletion</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>18</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Age at latest follow-up (years)</td>
<td>45.7 ± 11.6 (34–73)</td>
<td>37.5 ± 15 (13–57)</td>
<td>NS</td>
</tr>
<tr>
<td>Age at first manifestation (years)</td>
<td>27.7 ± 13.9 (1–58)</td>
<td>17.8 ± 14.8 (1–52)</td>
<td>0.04</td>
</tr>
<tr>
<td>Diabetess</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at diagnosis (years)</td>
<td>29.5 ± 12.2 (10–58)</td>
<td>25.7 ± 15 (10–53)</td>
<td>NS</td>
</tr>
<tr>
<td>Symptoms at diagnosis (yes/no)</td>
<td>4/14</td>
<td>7/3</td>
<td>0.02</td>
</tr>
<tr>
<td>BMI at diagnosis (kg/m²)</td>
<td>24.9 ± 3.3 (17.8–32)</td>
<td>21.6 ± 4.0 (15–26.9)</td>
<td>NS</td>
</tr>
<tr>
<td>HbA1c at last examination (%)</td>
<td>7.2 ± 1.0 (5.6–9.2)</td>
<td>7.1 ± 1.6 (5.5–9.5)</td>
<td>NS</td>
</tr>
<tr>
<td>Insulin therapy (yes/no)</td>
<td>14/4</td>
<td>6/4</td>
<td>NS</td>
</tr>
<tr>
<td>Kidney disease</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at diagnosis (years)</td>
<td>32.4 ± 16.2 (1–63)</td>
<td>22.7 ± 17.2 (1–57)</td>
<td>NS</td>
</tr>
<tr>
<td>Kidney cysts (yes/no)</td>
<td>14/4</td>
<td>7/3</td>
<td>NS</td>
</tr>
<tr>
<td>Other kidney morphological abnormalities</td>
<td>8/10</td>
<td>6/4</td>
<td>NS</td>
</tr>
<tr>
<td>Creatinine clearance at latest examination (ml/min of 1.73 m²)</td>
<td>40 ± 20.1 (5.5–74)</td>
<td>61.0 ± 31.6 (15.8–108)</td>
<td>NS</td>
</tr>
<tr>
<td>Systolic blood pressure (mmHg)</td>
<td>127 ± 18 (104–170)</td>
<td>127 ± 15 (106–150)</td>
<td>NS</td>
</tr>
<tr>
<td>Diastolic blood pressure (mmHg)</td>
<td>76 ± 8 (60–90)</td>
<td>73 ± 6 (60–80)</td>
<td>NS</td>
</tr>
<tr>
<td>Genital abnormalities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genital abnormalities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genital abnormalities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kidney disease</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liver disease</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liver tests abnormalities (yes/no)</td>
<td>14/3</td>
<td>9/1</td>
<td>NS</td>
</tr>
</tbody>
</table>

Data are means ± SD (range).

mouse, Xenopus, and salmon TCF2 sequences. A panel of 212 control chromosomes and 250 chromosomes of patients with type 2 diabetes was screened to exclude that missense mutations may be rare polymorphisms.

Clinical characteristics of the patients with TCF2 abnormalities. A wide clinical spectrum was observed in the 28 MODY5 patients (Table 1). The first recorded manifestation of the disease in 70% of patients was diabetes, with or without associated renal disease. In 86% of cases, diabetes was diagnosed before the age of 40, and in 71% of cases, subjects were lean (BMI < 25 kg/m²) at diagnosis. At the latest follow-up, 71% of the patients required insulin therapy and all but two patients had some degree of renal impairment not consistent with diabetic nephropathy. The silent course of the renal disease may explain the wide range of the age at diagnosis (1–63 years).

A family history of diabetes in first-degree relatives was present in the majority of the patients. Various genital tract abnormalities were found in the 80% of the patients who were tested. Pancreas atrophy was observed in 60% of the patients. Lastly, 85% of the patients presented with fluctuating and increased plasma levels of liver enzymes.

A striking finding of this study was that similar clinical features were observed in carriers of point mutations and carriers of large genomic rearrangements (Table 3). Diabetes and renal disease were not more severe in the patients with TCF2 deletions than in subjects with a point mutation. Although the disease was diagnosed earlier and diabetes was more often associated with symptoms at onset in the former, a large overlap was observed between the two groups for all other clinical characteristics. By contrast, different clinical profiles were observed in patients with no detected abnormality of TCF2 despite a suggestive phenotype and in those with TCF2 alterations (Table 1). Subjects without TCF2 alterations were older at diagnosis of diabetes and renal disease, had a higher BMI at diagnosis of diabetes, less often required insulin therapy (an indirect index of insulin deficiency), and less often had liver test abnormalities. They also reported more often a family history of diabetes. Lastly, in all of these patients but one, renal morphological abnormalities were restricted to the presence of cysts.

DISCUSSION
We have shown in this report that large genomic rearrangements of TCF2 cause MODY5 and that whole-gene deletion is the most frequent molecular alteration observed in MODY5 patients. In this series, 70% of the patients presenting with a clinical phenotype consistent with MODY5, i.e., presenting with hyperglycemia suggesting a MODY phenotype, associated with either abnormalities of kidney morphology or impaired renal function, were carriers of TCF2 molecular alterations. Importantly, large deletions of TCF2 accounted for one-half the patients with a phenotype suggestive of MODY5 but who were so far classified as MODY-X because conventional screening had failed to detect a TCF2 mutation. These findings demonstrate that MODY5 is actually more frequent than previously estimated. Moreover, point mutations of TCF2 are sometimes associated with phenotypes restricted to a single organ (8,15). Thus, one cannot exclude TCF2 gene deletions being associated with more restricted phenotypes, particularly in young subjects with pure renal involvement or in patients with abnormal...
genital tract or pancreas development. In our series, no TCF2 molecular alteration was detected in 30% of the patients. At least two hypotheses may account for this observation. Molecular alterations of other gene(s) may lead to a phenotype similar to that of MODY5. Alternatively, in adult patients, type 2 diabetes and other abnormalities such as renal failure, renal cysts, and liver test abnormalities may be etiologically unrelated.

Although most of the TCF2 mutations reported so far were private, mutations affecting the donor splicing site of exon 2 were found in different populations, suggesting that this region is a hot spot for mutations (20). Our results suggest the presence of a second hot spot, because one-fourth of our patients were carriers of a large TCF2 deletion. The microsatellite and STS analysis revealed in all patients a minimal deleted region of 1.2 megabases. Given that gross genomic rearrangements are not randomly distributed in the human genome, we could hypothesize that the recurrence of this anomaly is due to a common mutational mechanism that could be either a homologous unequal recombination, mediated by similar nonallelic regions such as repetitive sequences, or a nonhomologous recombination occurring between short motifs of homology (26,27). The first mechanism is more frequently involved in gross deletions. By contrast, the presence of short sequences located at the breakpoints and on both sides of the inserted sequence may explain the rearrangement observed in the patient with the deletion of exon 5. Moreover, these sequences are located within AT-rich sequences known to promote secondary structural formation and frequently involved in the occurrence of deletion (25).

In our series, gross rearrangements of TCF2 were associated with no obvious specific phenotype. We could therefore hypothesize that TCF2 deletions lead to MODY5 through haplo-insufficient expression of HNF-1β protein, as shown for several TCF2 mutations (21). Interestingly in MODY1, it has been reported that a balanced translocation that did not disrupt the HNF-4α coding region could affect the gene expression by disrupting regulatory transcriptional elements (28).

The absence of associated clinical defects is remarkable, because the large deletions observed in nine patients span over 1.2 megabases. Similar observations have been reported in other diseases, in particular with cancer-related genes in which genomic deletions of several hundred kilobases have been associated with the classical disease phenotype (29,30). The large deletions we observed involve two sequences encoding putative proteins of unknown function and seven genes for which no link with the pathogenesis of MODY5 is obvious (Fig. 2.).

Our study shows that genomic rearrangements of TCF2 represent 30% of the molecular defects identified in MODY5 patients. These findings need to be replicated in other populations because they may have important implications for the choice of the diagnosis strategy of MODY5. As already mentioned, such alterations are missed by conventional amplification and sequencing. Given that clinical presentation provides no firm indication on the molecular mechanism and that QMPSF is more rapid and more cost effective than sequencing, we suggest that the routine molecular screening of TCF2 should be started with the QMPSF assay. Sequencing should be used in patients for whom the QMPSF assay is negative.

Finally, large genomic rearrangements might be implicated in other MODY subtypes. MODY3, caused by mutations in TCF1, is by far the most prevalent MODY subtype in adults. However, point mutations and small insertion/deletions do not account for a large number of patients with a phenotype consistent with MODY3 (31). We suggest that this could be due, at least in part, to TCF1 deletions. Actually, several studies have shown a linkage between early-onset autosomal-dominant type 2 diabetes and the region including the MODY3 locus. However, so far neither mutations in TCF1 nor identification of novel locus have been found (32,33). We suggest that patients with the so-called MODY-X should be systematically reassessed for the presence of deletions in the corresponding candidate genes.

ACKNOWLEDGMENTS

We thank Jessica Zucman-Rossi for the gift of TCF2 control DNA samples.

REFERENCES

14. Lindner TH, Njolstad PR, Horikawa Y, Bostad L, Bell GI, Sovik O: A novel syndrome of diabetes mellitus, renal dysfunction and genital malformation...


