Isomers of the TCF1 gene encoding hepatocyte nuclear factor-1 alpha show differential expression in the pancreas and define the relationship between mutation position and clinical phenotype in monogenic diabetes

Lorna W. Harries1,*, Sian Ellard1, Amanda Stride1, The European MODY consortium, Noel G. Morgan2 and Andrew T. Hattersley1

1Institute of Biomedical and Clinical Sciences, Peninsula Medical School, Barrack Road, Exeter, Devon, UK and
2Institute of Biomedical and Clinical Sciences, Peninsula Medical School, Research Way, Plymouth. Devon, UK

Received April 25, 2006; Revised May 23, 2006; Accepted June 1, 2006

The generation of multiple transcripts by mRNA processing has the potential to moderate differences in gene expression both between tissues and at different stages of development. Where gene function is compromised by mutation, the presence of multiple isoforms may influence the resulting phenotype. Heterozygous mutations in the transcription factor hepatocyte nuclear factor-1 alpha (HNF1A or TCF1 gene) result in early-onset diabetes as a result of pancreatic beta-cell dysfunction. We investigated the expression of the three alternatively processed isoforms of the HNF1A gene and their impact on the phenotype associated with mutations. Real-time PCR demonstrated variation in tissue expression of HNF1A isoforms: HNF1A(A), with the lowest transactivation activity compared with the truncated isoforms HNF1A(B) and HNF1A(C), is the major isomer in liver (54%) and kidney (67%) but not in adult pancreas (24%) and islets (26%). However, in fetal pancreas HNF1A(A) is the major transcript (84%), which supports developmental regulation of isomer expression. We examined whether the isomers affected by the mutation altered the diabetes phenotype in 564 subjects with 123 mutations in HNF1A. Mutations that affected only isomer HNF1A(A) (exons 8–10) were diagnosed later (25.5 years) than mutations affecting all three isomers (exons 1–6) (18.0 years) (P = 0.006). This first genotype/phenotype relationship described for patients with HNF1A mutations is explained by isomer structure and not by either mutation type or functional domain. We conclude that all three isomers may be critical for beta-cell function and could play a role in both the developing and mature beta cell.

INTRODUCTION

Messenger RNA processing pathways provide fine control over the temporal and spatial expression patterns of many genes. Approximately 74% of all human genes are thought to produce more than one transcript by variation in exon usage, or by the choice of variant promoters or polyadenylation sites (1). This is in agreement with estimates from EST-based database analysis that indicates that ~35–60% of human genes are alternatively spliced (2). Most genes code for two or three transcripts (3), but some, such as the human neurexin3 gene, produce many more potential isomers (4). Little is known about the factors that regulate the choice of splice or polyadenylation sites (3). Exon usage varies during development and differentiation (5) and also in response to environmental stimuli such as cellular stress (6,7), temperature (8) or activation of signalling pathways (9). Alternate isomers may have different DNA, protein or ligand binding properties.
variable subcellular localizations, different mRNA stabilities or exhibit variation in enzymatic or signalling activities (3). Differences in tumourigenic capacity of oncogenic cell lines brought about by alternate splicing have been reported (10,11), as has altered rates of neuronal degeneration associated with pathological TAOK2 isoforms in Alzheimer’s disease (12).

The hepatocyte nuclear factor-alpha (HNF1A) gene (approved gene name TCF1) encodes a crucial member of an autoregulatory transcription circuit in mature and developing pancreas (13,14). The HNF1A gene contains 10 exons and codes for a 631 amino acid homeoprotein with transcriptional activation properties. The HNF-1α protein consists of a dimerization domain (amino acids 1–33), a bipartite DNA-binding domain (homeo domain amino acids 100–184; POU domain amino acids 198–281) and a transactivation domain (amino acids 282–631) (15,16). The HNF1A gene has been reported to produce three alternate isomers [HNF1A(A) to HNF1A(C)] by differential splicing and polyadenylation (17). The HNF1A(B) and HNF1A(C) transcripts contain 6 and 7 exons, respectively (Fig. 1) and arise from alternative splicing and differential polyadenylation of transcripts (17). Transactivation studies suggest that all three isoforms have activity, but that the activity of isoforms HNF1A(B) and HNF1A(C) is approximately 5-fold greater than that of HNF1A(A). The expression profile of HNF1A(A) is well defined, but there are very few data on either the expression profile or the developmental impact of HNF1A(B) or HNF1A(C) transcripts.

In the developing mouse embryo, the Tcf1 gene is expressed relatively late in pancreatic development (around day E10.5), after the appearance of Ipf1, Isl1, Pax6 and Nkx2-2 (18,19). Expression is noted initially in most epithelial cells but is later confined to acinar cells and developing islets (20). These findings suggest that HNF-1α is not required for the initial stages of embryonic development but may play a role in maintaining the differentiation profile of developing islets after the secondary transition at day E14 (20). Accordingly, the islets of Tcf1 null mice and transgenic animals over-expressing mutant HNF-1α show abnormalities of morphology and size (21–23).

Tcf1 expression is noted in the pancreas, gut, liver and kidney of adult mice (21). In mature islets, HNF-1α regulates the expression of a large number of beta cell genes such as insulin, the glucose transporter molecule Glut2 and pyruvate kinase (20,24–27) and is important for the maintenance of mature pancreatic function. Homozygous knockout mice show reductions in glucose-stimulated insulin release and hyperglycaemia, which is attributed to defective glycolytic signalling proximal to mitochondrial oxidation (21,28,29) and also possibly defective mitochondrial function (30). No phenotype is observed in heterozygous Tcf1 knockout animals (21,28).

In humans, heterozygous mutations in the HNF1A gene are the most common cause of maturity-onset diabetes of the young (MODY) in most populations (31). This is a form of monogenic diabetes characterized by autosomal dominant inheritance, an early age of onset (often before 25 years) and a progressive failure of beta-cell function which is due to a failure to appropriately increase insulin secretion in the face of hyperglycaemia (32).

Over 193 mutations have been identified in the HNF1A gene (33) but to date, the phenotype was not correlated with genotype (either mutation type, mutation position or the location of the protein relative to the functional domain of the protein in a study of 34 families) (34). There is evidence that the age of onset is influenced by undefined genetic modifiers on chromosomes 14q24, 5p15 and 9q22 (35) and also maternal hyperglycaemia (36,37).

We have previously reported that abnormalities of HNF1A mRNA processing and/or mRNA stability can result in MODY (38,39). In this study, we aimed to determine whether the clinical phenotype of HNF1A mutations is influenced by their differential impact on alternate isoforms. First, we established the expression profile of each isomer in

Figure 1. Alternately processed transcripts of the HNF1A gene. The predicted structure of the HNF1A gene and of the three HNF1A isoforms is illustrated here. Isoforms HNF1A(B) and HNF1A(C) are generated by the use of alternate polyadenylation sites. The functional domains are identified on the genomic sequence in dark grey (dimerization domain), mid-grey (DNA binding domain) and light grey (transactivation domain). Processed exons in the mRNA transcripts are given by white boxes and incorporated intronic sequences by hatched boxes. The presence of mutations in exons 1–6, exon 7 and exons 8–10 are indicated by black stars, grey stars and white stars, respectively.
the fetal and adult pancreas to assess if specific isomers were more likely to play a role in beta-cell development or in the maintenance of mature beta-cell function. To do this we used real-time PCR assays specific to each HNF1A isoform and also to the endogenous housekeeping gene beta-2 micro-globulin (B2M) for comparative purposes. The impact of mutations in alternative isoforms was studied by examining the relationship between the location of mutations with respect to isomer structure and the age of diagnosis in a large cohort of patients carrying HNF1A mutations.

RESULTS

The HNF1A gene produces three alternatively processed isomers

HNF1A isomers were isolated from pooled pancreas, liver and kidney mRNA by long range reverse transcriptase–PCR. We were able to isolate all three transcripts corresponding to the previously described HNF1A isomer sequences (17) and did not find any novel isomers. There was no evidence for the existence of additional transcripts (Fig. 2).

Validation of real-time PCR for the detection and quantification of alternatively spliced transcripts

Real-time PCR assays specific to HNF1A isoforms identified their targets selectively across a dynamic range of more than seven 1:2 serial dilutions. The efficiencies of detection were \( r^2 = 0.96 \) and \( r^2 = 0.98 \) for HNF1A(A); \( r^2 = 0.96 \) and HNF1A(C); \( r^2 = 0.95 \).

HNF1A isoforms have different expression profiles in different tissues

Expression of all three HNF1A isoforms was noted in kidney, liver, pancreas and isolated islets. As expected, total HNF1A expression was extremely low in skeletal muscle and in visceral adipose tissues (Fig. 3).

Although the total level of HNF1A expression was comparable between liver, islet and pancreas, the distribution patterns of isomers were different. In both kidney and liver, the predomi-nant transcript was HNF1A(A); comprising 67 and 54% of total HNF1A expression, respectively. Isomers HNF1A(B) and HNF1A(C) were less represented, with levels of only 25 and 8% (kidney) and 35 and 11% (liver). Conversely, in total pancreas RNA and isolated islet cell RNA, HNF1A(A) expression was only 24 and 26% of total HNF1A expression, respectively. The predominant transcript was HNF1A(B), with levels of 55% (pancreas) and 49% (pooled islets). Levels of HNF1A(C) were also correspondingly higher [21% (pancreas) and 25% (pooled islets)] (Fig. 3). As most tissue samples were pooled at source, we were unable to determine the degree of variation between individuals.

HNF1A isoforms show different expression patterns during pancreatic development

The expression levels of the different isomers differed between adult and fetal total pancreas mRNA (Fig. 3). In mature pancreas, the most abundant isomer was HNF1A(B) (55%) when compared with HNF1A(A) (24%) or HNF1A(C) (21%). In fetal pancreas, the dominant isomer was HNF1A(A) (84%), with lower expression of HNF1A(B) (13%) and HNF1A(C) (3%).

Mutation position influences age of diagnosis of MODY

We went on to assess if the position of mutations in the HNF1A gene relative to isomer structure altered the age of diabetes diagnosis in patients. Mutations were grouped as; exons 1–6 [which altered isomers HNF1A(A), HNF1A(B) and HNF1A(C)], exon 7 [isomers HNF1A(A) and HNF1A(B)] and exons 8–10 [isomer HNF1A(A) only].
The ages of diagnosis expressed either as family median, or as the age of diagnosis of the youngest member are given in Table 1. The family median age at diagnosis for all mutations in the first six exons of the gene was 18 years. Patients with mutations in exon 7 were diagnosed with diabetes at 19 years, whereas those in exons 8–10 produced a later age of diagnosis of 25.5 years ($P = 0.006$). This effect was consistent regardless of whether the total dataset, the family medians or the age of diagnosis of the youngest family member was used for analysis (Table 1). Pairwise testing indicated that compared with mutations in exons 1–6, the impact on age of diagnosis was limited to mutations in exons 8, 9 or 10 ($P = 0.002$); as mutations in exon 7 did not alter the age of diagnosis ($P = 0.50$).

The effect of mutation position on age of diagnosis was restricted to missense alterations as mutations in the first six exons were associated with a median age of diagnosis of 18 years, whereas mutations in exons 8–10 caused diabetes to appear 12 years later (30 years; $P = <0.0001$) (Table 1); no significant correlation was noted between mutation position and age at diagnosis for nonsense or frameshift mutations generating premature termination codons (exons 1–6, age of diagnosis 18 years; exon 7, age of diagnosis 20 years; and exons 8, 9 or 10, age of diagnosis 19 years; $P = 0.87$).

**Age of diagnosis effect is not related to the functional domains of the protein or the mutation type**

An effect of mutation position on age at diagnosis could also be mediated by disruption of different functional domains. We thus correlated age of diagnosis and the position of the mutation with respect to the dimerization, DNA binding and transactivation domains of the $HNF1A$ gene. Considering the dataset as a whole, no correlations were noted between mutations located in the dimerization domain (age at diagnosis 18.5 years), DNA binding domains (age at diagnosis 18 years) or the transactivation domain (age at diagnosis 18 years) ($P = 0.35$). However, when missense mutations alone were considered, a statistically significant association was noted between the transactivation domain and age at diagnosis ($P = 0.001$). However, as all mutations in exons 8, 9 or 10 are located in this region, the association with this domain could be indirect. To determine whether functional domain alone could account for the variation in age at diagnosis, we carried out a linear regression analysis including both isomer structure and the functional domains of the protein. The association with isomer structure remained significant ($P = <0.0001$ for all mutations, $P = 0.003$ for missense mutations only), whereas that with the transactivation domain did not ($P = 0.43$ for all mutations, $P = 0.14$ for missense mutations only). In keeping with this, the variation in age at diagnosis was still concordant with isomer structure when we restricted our analysis to mutations within the transactivation domain; individuals with mutations in exons 4, 5 and 6 had a mean age of 17 years, whereas those with mutations in exons 7, 8, 9 or 10 had a mean age of 24.8 years ($P = 0.0009$).

Similarly, age of diagnosis effects could possibly be mediated by variation in mutational mechanism between different mutation types. However, when missense mutations were compared with nonsense or frameshift mutations, no effect of mutation type on age of diagnosis was seen (20 and 18 years for missense and prematurely terminating mutations, respectively; $P = 0.1$). Further confirmation of the role of alternate mRNA processing in the determination of $HNF1A$ phenotype could be achieved by analysis of individuals with mutations in the intronic regions specific to isomers $HNF1A(B)$ and $HNF1A(C)$. We examined the partial sequence of IVS6 and IVS7 in a cohort of 272 patients in whom sequencing of the exons and conserved splice sites of the $HNF1A$ gene did not identify a mutation. We only identified a single synonymous

**Table 1. Age of diagnosis according to mutation position**

<table>
<thead>
<tr>
<th></th>
<th>All data</th>
<th></th>
<th>Family median</th>
<th></th>
<th>Youngest member</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Median age of diagnosis (years)</td>
<td>N</td>
<td>Median age of diagnosis (years)</td>
<td>N</td>
<td>Median age of diagnosis (years)</td>
</tr>
<tr>
<td>Exons 1–6</td>
<td>480</td>
<td>18 (12–24)</td>
<td>215</td>
<td>18 (13–23)</td>
<td>215</td>
<td>15 (11–19)</td>
</tr>
<tr>
<td>Exon 7</td>
<td>39</td>
<td>20 (13–27)</td>
<td>20</td>
<td>19 (13.6–24.4)</td>
<td>20</td>
<td>16.5 (11–22)</td>
</tr>
<tr>
<td>Exons 8–10</td>
<td>45</td>
<td>27 (16.7–37.3)</td>
<td>21</td>
<td>25.5 (16.5–34.5)</td>
<td>21</td>
<td>20 (12.2–27.8)</td>
</tr>
<tr>
<td>$P$-value</td>
<td>&lt;0.0001</td>
<td>0.006</td>
<td></td>
<td></td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Exons 1–6</td>
<td>220</td>
<td>18 (12.5–23.5)</td>
<td>95</td>
<td>18 (13–23)</td>
<td>95</td>
<td>15 (11–19)</td>
</tr>
<tr>
<td>Exon 7</td>
<td>23</td>
<td>18 (7.5–28.5)</td>
<td>9</td>
<td>18 (7.7–28.3)</td>
<td>9</td>
<td>17 (13.9–20.1)</td>
</tr>
<tr>
<td>Exons 8–10</td>
<td>31</td>
<td>29 (19.5–38.5)</td>
<td>14</td>
<td>30 (20.9–39.1)</td>
<td>14</td>
<td>23 (16.5–29.5)</td>
</tr>
<tr>
<td>$P$-value</td>
<td>&lt;0.0001</td>
<td>0.002</td>
<td></td>
<td></td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>Exons 1–6</td>
<td>260</td>
<td>18 (11.6–24.4)</td>
<td>120</td>
<td>18 (13–23)</td>
<td>120</td>
<td>15 (11.5–18.5)</td>
</tr>
<tr>
<td>Exon 7</td>
<td>16</td>
<td>21 (15.5–26.5)</td>
<td>11</td>
<td>20 (14.7–25.3)</td>
<td>11</td>
<td>16 (10–22)</td>
</tr>
<tr>
<td>Exons 8–10</td>
<td>14</td>
<td>17 (4.1–29.9)</td>
<td>7</td>
<td>19 (12.7–25.3)</td>
<td>7</td>
<td>13 (8.5–17.5)</td>
</tr>
<tr>
<td>$P$-value</td>
<td>0.95</td>
<td>0.87</td>
<td></td>
<td></td>
<td>0.64</td>
<td></td>
</tr>
</tbody>
</table>

$N$, number of samples; $P$-values were obtained by Kruskal–Wallis testing. ‘All data’ column refers to the complete data set. ‘Family median’ refers to the median age of diagnosis of each family. ‘Youngest member’ refers to the age of diagnosis of the youngest affected individual in each family. PTC, mutations producing a premature termination codon. Data are medians +/- IQR.
variant at position IVS7nt + 65G>C which is unlikely to have functional significance. The lack of mutations rendered any further analysis unfeasible.

**DISCUSSION**

We have confirmed the expression of three alternate isoforms of the HNF1A gene, as reported in a previous study (17). HNF1A mRNA transcripts are differentially expressed in tissues in man that are relevant to diabetes. In the most important tissue, the pancreas, the relative expression of isomers differs between the fetus and the adult. Their importance in both the developing fetus and also in the maintenance of mature beta-cell function is strongly supported by a marked difference in the age of diagnosis between mutations which affect all three isomers and those which affect only one isoform. Understanding of HNF1A isomer expression therefore suggests a genotype:phenotype relationship which has not been seen previously in this commonest subtype of monogenic diabetes.

Our analysis indicates that all three HNF1A (TCF1) isoforms are present in human liver, kidney, pancreas, isolated islet and gut, but not in skeletal muscle or in visceral adipose tissues, at least at the mRNA level. Absolute confirmation of expression awaits protein analysis; this is currently difficult because of the lack of suitable antibodies and the high levels of cross reactivity with the HNF1beta gene products. Our data is in agreement with Tcf1 expression data from the mouse (21,29). There are differences in the relative spatial expression patterns of human isoforms (Fig. 3). These differences in tissue-specificity may reflect differences in the function of the isoforms. The HNF1A(A) isomer, predominant in liver and kidney, is thought to have a lower transactivation potential than the HNF1A(B) or HNF1A(C) isoforms on the basis of in vitro studies (17). Variation in the expression profiles may also reflect the need to activate different subsets of downstream effectors in these tissues.

The expression profiles of HNF1A isoforms in the pooled islet samples are similar to those obtained from total pancreas, both in the distribution of isoforms and in the magnitude of expression. Messenger RNA deriving from entire pancreases contains ductal and connective tissue in addition to the endocrine and exocrine derivatives. The similarity of expression profiles may indicate therefore that HNF1A isoforms are expressed at similar levels in both endocrine and exocrine tissues. However, we acknowledge that the present data cannot exclude the possibility that the expression profile may vary among the different endocrine cell types within the islet, or that it may be modified by factors at the translational level.

The HNF1A isoforms also demonstrate temporal or developmental variation in expression, with an apparent developmental switch from HNF1A(A) in the fetal pancreas to HNF1A(B) in the adult (Fig. 3). Prior to this study, information on expression has only been documented for the well recognized HNF1A(A) isomer. This is broadly expressed in developing pancreas as early as day E10.5, with hormone-positive endocrine cells and amylase-positive cells expressing HNF-1α around day E15.5 (18). This may be consistent with a role in fetal development. HNF1A gene knockout in transgenic mice does not profoundly affect embryogenesis (22), which may suggest that HNF-1α is not required for early endoderm commitment, but may play a role in maintaining the differentiation status of the developing islets (20).

The HNF-1α protein is a member of an autoregulatory circuit in pancreatic cells (13,14). In a model proposed by Ferrer in 2002, the circuit can exist either in an ‘on’ or an ‘off’ state (14). The transition from ‘on’ to ‘off’ is regulated by changes in the expression levels of any of the key regulatory genes in the network. A reduction in any of these genes could cause the network to be set to the ‘off’ position and set into play a chain of events that ultimately results in beta-cell dysfunction. HNF-1α is a key regulator of HNF-4α in pancreatic cells (13). HNF4A is expressed very early in embryogenesis and is doubtlessly important in early embryonic development, as knockout animals are not viable (40). The spatial and temporal variation in expression patterns of HNF1A isoforms we report in this study give rise to the possibility that variant HNF1A isoforms may have different regulatory roles in maintaining the transcriptional status of the network in the active state. Mutations that differentially affect HNF1A isoforms may therefore exert different influences on the development of beta-cell dysfunction and thus on phenotype.

We hypothesized that mutations in the proximal segment of the HNF1A gene should give rise to a more severe phenotype than mutations in the distal regions, because some isoforms may evade the effects of mutation in the latter. Mutations in exons 1–6 are expected to functionally impair all three isoforms, whereas those in exon 7 will affect isoforms HNF1A(A) and HNF1A(B) selectively. Mutations in exons 8, 9 or 10 will affect only isomer HNF1A(A) (as indicated in Fig. 1). Previous studies have failed to show variation in age of diagnosis with position of mutation (34). We found that missense mutations in exons 1–6 (affecting all three isoforms) and exon 7 [affecting isoforms HNF1A(AA) and HNF1A(CC)] showed no apparent differences in age of diagnosis (18 vs. 18 years; P = 0.13). However, where missense mutations were present in exons 8–10 [affecting only variant HNF1A(A)], the age of diagnosis was 12 years later (18 versus 30 years; P = 0.0003). These findings support the concept that the expression of isomer HNF1A(B) plays an important role in the continued maintenance of beta-cell function. This is consistent with our finding that HNF1A(B) is the most abundant isomer in mature islets. These findings were independent of the position of the mutation with reference to the functional domains of the protein, and were also unrelated to mutation type.

In our cohort, the genotype:phenotype relationship appeared to be restricted to missense mutations. Families carrying mutations predicted to lead to the generation of premature termination codons (nonsense, frameshifted or splice-site mutations) showed no difference in age of diagnosis, regardless of the position of the mutation (Table 1). This may seem difficult to reconcile with our hypothesis, as mutations leading to loss of function should show equivalent effects, regardless of mutational mechanism. However, a likely explanation is that nonsense and frameshift mutations are degraded by surveillance pathways such as nonsense-mediated decay.
(NMD) (41). We have already shown that this is the case for the commonest mutation, a C insertion at amino acid 291 (39), and that it is common for HNF1beta mutations (42). This could result in the mutated isomer not being expressed. The only exception to this would be mutations in exon 10, that are immune from NMD because of the requirement of a downstream exon to mark the PTC as premature (41). There were only two families with a truncating mutation in exon 10, that are immune from NMD because of the isomers that are mutated is the first description of a genotype: mutation in exon 10 and they had a median age at diagnosis of 17 years.

The alteration in the age of diagnosis according to the isomers that are mutated is the first description of a genotype: phenotype relationship for mutations in the HNF1A gene. This is consistent with the isomer structure that results from messenger RNA processing pathways defining the impact of mutations on the clinical phenotype. The observed differences in spatial and temporal expression of isoforms in the pancreas could reflect differences in their function. Our results are consistent with a critical role for all three HNF1A isoforms in maintaining beta-cell function, which may be mediated by differences in their function or in the precise location and timing of their expression.

**METHODS**

**Expression profiling**

RNA samples used for study. RNA samples were obtained from kidney (pooled from six Caucasian females, 28–52 years), skeletal muscle (pooled from seven male and female Caucasians, 20–68 years) (Clontech, Oxford, UK), liver (51-year male Caucasian), total pancreas (pooled from five male and female Caucasians, 24–77 years), fetal pancreas (pooled 19/27-week gestation), adipose tissue (male, 25 year, Caucasian) (AMS Bioscience, Abingdon, UK) and two separate DHAL3 islet preparations [National Disease Resource Interchange (NDRI), Philadelphia, USA]. All RNA samples were DNase-treated prior to reverse transcription using the TURBO DNase kit (Ambion, Huntingdon, UK).

**Isolation and identification of HNF1A isoforms**

A total of 5 μg of DNAse-treated RNA was reverse-transcribed from pooled pancreas, liver and kidney samples using the Thermoscript™ RT–PCR System (Invitrogen, Paisley, UK) in a total volume of 20 μl. cDNA samples were treated with 1 U RNase H at 37°C for 30 min (Invitrogen) to remove RNA prior to amplification. 2 μl DNA was then amplified using primers designed to amplify the entire sequence of HNF1A(A), HNF1A(B) or HNF1A(C) isoforms. Amplifications were as described in Frayling et al. (31). Primer sequences are given in Table 2. PCR products were then directly sequenced on an Applied Biosystems 3100 capillary sequencer using Big Dye V1.1 technology (Applied Biosystems, Foster City, USA).

**Assay design**

Reference sequences corresponding to HNF1A variants HNF1A(A), HNF1A(B) and HNF1A(C) were identified from the Entrez Nucleotides Database (http://www.ncbi.nlm.nih.gov). The ubiquitously expressed β2M gene was selected as a control on the basis of its constant expression levels across the tissues and developmental stages studied. An assay for this gene (Assay Hs00187842) was purchased by Assays-on-Demand from Applied Biosystems. Custom real-time PCR assays to HNF1A isoforms were obtained from the Assays-by-Design service available from Applied Biosystems. Probe and primer sequences are given in Table 2. Where possible, probes were chosen to span introns to ensure amplification of cDNA rather than genomic DNA. Assays were validated by standard curve analysis of serial 1:2 dilutions of pooled liver, kidney and pancreas cDNA. The efficiency of each assay was assessed by reference to the slope of the corresponding standard curve. In a maximally efficient reaction, there should be 3.3 cycles between each serial dilution. The gradient of such a curve will thus be −3.3. Therefore, the closer the gradient of the experimental standard curve to −3.3, the more efficient the reaction.

### Table 2. Probe and primer sequences for expression profiling

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>RT-HNF1A(ABC)F</em></td>
<td>5'-ATG GTT TCT AAA CTG AGC CAG-3'</td>
<td>None</td>
</tr>
<tr>
<td><em>RT-HNF1A(A)R</em></td>
<td>5'-TAC TGG GAG GAA GAG GC-3'</td>
<td>None</td>
</tr>
<tr>
<td><em>RT-HNF1A(B)R</em></td>
<td>5'-CAA TCA CGG AAA AAC ACT-3'</td>
<td>None</td>
</tr>
<tr>
<td><em>RT-HNF1A(C)R</em></td>
<td>5'-CAC CCA GGT ATC ACT GAT-3'</td>
<td>None</td>
</tr>
<tr>
<td>HNF1A(A)F</td>
<td>5'-CCT GGC CAG CCT CAC G-3'</td>
<td>None</td>
</tr>
<tr>
<td>HNF1A(A)R</td>
<td>5'-GTG TGA AGC CCG GAC TCA-3'</td>
<td>None</td>
</tr>
<tr>
<td>HNF1A(A)P</td>
<td>5'-CAC CAA GCA GGT CTT C-3'</td>
<td>5' 6-FAM, 3' MGB</td>
</tr>
<tr>
<td>HNF1A(B)F</td>
<td>5'-GAC CCA GAG CCC CTT CAT G-3'</td>
<td>None</td>
</tr>
<tr>
<td>HNF1A(B)R</td>
<td>5'-CCT CTA TCA TCA TCT CTT CAC GCT GTG T-3'</td>
<td>None</td>
</tr>
<tr>
<td>HNF1A(B)P</td>
<td>5'-CCC ACG GTG AGC ACC-3'</td>
<td>5' 6-FAM, 3' MGB</td>
</tr>
<tr>
<td>HNF1A(C)F</td>
<td>5'-GTT CCT AGC TTC ACC ACA-3'</td>
<td>None</td>
</tr>
<tr>
<td>HNF1A(C)R</td>
<td>5'-GTT GCC CAC CCA TCC C-3'</td>
<td>None</td>
</tr>
<tr>
<td>HNF1A(C)P</td>
<td>5'-CTG GTC ATC GGT AAG CTG-3'</td>
<td>5' 6-FAM, 3' MGB</td>
</tr>
</tbody>
</table>

Primers for isolation and sequencing of transcripts (marked with stars) and real-time probes and primers specific for HNF1A isoforms HNF1A(A), HNF1A(B) and HNF1A(C) are given below. Probes are labelled 5' with 6-fluorescein (6-FAM) and 3' with a minor groove-binding protein (MGB) to enhance probe specificity.
The correlation between cDNA concentration and crossing point ($r^2$ value) was assessed from the fit of the crossing point corresponding to each dilution point to the curve.

**Reverse transcription and real-time PCR amplification**

Real-time PCR reactions to determine $\beta 2M$, HNF1A(A), HNF1A(B) and HNF1A(C) expression levels were carried out using the ABI Prism 7000 platform. Each sample was amplified in triplicate to ensure accuracy of quantification. PCR reactions contained 10 $\mu$l TaqMan Universal Mastermix (no AMPerase) (Applied Biosystems), 0.9 $\mu$M each primer, 0.25 $\mu$M probe and 2 $\mu$l cDNA reverse-transcribed as above in a total volume of 20 $\mu$l. PCR conditions were a single cycle of 95°C for 10 min followed by 40 cycles of 95°C for 15 s and 60°C for 1 min.

**Expression profiling**

Expression levels of $\beta 2M$, HNF1A(A), HNF1A(B) and HNF1A(C) were calculated by reference to the average crossing point of the triplicate samples. For each RNA sample, the difference between the average crossing point (Ct) obtained from $\beta 2M$ amplification (Ct$^{\beta 2M}$) and the average crossing point obtained from the test transcript (Ct$^{test}$). This figure was termed the $\Delta$Ct value. $\Delta$Ct values were calculated for each isoform. The level of each transcript relative to $\beta 2M$ levels could then be calculated from the following equation: $2^{-\Delta \text{Ct}}$ (43) where $\Delta C t$ is the $\Delta$Ct value of the test transcript ($\Delta C t^{test}$) normalized to a reference transcript which was taken to be the levels of HNF1A(A) in human liver ($\Delta C t^{ref}$).

**HNF1A subjects and mutations**

Family data were gathered from three sources: families that had been referred for testing in our laboratory in the period between 1996 and 2005 (n = 328); data collected by the European MODY consortium (n = 66); and families reported in the literature for which age of diagnosis data were available (n = 170) (33). We studied the possible effect of alternate mRNA processing on phenotype in a total of 123 different mutations found in 564 subjects from 256 families. Of these, 85% mutations were located in the first six exons, 7% grouped in exon 7 and 8% resided in exons 8-10. Forty-nine percent mutations were missense changes, 48% were frame-shift or nonsense and 3% were splice-site mutations. Mutations amounting to 1.6% were located in the promoter region, 1.4% mutations were located in the dimerization domain, 2.3% were in regions with undefined function, 44.9% in the DNA binding domain and 49.8% in the transactivation domain.

**Correlation of phenotype with mutation position**

Age at diagnosis was taken as a measure of phenotype because it is less susceptible to subjective variation than features such as the treatment regime. To avoid bias due to possible earlier diagnosis for individuals in families with a previously reported HNF1A mutation, we carried out our analysis using both the family median age of diagnosis and the age at diagnosis of the youngest family member.

We correlated age at diagnosis with mutation position for missense mutations and prematurely terminating mutations both together and separately. In order to establish definitively that any effect noted was because of the effect of mutation on different HNF1A isomers and was not related to differences in mutational mechanism, we also correlated age at diagnosis with mutation type. An effect of mutation position on age at diagnosis could also arise from differences in the impact of inactivation of dimerization, DNA binding or transactivational activity. In order to examine this we also tested if the age at diagnosis was correlated with the location of the mutation with respect to the functional domains of the protein.

**Statistical methods**

Differences between the median age of diabetes diagnosis between the described subsets of HNF1A mutation carriers were examined for statistical significance using Kruskal–Wallis tests. Pairwise comparisons were achieved using Mann–Whitney U analysis. Correction for the effect of functional domain on age of diagnosis was produced by linear regression. All statistics were carried out using the SPSS software package (SPSS plc., Chicago, USA).

**Members of the European GIFT MODY consortium (44)**

M. Vaxillaire (Department of Human Genetics, Pasteur Institute, Paris), T. Tuomi (Helsinki University Central Hospital, Department of Medicine, Finland), E. Barbetti (IBCIT Biomedical Science Park, Rome), P.R. Njolstad, (Department of Paediatrics, University of Bergen, Norway), T. Hansen (Steno Diabetes Center, Denmark), A. Costa (Endocrinology and Diabetes Unit, University of Barcelona), I. Congret (Endocrinology and Diabetes Unit, University of Genoa), O. Pedersens (Steno Diabetes Center, Denmark), O. Sovik (Department of Paediatrics, University of Bergen, Norway), R. Lorini (Department of Paediatrics, University of Genoa), P. Froguel (Department of Human Genetics, Pasteur Institute, Paris).

**ACKNOWLEDGEMENTS**

We are grateful to Dr Neil Hanley for helpful discussion. We thank Kate Thomson and Kevin Colclough of the Molecular Genetics Laboratory at the Royal Devon and Exeter Hospital for identifying the mutations in the Exeter cohort. We are grateful to Dr Beverley Shields for statistical advice. This study was funded by a Diabetes UK small grant. We also thank the Wellcome Trust for additional financial support. A.T.H. is a Wellcome Trust research leave fellow and S.E. is funded by the Royal Devon and Exeter NHS Foundation Trust R&D Directorate.

**Conflict of interest statement.** The authors report no financial interests or connections, direct or indirect, or other situations that might raise the question of bias in the work reported or in the conclusions, implications or opinions drawn.
REFERENCES


