Title

Heterozygous *RFX6* protein truncating variants cause Maturity-Onset Diabetes of the Young (MODY) with reduced penetrance

Authors:

Kashyap A Patel[1], Markku Laakso[2,3], Alena Stančáková[3], Thomas W Laver[1], Kevin Colclough[4], Matthew B Johnson[1], Jarno Kettunen[5] ,Tiinamaija Tuomi[5,6], Miriam Cnop[7,8], Maggie H. Shepherd[1], Sarah E Flanagan[1], Sian Ellard[1], Andrew T Hattersley[1]*, Michael N Weedon[1]*#

*Joint contribution
corresponding author - M.N.Weedon@exeter.ac.uk

1) Institute of Biomedical and Clinical Science, University of Exeter Medical School, Exeter, UK.
2) Department of Medicine, Kuopio University Hospital, Kuopio, Finland.
3) Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland, Kuopio, Finland.
4) Department of Molecular Genetics, Royal Devon and Exeter National Health Service Foundation Trust, Exeter, UK
5) Department of Endocrinology, Abdominal Centre, Helsinki University Hospital; Folkhalsan Research Center; and Research Programs Unit, Diabetes and Obesity, University of Helsinki, Helsinki, Finland
6) Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland
7) ULB Center for Diabetes Research, Faculty of Medicine, Université Libre de Bruxelles, Brussels, Belgium
8) Division of Endocrinology, Erasmus Hospital, Université Libre de Bruxelles, Brussels, Belgium

Abstract

Finding new genetic causes of monogenic diabetes can help to understand development and function of the human pancreas. We aimed to find novel protein-truncating variants causing Maturity-Onset Diabetes of the Young (MODY), a subtype of monogenic diabetes. We used a combination of next-generation sequencing of MODY cases with unknown aetiology along with comparisons to the ExAC database to identify new MODY genes. In the discovery cohort of 36 European patients, we identified two probands with novel *RFX6* heterozygous nonsense variants. *RFX6* protein truncating variants were enriched in the MODY discovery cohort compared to the European control population within ExAC (Odds Ratio=130, $P=6\times10^{-28}$). We found similar results in non-Finnish European (n=348, OR=26, $P=6\times10^{-14}$) and Finnish (n=80, OR=22, $P=5\times10^{-18}$) replication cohorts. *RFX6* heterozygotes had reduced penetrance of diabetes compared to common *HNF1A* and *HNF4A* MODY mutations (27%, 70% and 55% at 25 years of age, respectively). Our study demonstrates that heterozygous *RFX6* protein truncating variants are a new cause of MODY with reduced penetrance.

Finding the genetic cause of rare familial diabetes (monogenic diabetes) is important as it provides new biological insights into human pancreas development and function, as well as potentially providing novel therapeutic targets with important treatment implications[1]. Maturity-Onset Diabetes of Young (MODY) is a sub-type of monogenic diabetes which usually presents under the age of 25 years. Typically, patients are non-obese, non-insulin dependent and have an autosomal dominant inheritance of diabetes[2]. Mutations in *HNF1A*, *HNF4A* and *GCK* are the commonest causes of MODY responsible for ~60 % of MODY aetiology[1]. Up to 11% of MODY patients lack mutations in the known MODY genes suggesting more genetic aetiologies are still to be discovered[1].

It is difficult to discover new genetic causes of MODY. There has been limited recent success in finding new MODY genes. *WFS1* heterozygous variants were recently shown to be a rare cause of MODY in Finnish families[3] and loss-of-function variants in the *APPL1* gene have recently been reported to cause MODY in two families[4]. The reason for this limited success is the difficulty of distinguishing monogenic diabetes patients from those with type 1 diabetes[5,6], or from the increasing number of patients with early-onset type 2 diabetes due to rising rates of obesity. Another important reason is the lack of large pedigrees with an autosomal dominant pattern of inheritance of diabetes that do not have an already known genetic cause. This prevents us from performing classical linkage analysis which was used to discover the most common forms of MODY such as *GCK*, *HNF1A* and *HNF4A*[7-10].

Rare-variant association testing is an important step to confirm the pathogenicity of novel variants in monogenic disease[11]. Rare-variant association testing particularly for comparing the frequency of novel protein-truncating variants (PTVs) in monogenic cases with unknown aetiology to the frequency in large control cohorts is now possible because of the availability of resources such as ExAC – a database of protein coding variants in large control populations[12]. This allows burden testing of the frequency of novel or rare coding variants in diseases of interest and a comparison to rates in controls to identify new genetic causes of monogenic disease.

In this study, we undertook next-generation sequencing of MODY cases with unknown aetiology and compared the frequency of protein truncating variants to large publicly available control cohorts to identify new MODY genes. Our study identified heterozygous *RFX6* protein truncating variants as a new cause of MODY.

Results

Heterozygous *RFX6* protein-truncating variants were identified in MODY patients with unknown aetiology. To identify patients with novel heterozygous protein truncating variants causing MODY, we first assessed 38 European (non-Finnish) probands with a strong MODY-like phenotype who did not have mutations in the common MODY genes (*GCK*, *HNF1A*, *HNF4A*) when assessed by Sanger sequencing (**Supplementary Table 1)**. To exclude the other known/less common causes of monogenic diabetes, these patients underwent comprehensive targeted-next generation sequencing (NGS) for all 29 known monogenic diabetes genes, including the genes for neonatal diabetes, MODY and mitochondrial

diabetes, lipodystrophy or other forms of syndromic diabetes)[13] (**Supplementary Table 2**). We identified two probands with mutations in the known MODY gene *HNF1B*[13,14]. The analysis of heterozygous PTVs in the 29 genes on the targeted panel identified two unrelated probands with a novel heterozygous nonsense variant in *Regulatory Factor X 6* (*RFX6*) (Family 1 - p.leu292Ter, Family 2 - p.lys351Ter) (**Table 1**)(**Fig. 1**)(**Supplementary Fig. 1**). We did not identify any rare (<1%) missense *RFX6* variants in this cohort. *RFX6* was part of the targeted sequencing panel because recessive *RFX6* variants (missense and/or protein-truncating) are a known cause of syndromic neonatal diabetes[15], but heterozygotes were not previously known to have any phenotype.

Heterozygous *RFX6* protein-truncating variants were enriched in a MODY discovery cohort compared to population controls from ExAC. We compared the frequency of *RFX6* PTVs in our discovery cohort to a large control population from ExAC[12]. Neither of the *RFX6* variants from the discovery cohort were present in the 60,706 individuals in ExAC. There were 15 individuals with *RFX6* PTV in the 33,346 ExAC Non-Finnish European control population (**Supplementary Table 3).** The frequency of the *RFX6* PTVs in the MODY discovery cohort was significantly higher than the ExAC Non-Finnish European control population (5.5% vs 0.045%, Odds ratio=130, 95% CI 28.7-593.6, $P=6x10^{-28}$) (**Table 1**).

Heterozygous *RFX6* protein-truncating variants are associated with MODY in a Non-Finnish European replication cohort. To replicate the findings of our discovery cohort, we then examined 348 non-Finnish European probands who were routinely referred for MODY genetic testing to the Molecular Genetics Laboratory, Exeter, UK and in whom the common causes of MODY were excluded using targeted-NGS assay (**Supplementary Table 1**). The analysis of heterozygous PTVs identified four unrelated probands with two novel *RFX6* nonsense variants (p.Gln25Ter, p.Arg377Ter) (**Supplementary Table 1**)(**Supplementary Fig. 1**). Similarly to the discovery cohort, the MODY replication cohort was enriched for *RFX6* PTVs compared to the ExAC non-Finnish European control population (1.15% vs 0.045%, Odds ratio=26, 95% CI 8.5- 78.2, $P=6x10^{-14}$) (**Table 1**). *RFX6* is well covered in EXAC, except exon 1 where the median coverage is 18X compared to 86X for the remaining exons (http://exac.broadinstitute.org/gene/ENSG00000185002). If we exclude exon 1 from the analysis the association of *RFX6* PTVs with MODY in the replication cohort is maintained ($P=0.0014$). None of the monogenic diabetes patients with mutations in the known genes had *RFX6* PTVs (n=131).

Finnish individuals had ~10-fold higher frequency of *RFX6* PTVs compared to Non-Finnish Europeans due to a founder effect. The ExAC database showed a relative abundance of *RFX6* PTVs in Finnish-European (15/3305, 0.45%) compared to non-Finnish European (15/33,346, 0.045%) (**Supplementary Table 3**). All of the Finnish individuals in ExAC with *RFX6* PTV had the same variant, p.His293Leufs. To further validate this finding in a larger Finnish control population, we analysed *RFX6* PTVs in 7040 control individuals from the METSIM study[16]. There were 26 individuals with *RFX6* PTVs in this cohort and all had the p.His293Leufs variant. The frequency of p.His293Leufs was not significantly different from the ExAC Finnish population frequency (0.37% vs 0.45%, $P=0.63$) (**Supplementary Table 3**). The METSIM study has contributed to the ExAC Finnish cohort so to prevent duplication, we used the data from the larger METSIM study for further analysis[12].

The *RFX6* p.His293Leufs is strongly enriched in Finnish MODY patients with unknown aetiology. To assess whether the p.His293Leufs variant is a cause of MODY in Finnish patients, we genotyped the *RFX6* p.His293Leufs variant in 80 Finnish probands who were routinely referred for MODY genetic testing to Genome Center of Eastern Finland, University of Eastern Finland and did not have mutations in the most common MODY genes (*GCK*, *HNF1A*, *HNF4A* and *HNF1B*) (**Supplementary Table 1**). We identified six probands with the p.His293Leufs variant. The frequency of this variant was significantly higher in the Finnish MODY cohort compared to the METSIM controls (7.5% vs 0.37%, Odds ratio=22, 95% CI 8.7-54.7, P=5x10^{-18}) (**Table 1**).

***RFX6* PTVs co-segregation with diabetes is consistent with reduced penetrance.** To further assess the causality of *RFX6* PTVs, we conducted co-segregation analysis in families with genetic data available on more than three affected individuals. We had one family with >3 affected individuals with genetic data (**Fig. 1** and **Supplementary Fig. 1**). The analysis in this large 4 generation pedigree (Family 1) showed that the *RFX6* variant p.Leu292Ter co-segregated in 9 out of 10 individuals with diabetes (maximum LOD score = 2.82). One individual without the *RFX6* variant had diabetes which is likely to be a phenocopy of type 2 diabetes considering the large pedigree, age of diagnosis and obesity (51 years, BMI 30 kg/m^2). There were two family members with an *RFX6* variant but with normal HbA1c level at the time of study (18 and 57 years) suggesting that *RFX6* PTVs may have reduced penetrance.

***RFX6* PTVs showed reduced penetrance compared to *HNF1A* and *HNF4A* MODY.** To assess the penetrance of *RFX6* PTVs for diabetes compared to common causes of MODY, we combined data for all six non-Finnish European proband families. There were 18 *RFX6* heterozygotes of whom five had not developed diabetes at study entry. 27% (95% CI 11-58) developed diabetes by the age of 25 years and 78% (95% CI 55-95) by 51 years (**Fig. 2**). Two out of six probands did not have affected parents at study entry (**Supplementary Fig. 1**). The penetrance of diabetes for *RFX6* heterozygotes was substantially reduced compared to pathogenic variants of *HNF1A* (70%, 95% CI 67-72 by the age of 25 years and 97%, 95% CI 96-98 by 50 years) and moderately lower than pathogenic variants of *HNF4A* (55%, 95% CI 50-60 by the age of 25 years and 91%, 95% CI 88-94 by 50 years) (**Fig. 2**). Similar to European proband families, the Finnish founder *RFX6* p.His293Leufs variant also showed reduced penetrance in Finnish families (**Supplementary Fig. 2**). In a previously reported family with a homozygous p.Arg181Gln *RFX6* neonatal diabetes child [15,17], the limited genetic information available on *RFX6* heterozygous family members was also suggestive of reduced penetrance of diabetes (**Supplementary Fig. 3**).

***RFX6* PTVs are not associated with type 2 diabetes.** The reduced penetrance and later age of onset of diabetes with *RFX6* PTVs raised the possibility that these variants may be associated with type 2 diabetes. To assess this, we used freely available data from the Type 2 diabetes Knowledge Portal which contains whole-exome data on type 2 diabetes patients[18]. Burden testing of *RFX6* PTVs for exome sequencing data from 8373 type 2 diabetes cases and 8466 controls showed no significant association with type 2 diabetes (0.14% vs 0.083%, Odds ratio=1.79, 95% CI 0.7-4.57, P=0.22)[18](**Supplementary Table 2**).

Phenotype of *RFX6* heterozygotes with diabetes. The median age of diagnosis of diabetes in 13 non-Finnish European heterozygotes was 36 years (IQR 21-41, range 13-64 years) (**Table 2**). They had a median 15 years (IQR 11-24) of diabetes with median BMI of 24 kg/m^2 (IQR 23-26). Significant endogenous insulin was present in 10/11 patients at recruitment. 69% (n=9) of patients were treated with insulin (**Table 2**). There was no history of sulphonylurea sensitivity and they did not have islet autoantibodies (GAD/IA2). All patients had isolated diabetes and there were no reports of the other features of homozygous *RFX6* mutations, duodenal or gall bladder atresia. The Finnish *RFX6* p.His293Leufs heterozygotes with diabetes (n=8) had similar phenotypes (**Supplementary Table 4**).

Discussion

Heterozygous *RFX6* protein-truncating variants are a new cause of MODY. The *RFX6* PTVs are predicted to be pathogenic according to the published guildlines[11,19]. We identified these variants in unrelated MODY patients in whom the known causes of monogenic diabetes had been excluded. These variants were enriched in Finnish and non-Finnish European MODY probands. They were rare in the respective control cohorts and in patients with type 2 diabetes. We observed co-segregation within pedigrees albeit with reduced penetrance. Finally, these variants are likely to have a functional effect due nonsense mediated decay causing haploinsufficiency[12,20]. This is further supported by the studies that showed that the homozygous *RFX6* PTVs cause neonatal diabetes[15,21,22]. In line with higher frequency in the control Finnish population, *RFX6* PTVs were ~7 fold higher in Finnish MODY cases compared to European MODY cases. They were responsible for 7.5% MODY cases compared to only ~1% of MODY European cases. This suggests that *RFX6* heterozygous PTVs are a relatively common cause of MODY in the Finnish population.

Large-scale control cohorts such as ExAC in combination with next-generation sequencing of well-characterised cases is a useful strategy for identifying new causes of monogenic disease. The very large ExAC database now provides sufficient power for reliable burden testing of rare variants in monogenic disease [12]. This is a particularly useful strategy for genetically heterogenous monogenic diseases as it does not require large pedigrees or linkage analysis. *RFX6* PTVs were highly enriched in both discovery and replication cohorts compared to control cohorts, supporting their pathogenicity. The ExAC database has been very useful in identifying benign variants because of an unusually high frequency in the population compared to frequency of the disease in question[11,12,23]. However, caution is required for reduced penetrance variants as their frequency can be higher than estimated disease frequency in the general population. *RFX6* PTVs are an example where the reduced penetrance explains the higher frequency in control cohorts compared to the estimated frequency of MODY (0.01%) in the general population[24].

Our study highlights the importance of population specific control and disease cohorts. The frequency of *RFX6* PTVs is ~10-fold higher in the Finnish population compared to non-Finnish populations due to a founder effect. It is ~40-fold higher than the estimated MODY frequency in the European population [24]. Based on this data, *RFX6* PTVs may be misclassified as benign variants due to this unusually high frequency in control population. However, using a population specific disease cohort, we showed that *RFX6* PTVs were enriched in Finnish MODY patients compared to a Finnish control population. In addition, evidence of

the functional impact of this variant (a Finnish neonatal diabetes patient with homozygous p.His293Leufs[21]) strongly supports its causality. The higher frequency of PTVs in genetic bottleneck population is not unexpected. It has been shown that the Finnish population has a higher burden of genome-wide PTVs (0.5-5%) compared to non-Finnish Europeans [25] and some of these have been associated with disease[25,26].

***RFX6* protein-truncating variants cause reduced penetrance MODY.** The *RFX6* PTV heterozygotes showed reduced penetrance for diabetes compared to the more common *HNF1A* and *HNF4A* MODY. Only 27% of heterozygotes developed diabetes by age 25 years and 78% by age 51 years. This reduced penetrance explains the lack of complete co-segregation in the *RFX6* pedigrees. It also clarifies why diabetes was only reported in the parents or grandparents (obligate heterozygous for functional *RFX6* variants) in 7 out of the 12 published recessive *RFX6* neonatal diabetes pedigrees [15,17,22,27-33]. The lack of enrichment of *RFX6* PTVs in type 2 diabetes patients compared to control patients further supports *RFX6* as cause of a reduced penetrance MODY rather than type 2 diabetes. Our study may have overestimated the penetrance of *RFX6* PTVs due to ascertainment bias as MODY genetic testing is commonly considered for patients with a parental history of diabetes. There are other examples of reduced-penetrance MODY mutations such as the recently published *HNF4A* p.R114W variant[34]. Homozygous PTVs in transcription factors such as *NEUROD1* and *PDX1* cause neonatal diabetes and, similar to *RFX6*, heterozygous PTVs in these two genes have been shown to cause reduced penetrance MODY[35,36].

There are differences as well as similarities between *RFX6* MODY and *HNF1A*/*HNF4A* MODY. In contrast to *HNF1A* and *HNF4A* MODY patients, *RFX6* MODY patients do not show enhanced sensitivity to sulphonylureas[37] and have redeuced penetrance. All patients with *HNF1A*/*HNF4A* MODY have significant endogenous insulin more than 3-5 years post diagnosis[38]. *RFX6* MODY patients showed a similar pattern. However we had one patient who did not have significant endogenous insulin. Similar to *HNF1A*/*HNF4A* MODY subtypes [6,37], *RFX6* MODY patients also lack islet autoantibodies and have isolated diabetes. This suggests that the presence of persistent C-peptide and lack of islet autoantibodies, which is currently used to distinguish common forms of MODY from type 1 diabetes, can also be used to identify *RFX6* MODY. However, a similar strategy will not help to distinguish late onset *RFX6* MODY from type 2 diabetes. Further studies are needed to understand the pathophysiology of *RFX6* MODY.

Our study supports the role of *RFX6* in the adult human pancreas. *RFX6* is from a family of transcription factors that contains winged-helix DNA binding domains[15]. *RFX6* is expressed almost exclusively in the pancreatic islets, small intestine and colon[15]. It acts downstream of *NEUROG3*, regulates islet cell differentiation and the development of the endocrine pancreas [15]. The homozygous *RFX6* missense and protein-truncating variants cause syndromic neonatal diabetes (gall bladder aplasia, gut atresia and diabetes)[15]. *RFX6* whole-body null mice show phenotypes consistent with human disease and die soon after birth, but the heterozygous whole-body *RFX6* mouse has not been reported to develop diabetes[15]. This is not surprising considering the lack of phenotype in heterozygous null mice of *HNF1A* and *HNF1B*[39]. Interestingly, the defect in glucose induced insulin secretion was present in the models that are more akin to haploinsufficiency of *RFX6* in adult humans[30,40]. 80% depletion of the RFX6 protein in the adult mouse pancreas *in vivo* as well as in human beta

cells *in vitro* showed that this defect was due to reduced expression of ABCC8, GCK and Ca^{2+} channels in beta cells and disruption of calcium mediated insulin secretion[30,40]. These data support the role of RFX6 in adult beta cells and may explain the cause of diabetes in our patients.

In conclusion, heterozygous *RFX6* protein-truncating variants are a novel cause of reduced penetrance MODY.

Methods

Study population:

Discovery MODY cohort: The discovery cohort comprises of 38 European probands with strong MODY-like phenotype who did not have mutations in the three most common MODY genes (*GCK*, *HNF1A* and *HNF4A*) (**Supplementary Table 1**). They were diagnosed <25 years of age, non-obese, had ≥ 3 generation history of diabetes, non-insulin treated or insulin treated with C-peptide> 200 pmol/L (if available), and lacked islet autoantibodies.

Non-Finnish European replication MODY cohort: The replication cohort was derived from 474 non-Finnish European routine MODY diagnostic referrals to the Molecular Genetic Laboratory, Exeter, UK. A monogenic aetiology in a known monogenic diabetes gene was identified in 131 patients and the remaining 343 patients with unknown aetiology comprised the replication cohort (**Supplementary Table 1**).

Finnish-European MODY cohort: This cohort consisted of 80 patients who were routinely referred for MODY diagnostic testing to the Genome Center of Eastern Finland, University of eastern Finland in whom no mutation was found in the common MODY genes (*GCK*, *HNF1A*, *HNF1B* and *HNF4A*) when assessed by Sanger sequencing (**Supplementary Table 1**). Two of the identified patients with *RFX6* mutations participated in the FINNMODY Study (www.botnia-study.org/finnmody) recruiting patients with MODY-like diabetes and their relatives in Finland, and their families were studied.

Finnish-European control cohort: Individuals of this cohort were part of the METSIM study (n=7040). They were all males aged 45 to 70 years, randomly selected from the population register of the Kuopio town, Eastern Finland and have been described previously[16].

Cohort of people with pathogenic HNF1A and HNF4A variants: This cohort included probands and their family members referred to the Molecular Genetics Laboratory, Exeter, UK for MODY genetic testing and were identified to have a pathogenic *HNF1A* (n=1265) or *HNF4A* (n=427) variant.

DNA analysis:

Targeted-next generation sequencing:

The analysis of all known monogenic diabetes genes in European cohorts was conducted using targeted-next generation sequencing as described previously[13]. The panel included 29 genes in which variants are known to cause monogenic neonatal diabetes, MODY, mitochondrial diabetes, lipodystrophy or other forms of syndromic diabetes[13]

(**Supplementary Table 2**). The *RFX6* PTVs identified by targeted-NGS were confirmed using a gold-standard Sanger sequencing. The essential splice site, nonsense and frameshift variants excluding the last exon were considered protein-truncating variants in this study[12,20]. The targeted-NGS assay covered 100% bases of the *RFX6* coding region >10x read depth for all the samples.

Sanger sequencing:
Genomic DNA was extracted from whole blood using standard procedures and the coding region and intron/exon boundaries of the *RFX6* gene were amplified by PCR. Amplicons were sequenced using the Big Dye Terminator Cycler Sequencing Kit v3.1 (Applied Biosystems, Warrington, UK) according to manufacturer's instructions and reactions were analysed on an ABI 3730 Capillary sequencer (Applied Biosystems, Warrington, UK). Sequences were compared with the reference sequences (NM_173560.3) using Mutation Surveyor v3.24 software (So Genetics, State College, PA).

The Finnish-European MODY cohort was analysed for p.His293Leufs variant using Sanger sequencing as described above. Family co-segregation analysis was performed in available family members using a Sanger sequencing assay for the specific *RFX6* variant identified in that family.

DNA analysis of the METSIM study has been described before [21].

The DNA analysis of the ExAC cohort has been described previously[12]. >95% bases of the *RFX6* coding region was covered with >10x read depth for >90% of samples. The lowest coverage was seen in exon 1 with only 80% of bases having >10x coverage in >50% of samples. We only used high quality filtered variants for the analysis.

Ethics:
The study is approved by the North Wales Research Ethics Committee, UK. The FINNMODY study is approved by the Research Ethics committee for Medicine of the Helsinki University Hospital.

Statistical analysis:
The chi-squared test was used to compare the frequency of *RFX6* PTVs. The penetrance of diabetes was assessed using survival time analysis method. The statistical analysis was conducted using Stata 14 (StataCorp, Texas, USA). Singlepoint non-parametric linkage analyses were performed using MERLIN 1.1.2 [41]. The Z score was converted into a LOD score by use of the Kong and Cox exponential model implemented in MERLIN[41,42].

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

Author Contributions. K.A.P and M.N.W. researched data and performed statistical analyses. K.A.P. wrote the first draft of the manuscript, which was modified by all authors. All authors contributed to the discussion and reviewed or edited the manuscript. M.N.W. is the

guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Acknowledgement:
We thank Drs Marc Abramowicz and Julie Desir for sharing genetic information on *RFX6* neonatal diabetes family from Brussels.

Funding: K.A.P. has a postdoctoral fellowship funded by the Wellcome Trust (110082/Z/15/Z). S.E.F. has a Sir Henry Dale Fellow-ship jointly funded by the Wellcome Trust and the Royal Society (105636/Z/14/Z). S.E. and A.T.H. are Wellcome Trust Senior Investigators (WT098395/Z/12/Z), and A.T.H. is also supported by a NIHR Senior Investigator award. M.N.W. is supported by the Wellcome Trust Institutional Strategic Support Fund (WT097835MF) and the Medical Research Council (MR/M005070/1). MS is supported by the NIHR Exeter Clinical Research Facility. Additional support came from the University of Exeter and the NIHR Exeter Clinical Research Facility. MC is supported by the Fonds National de la Recherche Scientifique (FNRS), Belgium. The FINNMODY Study has been financially supported by the Folkhalsan Research Foundation, Helsinki University Central Hospital Research Foundation, Finnish Diabetes Research Foundation, Finnish Medical Society, Foundation for Pediatric Research, Ahokas Foundation, Diabetes Wellness Foundation. The views expressed are those of the authors and not necessarily those of the Wellcome Trust, the National Health Service, the NIHR, or the Department of Health.

References:

1. McCarthy, M. I. & Hattersley, A. T. Learning from molecular genetics: novel insights arising from the definition of genes for monogenic and type 2 diabetes. *Diabetes* **57,** 2889–2898 (2008).
2. Murphy, R., Turnbull, D. M., Walker, M. & Hattersley, A. T. Clinical features, diagnosis and management of maternally inherited diabetes and deafness (MIDD) associated with the 3243A>G mitochondrial point mutation. *Diabet. Med.* **25,** 383–399 (2008).
3. Bonnycastle, L. L. *et al.* Autosomal dominant diabetes arising from a Wolfram syndrome 1 mutation. *Diabetes* **62,** 3943–3950 (2013).
4. Prudente, S. *et al.* Loss-of-Function Mutations in APPL1 in Familial Diabetes Mellitus. *Am. J. Hum. Genet.* **97,** 177–185 (2015).
5. Oram, R. A. *et al.* Most people with long-duration type 1 diabetes in a large population-based study are insulin microsecretors. *Diabetes Care* **38,** 323–328 (2015).
6. McDonald, T. J. *et al.* Islet autoantibodies can discriminate maturity-onset diabetes of the young (MODY) from Type 1 diabetes. *Diabet. Med.* **28,** 1028–1033 (2011).
7. Froguel, P. *et al.* Close linkage of glucokinase locus on chromosome 7p to early-onset non-insulin-dependent diabetes mellitus. *Nature* **356,** 162–164 (1992).
8. Hattersley, A. T., Turner, R. C., Patel, P. & O'Rahilly, S. Linkage of type 2 diabetes to the glucokinase gene. *The Lancet* **339,** 1307–1310 (1992).
9. Yamagata, K. *et al.* Mutations in the hepatocyte nuclear factor-1alpha gene in maturity-onset diabetes of the young (MODY3). *Nature* **384,** 455–458 (1996).
10. Yamagata, K. *et al.* Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young (MODY1). *Nature* **384,** 458–460 (1996).

11. MacArthur, D. G. *et al.* Guidelines for investigating causality of sequence variants in human disease. *Nature* **508,** 469–476 (2014).
12. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536,** 285–291 (2016).
13. Ellard, S. *et al.* Improved genetic testing for monogenic diabetes using targeted next-generation sequencing. *Diabetologia* **56,** 1958–1963 (2013).
14. Shepherd, M. *et al.* Systematic Population Screening, Using Biomarkers and Genetic Testing, Identifies 2.5% of the U.K. Pediatric Diabetes Population With Monogenic Diabetes. *Diabetes Care* **39,** 1879–1888 (2016).
15. Smith, S. B. *et al.* Rfx6 directs islet formation and insulin production in mice and humans. *Nature* **463,** 775–780 (2010).
16. Stančáková, A. *et al.* Changes in insulin sensitivity and insulin release in relation to glycemia and glucose tolerance in 6,414 Finnish men. *Diabetes* **58,** 1212–1221 (2009).
17. Martinovici, D. *et al.* Neonatal hemochromatosis and Martinez-Frias syndrome of intestinal atresia and diabetes mellitus in a consanguineous newborn. *European Journal of Medical Genetics* **53,** 25–28 (2010).
18. *AMP-T2D Program; T2D-GENES Consortium, SIGMA T2D Consortium. RFX6.* (type2diabetesgenetics.org). Available at: http://www.type2diabetesgenetics.org/gene/geneInfo/RFX6#. (Accessed: 25 November 2016)
19. Richards, S. *et al.* Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. in **17,** 405–424 (2015).
20. MacArthur, D. G. *et al.* A systematic survey of loss-of-function variants in human protein-coding genes. *Science* **335,** 823–828 (2012).
21. Huopio, H. *et al.* Clinical, Genetic, and Biochemical Characteristics of Early-Onset Diabetes in the Finnish Population. *The Journal of Clinical Endocrinology & Metabolism* **101,** 3018–3026 (2016).
22. Skopkova, M. *et al.* Two novel RFX6 variants in siblings with Mitchell-Riley syndrome with later diabetes onset and heterotopic gastric mucosa. *European Journal of Medical Genetics* **59,** 429–435 (2016).
23. Manrai, A. K. *et al.* Genetic Misdiagnoses and the Potential for Health Disparities. *N. Engl. J. Med.* **375,** 655–665 (2016).
24. Shields, B. M. *et al.* Maturity-onset diabetes of the young (MODY): how many cases are we missing? *Diabetologia* **53,** 2504–2508 (2010).
25. Lim, E. T. *et al.* Distribution and medical impact of loss-of-function variants in the Finnish founder population. *PLoS Genet* **10,** e1004494 (2014).
26. Akinrinade, O., Koskenvuo, J. W. & Alastalo, T.-P. Prevalence of Titin Truncating Variants in General Population. *PLoS ONE* **10,** e0145284 (2015).
27. Concepcion, J. P. *et al.* Neonatal diabetes, gallbladder agenesis, duodenal atresia, and intestinal malrotation caused by a novel homozygous mutation in RFX6. *Pediatr Diabetes* **15,** 67–72 (2014).
28. Spiegel, R. *et al.* Clinical characterization of a newly described neonatal diabetes syndrome caused by RFX6 mutations. *Am. J. Med. Genet.* **155A,** 2821–2825 (2011).
29. Chappell, L. *et al.* A further example of a distinctive autosomal recessive syndrome comprising neonatal diabetes mellitus, intestinal atresias and gall bladder agenesis.

Am. J. Med. Genet. **146A,** 1713–1717 (2008).

30. Chandra, V. *et al.* RFX6 regulates insulin secretion by modulating Ca2+ homeostasis in human β cells. *CellReports* **9,** 2206–2218 (2014).
31. Sansbury, F. H. *et al.* Biallelic RFX6 mutations can cause childhood as well as neonatal onset diabetes mellitus. *Eur. J. Hum. Genet.* **23,** 1744–1748 (2015).
32. Zegre Amorim, M. *et al.* Mitchell-Riley Syndrome: A Novel Mutation in RFX6 Gene. *Case Rep Genet* **2015,** 937201 (2015).
33. Mitchell, J. *et al.* Neonatal diabetes, with hypoplastic pancreas, intestinal atresia and gall bladder hypoplasia: search for the aetiology of a new autosomal recessive syndrome. *Diabetologia* **47,** 2160–2167 (2004).
34. Laver, T. W. *et al.* The Common p.R114W HNF4A Mutation Causes a Distinct Clinical Subtype of Monogenic Diabetes. *Diabetes* **65,** 3212–3217 (2016).
35. Malecki, M. T. *et al.* Mutations in NEUROD1 are associated with the development of type 2 diabetes mellitus. *Nature Genetics* **23,** 323–328 (1999).
36. Stoffers, D. A., Ferrer, J., Clarke, W. L. & Habener, J. F. Early-onset type-II diabetes mellitus (MODY4) linked to IPF1. *Nature Genetics* **17,** 138–139 (1997).
37. Murphy, R., Ellard, S. & Hattersley, A. T. Clinical implications of a molecular genetic classification of monogenic beta-cell diabetes. *Nat Clin Pract Endocrinol Metab* **4,** 200–213 (2008).
38. Besser, R. E. J. *et al.* Urinary C-Peptide Creatinine Ratio Is a Practical Outpatient Tool for Identifying Hepatocyte Nuclear Factor 1- /Hepatocyte Nuclear Factor 4- Maturity-Onset Diabetes of the Young From Long-Duration Type 1 Diabetes. *Diabetes Care* **34,** 286–291 (2011).
39. Maestro, M. A. *et al.* Distinct roles of HNF1beta, HNF1alpha, and HNF4alpha in regulating pancreas development, beta-cell function and growth. *Endocr Dev* **12,** 33–45 (2016).
40. Piccand, J. *et al.* Rfx6 maintains the functional identity of adult pancreatic β cells. *CellReports* **9,** 2219–2232 (2014).
41. Abecasis, G. R., Cherny, S. S., Cookson, W. O. & Cardon, L. R. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nature Genetics* **30,** 97–101 (2002).
42. Kong, A. & Cox, N. J. Allele-sharing models: LOD scores and accurate linkage tests. *Am. J. Hum. Genet.* **61,** 1179–1188 (1997).

Table 1: Frequency of heterozygous *RFX6* protein-truncating variants in all study cohorts and control populations.

	MODY cohorts, Frequency of *RFX6* PTV	Control population, Frequency of *RFX6* PTV	Odds ratio (95% CI)	*P*
Non-Finnish European	Discovery cohort (n=2/36), 5.5%	ExAC control population (n=15/33346), 0.045%	130 (28.7-593.6)	$6x10^{-28}$
	Replication cohort (n=4/348), 1.15%		26 (8.5-78.2)	$6x10^{-14}$
Finnish European	Replication cohort (n=6/80), 7.5%	METSIM control population (n=26/7040), 0.37%	22 (8.7-54.7)	$5x10^{-18}$

Fig. 1: Extended pedigree of family 1 showing inheritance of heterozygous *RFX6* p.Leu292Ter (NM_173560.3:c.875T>G) variant. Genotype is shown underneath each symbol; M and N denote mutant and wild-type alleles, respectively. Directly below the genotype is the age of diabetes onset in years, duration in years, BMI and treatment at study entry. Squares represent male family members, and circles represent female members. Black-filled symbols denote patients with diabetes. An arrow denotes the proband in the family. OHA, oral hypoglycaemic agents. *age at recruitment

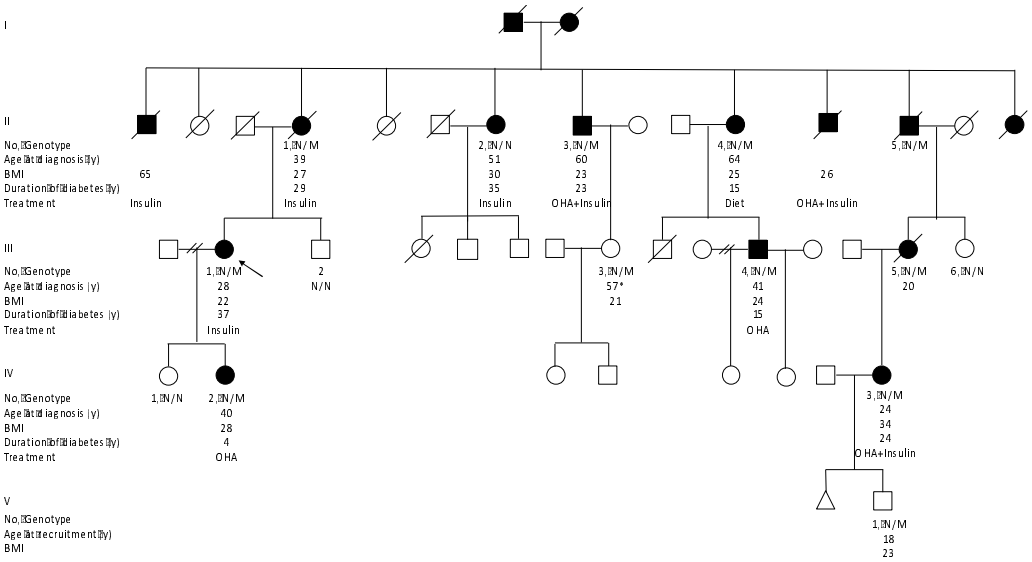

Fig. 2: Penetrance of diabetes in people with a heterozygous *RFX6* protein-truncating variant (n=18), pathogenic *HNF1A* variant (n=1265) or *HNF4A* variant (n=427).

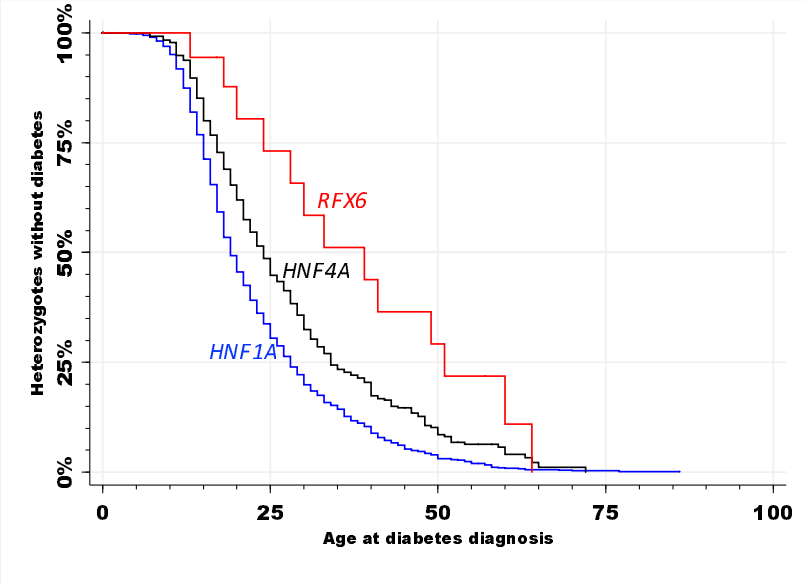

Table 2: Clinical characteristics of non-Finnish European patients with diabetes and *RFX6* heterozygous protein-truncating variant (n=13)

Characteristic	
Age at diagnosis (years), median (IQR), n=13	36 (21-41)
Duration of diabetes, median (IQR), n=13	15 (11-24)
Female, n (%), n=13	11 (85%)
BMI (kg/m^2), median (IQR), n=12	24 (23-26)
Initial treatment, n (%), n=13	
Diet	3 (23%)
Oral hypoglycaemic agents	6 (46%)
Insulin	3 (23%)
Insulin + Oral hypoglycaemic agents	1 (8%)
Current treatment, n (%), n=13	
Diet	1 (8%)
Oral hypoglycaemic agents	3 (23%)
Insulin	5 (38%)
Insulin + Oral hypoglycaemic agents	4 (31%)
HbA1C at recruitment, mmol/mol, median (IQR), n=11	52 (43-83)
Significant endogenous insulin at recruitment*, n (%), n=11	10 (91%)

* noninsulin treated or insulin treated with urine/blood random C-peptide >200 pmol/l.